Convex Analysis and Variational Problems

Ivar Ekeland
Roger Témam

C · L · A · S · S · I · C · S
In Applied Mathematics

siam
28

Convex Analysis and Variational Problems

SIAM's Classics in Applied Mathematics series consists of books that were previously allowed to go out of print. These books are republished by SIAM as a professional service because they continue to be important resources for mathematical scientists.

Classics in Applied Mathematics

C. C. Lin and L. A. Segel, *Mathematics Applied to Deterministic Problems in the Natural Sciences*
Johan G. F. Belinfante and Bernard Kolman, *A Survey of Lie Groups and Lie Algebras with Applications and Computational Methods*
James M. Ortega, *Numerical Analysis: A Second Course*
Anthony V. Fiacco and Garth P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*
F. H. Clarke, *Optimization and Nonsmooth Analysis*
George F. Carrier and Carl E. Pearson, *Ordinary Differential Equations*
Leo Breiman, *Probability*
R. Bellman and G. M. Wing, *An Introduction to Invariant Imbedding*
Abraham Berman and Robert J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*
Olvi L. Mangasarian, *Nonlinear Programming*
*Carl Friedrich Gauss, *Theory of the Combination of Observations Least Subject to Errors: Part One, Part Two, Supplement.* Translated by G. W. Stewart
Richard Bellman, *Introduction to Matrix Analysis*
U. M. Ascher, R. M. M. Mattheij, and R. D. Russell, *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*
K. E. Brenan, S. L. Campbell, and L. R. Petzold, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*
Charles L. Lawson and Richard J. Hanson, *Solving Least Squares Problems*
J. E. Dennis, Jr. and Robert B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*
Richard E. Barlow and Frank Proschan, *Mathematical Theory of Reliability*
Cornelius Lanczos, *Linear Differential Operators*
Richard Bellman, *Introduction to Matrix Analysis, Second Edition*
Beresford N. Parlett, *The Symmetric Eigenvalue Problem*

*First time in print.

Classics in Applied Mathematics (continued)

Richard Haberman, *Mathematical Models: Mechanical Vibrations, Population Dynamics, and Traffic Flow*
Peter W. M. John, *Statistical Design and Analysis of Experiments*
Tamer Başar and Geert Jan Olsder, *Dynamic Noncooperative Game Theory, Second Edition*
Emanuel Parzen, *Stochastic Processes*
Petar Kokotović, Hassan K. Khalil, and John O'Reilly, *Singular Perturbation Methods in Control: Analysis and Design*
Jean Dickinson Gibbons, Ingram Olkin, and Milton Sobel, *Selecting and Ordering Populations: A New Statistical Methodology*
James A. Murdock, *Perturbations: Theory and Methods*
Ivar Ekeland and Roger Témam, *Convex Analysis and Variational Problems*
Ivar Stakgold, *Boundary Value Problems of Mathematical Physics, Volumes I and II*
J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*
David Kinderlehrer and Guido Stampacchia, *An Introduction to Variational Inequalities and Their Applications*
F. Natterer, *The Mathematics of Computerized Tomography*
Avinash C. Kak and Malcolm Slaney, *Principles of Computerized Tomographic Imaging*
R. Wong, *Asymptotic Approximations of Integrals*
O. Axelsson and V. A. Barker, *Finite Element Solution of Boundary Value Problems: Theory and Computation*
David R. Brillinger, *Time Series: Data Analysis and Theory*
Joel N. Franklin, *Methods of Mathematical Economics: Linear and Nonlinear Programming, Fixed-Point Theorems*
Philip Hartman, *Ordinary Differential Equations, Second Edition*
Michael D. Intriligator, *Mathematical Optimization and Economic Theory*
Philippe G. Ciarlet, *The Finite Element Method for Elliptic Problems*
Jane K. Cullum and Ralph A. Willoughby, *Lanczos Algorithms for Large Symmetric Eigenvalue Computations, Vol. I: Theory*
M. Vidyasagar, *Nonlinear Systems Analysis, Second Edition*
Robert Mattheij and Jaap Molenaar, *Ordinary Differential Equations in Theory and Practice*
Shanti S. Gupta and S. Panchapakesan, *Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations*
Eugene L. Allgower and Kurt Georg, *Introduction to Numerical Continuation Methods*
Heinz-Otto Kreiss and Jens Lorenz, *Initial-Boundary Value Problems and the Navier-Stokes Equations*
J. L. Hodges, Jr. and E. L. Lehmann, *Basic Concepts of Probability and Statistics, Second Edition*

This page intentionally left blank

Convex Analysis and Variational Problems

Ivar Ekeland
Université Paris-Dauphine
Paris, France

Roger Témam
Université Paris-Sud
Orsay, France

Indiana University
Bloomington, Indiana

siam
Society for Industrial and Applied Mathematics
Philadelphia

This SIAM edition is an unabridged, corrected republication of the work first published in English by North-Holland and American Elsevier, Amsterdam and New York, 1976.

10 9 8 7 6 5 4 3 2

Library of Congress Cataloging-in-Publication Data

Ekeland, I. (Ivar), 1944-
[Analyse convexe et problèmes variationnels. English]
Convex analysis and variational problems / Ivar Ekeland, Roger Temam.
p. cm. – (Classics in applied mathematics ; 28)
English language ed. originally published: Amsterdam : North-Holland Pub. Co.; New York : American Elsevier Pub. Co. [distributor], 1976, in series: Studies in mathematics and its applications ; v.1.
Includes bibliographical references and index.
ISBN 0-89871-450-8 (pbk.)
1. Mathematical optimization. 2. Convex functions. 3. Calculus of variations. I. Temam, Roger. II. Title. III. Series

QA402.5 .E3813 1999
519.3–dc21 99-046956

CONTENTS

This page intentionally left blank

PREFACE TO THE CLASSICS EDITION

This edition of the book is the same as the initial one, except for a few corrections and the addition of a small number of references.

Several parts of the book cover basic material which turns out to be useful in a number of applications and which is not expected to evolve; as far as we know this material has not appeared elsewhere in book form since it was published in this book.

Here are some of the topics developed in this book, and their present standing in theoretical and applied research; references are provided in the "Additional References" at the end of the book:

1. Duality in the calculus of variation (convex variational problems in infinite dimension).

   Duality has important applications in mathematical economy, in continuum mechanics, in numerical analysis (mixed finite elements), and in control theory. New developments have occurred in convex analysis in finite dimension: we refer, for instance, to semi-definite programming. Duality for some (finite or infinite dimensional) nonconvex problems has been developed. Systematic use of duality in solid mechanics for plasticity related problems has been made. The (infinite dimensional) nonconvex problems of calculus of variation appearing in nonlinear elasticity have attracted much attention and effort (with little or no reference to duality).

2. Generalized solutions of minimal surface problems. Important developments have occurred in the parametric case (not considered in this book) with geometrical ideas totally different from those used here. Extensions of the methods of this book have been developed and studied for the time dependent (evolution nonparametric) minimal surface problem.

3. The minmax theorems stated in this book have also many useful applications— in particular in relation with duality for the same topics as in point 1 and most recently for control theory, namely the robust control of partial differential equations in finite time horizon.

Ivar Ekeland

Roger Témam

This page intentionally left blank

PREFACE

In recent years, there has been a considerable expansion in the field of Convex Analysis, in conjunction with the development of various mathematical tools, certain of which have become standard. The initial motivation was provided by operations research: the success of linear programming, due to the duality theorem and the simplex method, has aroused the interest of managers and engineers in this type of problem, and similar results have therefore been sought for non-linear optimization. To this primary objective of Convex Analysis, others have been added, ranging from mathematical economics to mechanics, and encompassing more strictly mathematical problems such as the study of functional equations of monotone type. The result has been a deeper understanding of convex functions, together with the introduction of new concepts such as those of subdifferentiability and conjugate convex functions. The subdifferential of a convex function is a generalization of the notion of derivative, and has provided the theory of maximal monotone operators, so useful in the study of partial differential and integral equations, with its first examples. The concept of conjugate convex functions has emerged as an elegant and general formalization of duality in optimization.

However, it does not seem to us that these methods have been used to their full advantage in the study of variational problems, that is, optimization problems in concrete functional spaces. The object of this book is to fill this gap in two main directions:

by dualization of convex variational problems,

by relaxation of non-convex variational problems.

Duality allows us to associate a dual problem with a variational problem and to study the relationship between the two problems. This is useful in mathematical economics where the dual problem can be stated in terms of the price; in mechanics where the primal and the dual problems are two well-known forms of the conservation principles, characterizing the displacements and the constraints respectively; in numerical analysis where the dual problem may help us to solve the primal problem. In addition to these standard applications of duality, we have a new use for the calculus of variations in mind: the dual problem enables us to define the generalized solution of a variational problem which has no classical solution.

The relaxation consists in associating a "convexified" problem with a non-convex variational problem. This approach to non-convex problems has been considerably developed for optimal control problems; the study is developed here in a more general framework, adapted to problems in the calculus of variations. The solutions of the "convexified" problem arise as generalized solutions of the initial problem.

Thus we see that these two directions meet in the concept of generalized solutions: these are the cluster points of the minimizing sequences of the problem under consideration.

The book is divided into three parts, dealing with a summary of convex analysis, duality for convex variational problems, and the relaxation of non-convex variational problems respectively. We shall now describe the contents of the various chapters in more detail.

Chapter I summarizes the essentials of the theory of convex functions. We have omitted those points which are not directly useful to us, such as the inf-convolution, so that we can concentrate on the fundamental concepts of conjugate convex functions and of subdifferentiability.

Chapter II deals with the minimization of convex functions. Here we recall the principal results which guarantee the existence and uniqueness of the point where a convex function attains its minimum, characterizing it as the solution of a variational inequality.

Chapter III develops the theory of duality in convex optimization following R. T. Rockafellar. Given a convex optimization problem, we embed it in a family of perturbed problems, and by using conjugate convex functions we associate a dual problem with it. This very flexible abstract theory can be adapted to a wide variety of situations.

Chapter IV describes the application of duality to several problems in the calculus of variations, of mathematical physics, of mechanics and of filtering theory. In each case, we state the dual problem explicitly together with its relationship to the primal problem.

Chapter V describes the application of duality to the classical problem of minimal hypersurfaces and to problems of related type. The dual problem still has a unique solution, and the primal-dual relationship enables us to associate a generalized solution of the primal problem with it. Hence we obtain the existence of a generalized solution to the problem of minimal hypersurfaces, for which it is well known that there is generally no classical solution. In addition to the systematic use of duality, we here have an unexpected application of ε-subdifferentiability.

Chapter VI describes a different theory of duality, based on the minimax theorems. This approach, which is older than Rockafellar's, adds nothing new and we develop it briefly for completeness.

Chapter VII describes the application of duality to problems of numerical analysis, optimal control, mechanics and mathematical economics. All these applications of duality are still being developed at the present time and we do not pretend to be exhaustive. We have restricted ourselves to illustrating some typical methods with specific examples.

Chapter VIII tackles non-convex variational problems, by studying those cases where the existence of a classical solution is assured. The existence theorem which we obtain is illustrated by examples taken from optimal control and the calculus of variations.

Chapter IX considers variational problems devoid of a classical solution. We then define, by partial convexification, a relaxed problem, which can be shown to be near to the initial problem. In particular, the relaxed problem possesses classical solutions which are none other than the generalized solutions of the initial problem. These results are applied to a number of problems of optimal control and of the calculus of variations.

Chapter X deals separately with the fundamental problem of the calculus of variations in dimension $n > 1$. Although it is not amenable to the methods of the preceding chapter, the results obtained are similar: obtaining the relaxed problem by partial convexification, characterizing the classical solutions of the relaxed problem as generalized solutions of the initial problem. Finally, we conclude with a study of variational equations. Clearly, if a problem in the calculus of variations has no classical solution, the corresponding Euler equation has no solution in general. However, we show that approximate solutions always exist.

Chapters I to VII were the subject of a postgraduate course given by the second author at the Université de Paris XI in 1970–71 and 1971–72. They contain some new results and others which have only recently been published (see R. Temam [1]–[4]). Chapters VIII to X continue and develop previous work of the first author (see I. Ekeland [1]) and also contain large borrowings from the work of H. Berliocchi and J. M. Lasry (see [1]). Obviously all the first part of this book owes much to J. J. Moreau and R. T. Rockafellar. We offer them our thanks.

Our purpose was not to produce a systematic exposition of the topics considered here. We have only attempted to describe some methods linked with convex analysis, methods which have already been found to be productive and still seem to be promising. This should be the subject of future research.

We dedicate this book to J. L. Lions who has profoundly influenced our mathematical thinking and to whom we owe much.

We thank M. P. Lelong for welcoming our work into the series of which he is the general editor and also for his advice and suggestions.

Our thanks also go to Mme Cartier and Mme Maynard who typed the greater part of the manuscript.

Finally, we wish to thank Editions Dunod for their excellent typographic work

Paris, November 1973.

The English translation has been updated by incorporating recent work of the authors (Appendixes I and II) and their students. It has also benefited from the improvements suggested by the readers of the French edition.

PART ONE

Fundamentals of Convex Analysis

This page intentionally left blank

CHAPTER I

Convex Functions

Introduction

This chapter assumes a basic knowledge of topological vector spaces. Moreover, we shall recall in Section 1 several fundamental aspects of this theory which will be constantly used in what follows. These reminders are in no way systematic and are centred on the notion of a convex set. We go on to consider convex functions (Sections 2–4) and their differentiability (Sections 5 and 6). All the vector spaces studied here are real.

1. CONVEX SETS AND THEIR SEPARATION

1.1. Convex sets

Let V be a vector space over $\mathbf{R}$. If u and v are two points of V, u and v are called the endpoints of the line-segment denoted by $[u,v]$ where

$$[u, v] = \{ \lambda u + (1 - \lambda)v \mid 0 \leqslant \lambda \leqslant 1 \}.$$

A set $\mathscr{A} \subset V$ is said to be *convex* if and only if for every pair of elements (u,v) of $\mathscr{A}$ the segment $[u,v]$ is contained in V. We know that a set $\mathscr{A} \subset V$ is convex if and only if for every finite subset of elements $u_1, \ldots, u_n$ of $\mathscr{A}$, and for every family of real positive numbers $\lambda_1, \ldots, \lambda_n$ with sum unity, we have

$$\sum_{i=1}^{n} \lambda_i u_i \in \mathscr{A}.$$

The whole of the space V is convex and, conventionally, so is the empty set. Every intersection of convex sets is convex, but in general the union of convex sets is not convex.

If $\mathscr{A}$ is any subset of V, the intersection of all the sets containing $\mathscr{A}$ is a convex set, and it is the smallest convex set containing $\mathscr{A}$. It is called the *convex hull* of $\mathscr{A}$ and is denoted by co $\mathscr{A}$. It is also the set of all the convex combinations of the elements of $\mathscr{A}$, i.e.,

$$\operatorname{co} \mathscr{A} = \left\{ \sum_{i=1}^{n} \lambda_i u_i \,\middle|\, n \in \mathbf{N}, \sum_{i=1}^{n} \lambda_i = 1, \lambda_i \geqslant 0, u_i \in \mathscr{A}, 1 \leqslant i \leqslant n \right\}.$$

Let $\mathscr{H}$ be an affine hyperplane with equation $\ell(u) = \alpha$ where ℓ is a non-zero linear form on V and $\alpha \in \mathbf{R}$. The sets

$$\{ u \in V \mid \ell(u) < \alpha \}, \qquad \{ u \in V \mid \ell(u) > \alpha \},$$
$$\{ u \in V \mid \ell(u) \leqslant \alpha \}, \qquad \{ u \in V \mid \ell(u) \geqslant \alpha \}$$

are called respectively *open* and *closed* half-spaces bounded by $\mathscr{H}$. These are two convex sets which depend only on $\mathscr{H}$, and not on the particular ℓ and α chosen for its equation.

1.2. Separation of convex sets

We recall that a *topological vector space* (t.v.s.) is defined as a vector space V endowed with a topology for which the operations

$$(u, v) \mapsto u + v \qquad \text{of } V \times V \text{ into } V,$$
$$(\lambda, u) \mapsto \lambda u \qquad \text{of } \mathbf{R} \times V \text{ into } V$$

are continuous. The neighbourhoods of any point may then be deduced from those of the origin by translation. A t.v.s. is said to be a *locally convex space* (l.c.s.) if the origin possesses a fundamental system of convex neighbourhoods. This is the case with normed spaces: it is sufficient to take the set of neighbourhoods formed by the balls centred on the origin. All the usual t.v.s. encountered in analysis are locally convex.

Now let V be a t.v.s. and $\mathscr{H}$ an affine hyperplane with equation $\ell(u) = \alpha$ where ℓ is a non-zero linear functional on V and $\alpha \in \mathbf{R}$. It can be shown that the set $\mathscr{H}$ is topologically closed if and only if the function ℓ is continuous. Under these conditions the open (closed) half-spaces determined by $\mathscr{H}$ will be topologically open (closed).

In a t.v.s. V, the closure of a convex set is convex and the interior of a convex set is also convex (possibly empty). More generally, if $\mathscr{A} \subset V$ is convex, if $u \in \mathring{\mathscr{A}}$ (the interior of $\mathscr{A}$) and if $v \in \bar{\mathscr{A}}$ (the closure of $\mathscr{A}$), then $[u,v[\subset \mathring{\mathscr{A}}$, from which we deduce that $\bar{\mathring{\mathscr{A}}} = \bar{\mathscr{A}}$ whenever $\mathring{\mathscr{A}} \neq \varnothing$. This suggests to us the introduction of the following definition: a point $u \in \mathscr{A}$ will be called *internal* if every line passing through u meets $\mathscr{A}$ in a segment $[v_1, v_2]$ such that $u \in]v_1, v_2[$. Hence every interior point is internal, and by the above argument, if $\mathring{\mathscr{A}} \neq \varnothing$ every internal point is interior.

If $\mathscr{A}$ is any subspace of V, the intersection of all the closed convex subsets containing $\mathscr{A}$ is the smallest closed convex subset containing $\mathscr{A}$. It is also the closure of the convex hull of $\mathscr{A}$ (and not the convex hull of the closure!); it is called the closed convex hull of $\mathscr{A}$ and is denoted by $\overline{\text{co}}\, \mathscr{A}$.

An affine hyperplane $\mathscr{H}$ is said to *separate* (*strictly separate*) two sets $\mathscr{A}$ and $\mathscr{B}$ if each of the closed (open) half-spaces bounded by $\mathscr{H}$ contains one of them. This may be written analytically as follows. If $\ell(u) = \alpha$ is the equation of $\mathscr{H}$, then we have

$$\ell(u) \leqslant \alpha, \quad \forall u \in \mathscr{A}, \qquad \ell(v) \geqslant \alpha, \quad \forall v \in \mathscr{B},$$

for separation,

$$\ell(u) < \alpha, \ \forall u \in \mathscr{A}, \qquad \ell(v) > \alpha, \ \forall v \in \mathscr{B}.$$

for strict separation.

We now recall the Hahn-Banach theorem in its geometric form, and its consequences for the separation of convex sets. Naturally it is in the context of l.c.s. that the most precise results are obtained.

Hahn–Banach Theorem. *Let V be a real t.v.s., $\mathscr{A}$ an open non-empty convex set, and $\mathscr{M}$ a non-empty affine subspace which does not intersect $\mathscr{A}$. Then there exists a closed affine hyperplane $\mathscr{H}$ which contains $\mathscr{M}$ and does not intersect $\mathscr{A}$.*

Corollary 1.1. *Let V be a real t.v.s., $\mathscr{A}$ an open non-empty convex set, $\mathscr{B}$ a non-empty convex set which does not intersect $\mathscr{A}$. Then there exists a closed affine hyperplane $\mathscr{H}$ which separates $\mathscr{A}$ and $\mathscr{B}$.*

Corollary 1.2. *Let V be a real l.c.s., $\mathscr{C}$ and $\mathscr{B}$ two non-empty disjoint convex sets with one compact and the other closed. Then there exists a closed affine hyperplane $\mathscr{H}$ which strictly separates $\mathscr{C}$ and $\mathscr{B}$.*

Here is an example of application of Corollary 1.1. Let $\mathscr{A}$ be a subset of V and $\mathscr{H}$ a closed affine hyperplane which contains at least one point $u \in \mathscr{A}$, such that $\mathscr{A}$ is completely contained in one of the closed half-spaces determined by $\mathscr{H}$; we say that $\mathscr{H}$ is a *supporting* hyperplane and u a *supporting point* of $\mathscr{A}$. Then

Corollary 1.3. *In a real t.v.s. V, let $\mathscr{A}$ be a convex set with non-empty interior. Then every boundary point of $\mathscr{A}$ is a supporting point of $\mathscr{A}$.*

An application of Corollary 1.2 is:

Corollary 1.4. *In a real l.c.s. V, every closed convex set is the intersection of the closed half-spaces which contain it.*

All these results have a fundamental importance in analysis because they allow a convenient theory of duality. Thus, if V is a Hausdorff l.c.s., the Hahn-Banach theorem allows us to assert the existence of non-zero continuous linear forms over V: it is sufficient to consider two points u_1 and u_2 of V, and to separate then by a closed affine hyperplane $\mathscr{H}$ (Cor. 1.2); if the equation of

$\mathcal{H}$ is $\ell(u) = \alpha$, the non-zero linear form ℓ is continuous since $\mathcal{H}$ is closed and $\ell(u_1) \neq \ell(u_2)$. The vector space V^* of continuous linear functionals over V is said to be the *topological dual*, or more simply the *dual* of V. The elements of V^* will be, in general, denoted by u^* or v^*, and $\langle u,u^* \rangle$ will denote the value at u of the continuous linear functional $u^* \in V^*$.

These new notations emphasize the fact that V and V^* play symmetrical roles. In $(u,u^*) \to \langle u, u^* \rangle$ we have a bilinear form over $V \times V^*$ which can be considered either as a family of linear forms on V depending on the parameter $u^* \in V^*$, or as a family of linear forms over V^* depending on the parameter $u \in V$. Thus V^* (resp. V) is a vector space of linear forms over V (resp. V^*), the point $u^* \in V^*$ being identified with the function $u \to \langle u, u^* \rangle$ (or the point $u \in V$ being identified with the function $u^* \to \langle u, u^* \rangle$). We can thus introduce over V^* (resp. V) the topology of weak convergence over V (resp. V^*). This will be termed *weak topology* of V^* (resp. V) associated with the duality between V and V^* and will be denoted $\sigma(V^*,V)$ (resp. $\sigma(V,V^*)$). It is a topology of Hausdorff l.c.s. and $\sigma(V, V^*)$ is the coarsest of the Hausdorff l.c.s. topologies over V with dual V^*; in particular it is coarser than the original topology over V. The weakly closed subsets of V will thus be closed, while the converse is generally false.

From Corollary 1.4, however, *every closed convex set is weakly closed.* In a Hausdorff l.c.s. the weakly closed convex sets are identical with the closed convex sets. In the context of normed spaces we shall sometimes use this result in the following form:

Mazur's Lemma. *Let V be a normed space and $(u_n)_{n \in \mathbf{N}}$ a sequence converging weakly to $\bar{u}$. Then there is a sequence of convex combinations $(v_n)_{n \in \mathbf{N}}$ such that*

$$v_n = \sum_{k=n}^{N} \lambda_k u_k, \quad \textit{where} \ \sum_{k=n}^{N} \lambda_k = 1 \ \textit{and} \ \lambda_k \geqslant 0, \quad n \leqslant k \leqslant N$$

which converges to $\bar{u}$ in norm.

Proof. For every $n \in \mathbf{N}$, $\bar{u}$ belongs to the weak closure of $\bigcup_{k=n}^{\infty} \{u_k\}$, and *a fortiori*, to the weak closure of co $\bigcup_{k=n}^{\infty} \{u_k\}$. But this is exactly the weakly closed convex hull of $\bigcup_{k=n}^{\infty} \{u_k\}$ which coincides with its closed convex hull by Corollary 1.4. Finally,

$$\bar{u} \in \overline{\mathrm{co}} \bigcup_{k=n}^{\infty} \{ u_k \}, \qquad \forall n \in \mathbf{N},$$

and it suffices to choose $v_n \in \mathrm{co} \bigcup_{k=n}^{\infty} \{u_k\}$ such that $\|v_n - \bar{u}_n\| \leqslant 1/n$. ■

1.3. Analytical form of the Hahn–Banach theorem

The analytical form of the Hahn-Banach theorem is more precise than the geometrical form. We consider a real vector space V and a sub-linear function j over V, i.e. a mapping j of V into $\bar{\mathbf{R}}$ which satisfies

$$\begin{aligned} j(\lambda u) &= \lambda j(u), && \forall u \in V,\ \forall \lambda > 0, \\ j(u+v) &\leqslant j(u) + j(v), && \forall (u, v) \in V \times V. \end{aligned}$$

Hahn–Banach Theorem. *Let V be a real vector space, j a sub-linear function over V, $\mathscr{M}$ a vector subspace of V, ℓ a linear functional over $\mathscr{M}$ which is everywhere less than j. Then there exists a linear functional $\tilde{\ell}$ over V which extends ℓ and is everywhere less than j.*

Corollary 1.5. *Let V be a normed space, $\mathscr{M}$ a topological vector subspace, ℓ a continuous linear functional over $\mathscr{M}$. Then ℓ can be extended into a continuous linear functional over V with the same norm.*

2. CONVEX FUNCTIONS

2.1. Definitions

As before, we take a real vector space V and consider mappings of $\mathscr{A} \subset V$ into $\bar{\mathbf{R}}$, that is, the values $+\infty$ and $-\infty$ are allowed to the functions under consideration.

Definition 2.1. Let $\mathscr{A}$ be a convex subspace of V, and F a mapping of $\mathscr{A}$ into $\bar{\mathbf{R}}$. F is said to be *convex* if, for every u and v in $\mathscr{A}$, we have:

$$F(\lambda u + (1-\lambda)v) \leqslant \lambda F(u) + (1-\lambda)F(v) \qquad \forall \lambda \in [0, 1] \tag{2.1}$$

whenever the right-hand side is defined.

The inequality (2.1) must therefore be valid unless $F(u) = -F(v) = \pm\infty$. By induction, it can be shown that if F is convex, for every finite set $u_1, \ldots, u_n$ of points of V and for every family $\lambda_1, \ldots, \lambda_n$ of real positive numbers with sum unity, then

$$F\left(\sum_{i=1}^{n} \lambda_i u_i\right) \leqslant \sum_{i=1}^{n} \lambda_i F(u_i) \tag{2.2}$$

whenever the right-hand side is defined.

It is easy to see that, if $F: V \to \bar{\mathbf{R}}$ is convex, the sections

$$\{ u \mid F(u) \leqslant a \} \quad \text{and} \quad \{ u \mid F(u) < a \} \tag{2.3}$$

are, for each $a \in \bar{\mathbf{R}}$, convex sets of V. The converse is, however, false. For instance, if F is convex and if $\phi: \bar{\mathbf{R}} \to \bar{\mathbf{R}}$ is an increasing function then $\phi \circ F: V \to \bar{\mathbf{R}}$ will also have convex sections but will not be convex in general.

For every mapping $F: V \to \bar{\mathbf{R}}$, we call the section:

$$\operatorname{dom} F = \{ u \mid F(u) < +\infty \} \tag{2.4}$$

the *effective domain* of F. The effective domain of a convex function is thus convex.

Why do we allow the value $+\infty$? If F is a mapping of $\mathscr{A} \subset V$ into $\mathbf{R}$, we can associate with it the function $\tilde{F}$ defined throughout V by:

$$\left| \begin{array}{lll} \tilde{F}(u) = F(u) & \text{if} & u \in \mathscr{A}, \\ \tilde{F}(u) = +\infty & \text{if} & u \notin \mathscr{A}. \end{array} \right. \tag{2.5}$$

Thus $\tilde{F}$ is convex if and only if $\mathscr{A} \subset V$ is convex and $F: \mathscr{A} \to \mathbf{R}$ is convex. In the theory of convex functions, because of this extension by $+\infty$, we need only consider those functions defined everywhere.

There is a further advantage. If $\mathscr{A}$ is a subset of V, the *indicator function* $\chi_{\mathscr{A}}$ of $\mathscr{A}$ is defined as:

$$\left| \begin{array}{lll} \chi_{\mathscr{A}}(u) = 0 & \text{if} & u \in \mathscr{A}, \\ \chi_{\mathscr{A}}(u) = +\infty & \text{if} & u \notin \mathscr{A}. \end{array} \right. \tag{2.6}$$

Clearly $\mathscr{A}$ is a convex subset if and only if $\chi_{\mathscr{A}}$ is a convex function. Thus the study of convex sets is naturally reduced to the study of convex functions.

On the other hand, convex functions which assume the value $-\infty$ are very special. If $F(\bar{u}) = -\infty$, then on every half-line starting from $\bar{u}$, either F is identically equal to $-\infty$, or F takes the value $-\infty$ between $\bar{u}$ and a point $\bar{v}$, any value at $\bar{v}$, and $+\infty$ beyond $\bar{v}$.

To distinguish these very special cases, we shall say that a convex function F of V in $\bar{\mathbf{R}}$ is *proper* if it nowhere takes the value $-\infty$ and is not identically equal to $+\infty$.

Definition 2.2. The *epigraph* of a function $F: V \to \bar{\mathbf{R}}$ is the set:

$$\operatorname{epi} F = \{ (u, a) \in V \times \mathbf{R} \mid f(u) \leqslant a \}. \tag{2.7}$$

It is the set of points of $V \times \mathbf{R}$ which lie above the graph of F. The projection of epi F on V is none other than dom F. The epigraph will be found a most useful concept in the study of convex functions because of the following result:

Proposition 2.1. *A function $F: V \to \bar{\mathbf{R}}$ is convex if and only if its epigraph is convex.*

Proof. Let F be convex and take (u,a) and (v,b) in epi F. Then, necessarily, $F(u) \leqslant a < +\infty$ and $F(v) \leqslant b < +\infty$, and for all $\lambda \in [0,1]$ from (2.1) we have

$$F(\lambda u + (1 - \lambda)v) \leqslant \lambda F(u) + (1 - \lambda)F(v) \leqslant \lambda a + (1 - \lambda)b,$$

which means precisely that $\lambda(u,a) + (1 - \lambda)(v,b) \in \text{epi } F$.

Conversely, let epi F be convex. Its projection dom F is therefore convex and it is sufficient to verify (2.1) over dom F. Let us therefore take u and v in dom F, $a \geqslant F(u)$ and $b \geqslant F(v)$. By hypothesis, $\lambda(u,a) + (1 - \lambda)(v,b) \in \text{epi } F$ for every $\lambda \in [0,1]$ so that:

$$F(\lambda u + (1 - \lambda)v) \leqslant \lambda a + (1 - \lambda)b.$$

If $F(u)$ and $F(v)$ are finite, it is sufficient to take $a = F(u)$ and $b = F(v)$. If either $F(u)$ or $F(v)$ is equal to $-\infty$ it is sufficient to allow a or b to tend to $-\infty$ and (2.1) is obtained in both cases. ■

It remains for us to consider the usual manipulations of convex functions. The results, which are trivial, have been collected together in the following proposition. In particular we infer from them that the set of convex functions is a convex cone.

Proposition 2.2. (i) *If $F: V \to \bar{\mathbf{R}}$ is convex and if λ is a real positive number, then λF is convex.*

(ii) *If F and G are convex functions from V into $\bar{\mathbf{R}}$, then $F + G$ is convex. We stipulate that $(F + G)(u) = +\infty$ if $F(u) = -G(u) = \pm\infty$.*

(iii) *If $(F_i)_{i \in I}$ is any family of convex functions of V into $\bar{\mathbf{R}}$, their pointwise supremum $F = \sup_{i \in I} F_i$ is convex.*

Also, let us recall the concept of a strictly convex function.

Definition 2.3. Let $\mathscr{A}$ be a convex set of V and F a mapping of $\mathscr{A}$ into $\mathbf{R}$. F is said to be *strictly convex* if it is convex and the strict inequality holds in (2.1), $\forall u, v \in \mathscr{A}, u \neq v$ and $\forall \lambda \in \,]0,1[$.

2.2. Lower semi-continuous functions

We now pass on to topological properties, so V will be a real l.c.s. We recall that a function $F: V \to \bar{\mathbf{R}}$ is said to be *lower semi-continuous* on V (l.s.c.), if it satisfies the two equivalent conditions:

$$(2.8) \qquad \forall a \in \mathbf{R}, \qquad \{ u \in V \mid F(u) \leqslant a \} \quad \text{is closed,}$$

$$(2.9) \qquad \forall \bar{u} \in V, \qquad \varliminf_{u \to \bar{u}} F(u) \geqslant F(\bar{u}).$$

Naturally F will be called upper semi-continuous (u.s.c.) if $-F$ is l.s.c. Thus for example the indicator function $\chi_{\mathscr{A}}(\cdot)$ of a set $\mathscr{A} \subset V$ will be l.s.c. (or u.s.c.) if and only if $\mathscr{A}$ is closed (or open). More generally, we can at once state:

Proposition 2.3. *A function F: $V \to \bar{\mathbf{R}}$ is l.s.c. if and only if its epigraph is closed.*

Proof. Let us define on $V \times \mathbf{R}$ a function ϕ by $\phi(u,a) = F(u) - a$. Then the statements that F is l.c.s. on V and that ϕ is l.c.s. on $V \times \mathbf{R}$ are equivalent to each other. Now for every $r \in \mathbf{R}$, the section $\phi([-\infty, r[)$ is the set obtained from epi F by the translation of vector $(r,0)$ and it is therefore closed if and only if epi F is closed. ∎

Let us further recall that a pointwise supremum of l.s.c. functions is l.s.c. This then leads to the following definition: for every mapping F: $V \to \bar{\mathbf{R}}$, the largest l.s.c. minorant of F will be called the l.s.c. regularization of F and will be denoted by $\bar{F}$. It exists as the pointwise supremum of those l.s.c. functions everywhere less than F, and is characterized by:

Corollary 2.1. *Let F: $V \to \bar{\mathbf{R}}$ and $\bar{F}$ be its l.s.c. regularization. We have*

$$(2.10) \qquad \text{epi } \bar{F} = \overline{\text{epi } F},$$

$$(2.11) \qquad \forall u \in V, \qquad \bar{F}(u) = \varliminf_{v \to u} F(v).$$

Proof. Since $\bar{F}$ is a l.s.c. function everywhere less than F, epi $\bar{F}$ is a closed set containing epi F, and hence $\overline{\text{epi } F}$.

Conversely, $\overline{\text{epi } F}$ is an epigraph.[1] Let G be a function such that epi $G = \overline{\text{epi } F}$; it is a l.s.c. function everywhere less than F, hence $G \leqslant \bar{F}$ and epi $G = \overline{\text{epi } F}$ contains epi $\bar{F}$. Thus (2.10) is proved and (2.11) follows directly from it. ∎

[1] If $(u,a) \in \overline{\text{epi } F}$, there exists a filter $(u_\alpha, a_\alpha) \in \text{epi } F$ which converges to (u,a). If $b > a$, then for some suitable α, $a_\alpha \leqslant b$ and since $F(u_\alpha) \leqslant a_\alpha$ it follows that $(u_\alpha, b) \in \text{epi } F$, and in the limit $(u,b) \in \overline{\text{epi } F}$.

The intersection of $\overline{\text{epi } F}$ with a straight line $\{u\} \times \mathbf{R}$ is an empty set or an interval $[a, +\infty[$; if we set $G(u) = +\infty$ in the former and $G(u) = a$ in the latter case, we see that epi $G = \overline{\text{epi } F}$.

The case of convex functions assumes a special interest since lower semi-continuity exists when the topology of V is weakened. This is an extremely valuable property.

Corollary 2.2. *Every l.s.c. convex function F of V in $\bar{\mathbf{R}}$ remains l.s.c. when V is supplied with its weak topology $\sigma(V, V')$.*

In fact, the epigraph of F being convex, this is equivalent to saying that it is closed and hence weakly closed (Cor. 1.4).

The case of l.s.c. convex functions which assume the value $-\infty$ (improper functions) is more specialized:

Proposition 2.4. *If $F: V \to \bar{\mathbf{R}}$ is a l.s.c. convex function and assumes the value $-\infty$, it cannot take any finite value.*

Proof. Let us suppose that there exists $\bar{u} \in V$ a point such that $F(\bar{u}) \in \mathbf{R}$. We then take $\bar{a} \in \mathbf{R}$ such that $\bar{a} < F(\bar{u})$, and we strictly separate $(\bar{u}, \bar{a})$ from the closed convex set epi F. Then there exists a continuous non-zero linear form ℓ over V and $\beta \in \mathbf{R}$ such that:

$$\forall (u, a) \in \text{epi } F, \qquad \ell(\bar{u}) + \alpha\bar{a} < \ell(u) + \alpha a. \tag{2.12}$$

Taking $u = \bar{u}$ and $a = F(\bar{u})$, we get $\alpha(F(\bar{u}) - a) > 0$, and hence $\alpha > 0$. The two members of (2.12) can thus be divided by α:

$$\forall u \in V, \qquad \frac{1}{\alpha}\ell(\bar{u} - u) + \bar{a} < F(u), \tag{2.13}$$

which is impossible, since the first member is everywhere finite and the second member takes the value $-\infty$ at one point at least. ■

2.3. Continuity of convex functions

The study of the continuity of convex functions is based on the following lemma:

Lemma 2.1. *If in the neighbourhood of a point $u \in V$, a convex function F is bounded above by a finite constant, then F is continuous at u.*

Proof. We reduce the problem by translation to the case where $u = 0$ and $F(0) = 0$. Let $\mathscr{V}$ be a neighbourhood of the origin such that $F(v) \leqslant a < +\infty$ for all v of $\mathscr{V}$. Let us define $\mathscr{W} = \mathscr{V} \cap -\mathscr{V}$ (which is a symmetric neighbourhood of the origin), and let us take $\varepsilon \in]0, 1[$. If $v \in \varepsilon\mathscr{W}$, we have, due to the convexity of F:

$$\frac{v}{\varepsilon} \in \mathscr{V}, \quad \text{hence} \quad F(v) \leqslant (1 - \varepsilon)F(0) + \varepsilon F(v/\varepsilon) \leqslant \varepsilon a,$$

$$-\frac{v}{\varepsilon} \in \mathscr{V}, \quad \text{hence} \quad F(v) \geqslant (1 + \varepsilon)F(0) - \varepsilon F(-v/\varepsilon) \geqslant -\varepsilon a.$$

Then $|F(v)| \leqslant \varepsilon a$ for every v in $\varepsilon \mathscr{W}$, whence we have the required continuity. ∎

A general conclusion may be drawn:

Proposition 2.5. *Let $F: V \to \bar{\mathbf{R}}$ be a convex function. The following statements are equivalent to each other:*

(i) *there exists a non-empty open set $\mathcal{O}$ on which F is not everywhere equal to $-\infty$ and is bounded above by a constant $a < +\infty$;*

(ii) *F is a proper function, and it is continuous over the interior of its effective domain, which is non-empty.*

Proof. Clearly (ii) implies (i). Conversely, if (i) is true, $\mathcal{O} \subset \widehat{\mathrm{dom}\, F}^{\circ}$. Let us take $u \in \mathcal{O}$ such that $F(u) > -\infty$. From Lemma 2.1, F will be continuous at u, and hence finite in a neighbourhood of u, and hence proper. For every $v \in \widehat{\mathrm{dom}\, F}^{\circ}$, there exists $\rho > 1$ such that $w = u + \rho(v - u)$ also belongs to $\widehat{\mathrm{dom}\, F}^{\circ}$. The homothety h with centre w and ratio $1 - 1/\rho$ transforms u into v and $\mathcal{O}$ into an open set $h(\mathcal{O})$ containing v. For every $v' \in h(\mathcal{O})$, we have by convexity:

$$F(v') \leqslant \frac{\rho - 1}{\rho} F \circ h^{-1}(v') + \frac{1}{\rho} F(w) \leqslant \frac{\rho - 1}{\rho} a + \frac{1}{\rho} F(w).$$

To sum up: every point $v \in \widehat{\mathrm{dom}\, F}^{\circ}$ possesses a neighbourhood $h(\mathcal{O})$ where F is bounded above by a finite constant. From Lemma 2.1, F is continuous at v, which concludes the proof. ∎

We can use this result more precisely in numerous special cases—spaces of finite dimension, normed spaces and barrelled spaces.

Corollary 2.3. *Every proper convex function on a space of finite dimension is continuous on the interior of its effective domain.*

Proof. If $\widehat{\mathrm{dom}\, F}^{\circ}$ is non-empty, it contains $n + 1$ affinely independent points u_1, $1 \leqslant i \leqslant n + 1$, where n is the dimension of V. From the inequality defining convexity, F is bounded above by $\max_{1 \leqslant i \leqslant n+1} F(u_1)$ over the open set:

$$\left\{ \sum_{i=1}^{n+1} \lambda_i u_i \,\middle|\, \sum_{i=1}^{n+1} \lambda_i = 1 \text{ and } \lambda_i > 0 \quad \forall i \right\}. \blacksquare$$

Corollary 2.4. *Let F be a proper convex function on a normed space. The following properties are equivalent to each other:*

(i) *there exists a non-empty open set over which F is bounded above;*

(ii) *$\widehat{\mathrm{dom}\, F}^{\circ} \neq \varnothing$ and F is locally Lipschitz there.*

Proof. It is obvious that (ii) $\Rightarrow$ (i). Conversely, if (i) is true, F is continuous over $\widehat{\mathring{\text{dom}}\, F}$ (Prop. 2.5). Taking $u \in \widehat{\mathring{\text{dom}}\, F}$, for every $r > 0$ we define:

$$\mathscr{B}(u; r) = \{ v \mid \|v - u\| \leqslant r \}.$$

By the continuity of F at u, there exists an $r_0 > 0$ such that:

$$\forall v \in \mathscr{B}(u; r_0), \qquad -\infty < m \leqslant F(v) \leqslant M < +\infty.$$

Let us now suppose that $r \in]0,r_0[$, and let us take $v_1 \in \mathscr{B}(u,r)$. Let us set

$$G(w) = F(w + v_1) - F(v_1),$$

so that $G(0) = 0$, and $\mathscr{W} = \{w \mid \|w\| \leqslant r_0 - r\}$. Then G is bounded above by $M - m$ over $\mathscr{W}$, and due to the proof of Lemma 1.1:

$$\text{(2.14)} \qquad \forall \varepsilon \in [0, 1], \ \forall w \in \varepsilon\mathscr{W}, \qquad |G(w)| \leqslant \varepsilon(M - m).$$

If $\|v - v_1\| \leqslant r_0 - r$, then $w = v - v_1 \in \varepsilon\mathscr{W}$, with $\varepsilon = \|v - v_1\|/(r_0 - r)$, and substituting back into (2.14):

$$\text{(2.15)} \quad \forall v \in \mathscr{B}(v_1; r_0 - r), \qquad |F(v) - F(v_1)| \leqslant \frac{M - m}{r_0 - r} \|v - v_1\|.$$

If finally $v_2 \in \mathscr{B}(u;r)$, we take equidistant points $u_1 = v_1, u_2, \ldots, u_{n-1}$, $u_n = v_2$ on the segment $[v_1,v_2] \subset \mathscr{B}(u;r)$, in sufficient number so that $\|u_n - u_{n+1}\| \leqslant r_0 - r$ for $1 \leqslant k \leqslant n$. From (2.15), we have

$$|F(u_k) - F(u_{k+1})| \leqslant \frac{M - m}{r_0 - r} \|u_k - u_{k+1}\| \quad \text{for} \quad 1 \leqslant k \leqslant n$$

and by adding each member, we obtain the local Lipschitz condition:

$$v_1 \text{ and } v_2 \in \mathscr{B}(u; r) \Rightarrow |F(v_1) - F(v_2)| \leqslant \frac{M - m}{r_0 - r} \|v_1 - v_2\|. \quad \blacksquare$$

Remark 2.1. Clearly the above proof gives an estimate of the Lipschitz constant for (ii).

Corollary 2.5. *Every l.s.c. convex function over a barrelled space (in particular a Banach space) is continuous over the interior of its effective domain.*

Proof. Let $u \in \widehat{\mathring{\text{dom}}\, F}$, which is assumed to be non-empty. By a suitable translation, we can shift u to the origin. Then let $a > F(0)$. The set $\mathscr{C} = \{u \in V \mid F(u) \leqslant a\}$ is closed and convex. Also it is absorbent, since the restriction of F to every straight line passing through the origin is continuous in the neighbourhood of the origin (Cor. 2.3). Thus $\mathscr{C} \cap -\mathscr{C}$ is a barrel and hence a neighbourhood of the origin. As F is bounded by a over $\mathscr{C}$, F is continuous at 0 (Prop. 2.5).

3. POINTWISE SUPREMUM OF CONTINUOUS AFFINE FUNCTIONS

3.1. Definition of $\Gamma(V)$

As usual, V is a real l.c.s. The affine continuous functions over V are functions of the type $v > \ell(v) + \alpha$, where ℓ is a continuous linear functional over V and $\alpha \in \mathbf{R}$.

Definition 3.1. The set of functions $F: V \to \overline{\mathbf{R}}$ which are pointwise supremum of a family of continuous affine functions is denoted by $\Gamma(V)$. $\Gamma_0(V)$ denotes the subset of $F \in \Gamma(V)$ other than the constants $+\infty$ and $-\infty$.

It follows immediately from this definition that all the functions $F \in \Gamma(V)$ are convex and l.s.c. Conversely:

Proposition 3.1. *The following properties are equivalent to each other*:

(i) $F \in \Gamma(V)$

(ii) *F is a convex l.s.c. function from V into $\overline{\mathbf{R}}$, and if F takes the value $-\infty$ then F is identically equal to $-\infty$.*

Proof. Note that the pointwise supremum of an empty family is $-\infty$ and that if the family under consideration is non-empty, F cannot take the value $-\infty$. Therefore we have (i) $\Rightarrow$ (ii).

Conversely, suppose that F is a convex l.s.c. function of V into $\overline{\mathbf{R}}$ not taking the value $-\infty$. If F is the constant $+\infty$, it is the pointwise supremum of all the continuous affine functions of V into $\mathbf{R}$. If $F \in \Gamma_0(V)$, for every $\bar{u} \in V$ and for every $\bar{a} < F(\bar{u})$ we will show that there is a continuous affine function of V into $\mathbf{R}$ whose value at u is located between $\bar{a}$ and $F(\bar{u})$, which establishes the result.

Now epi F is a closed convex set which does not contain the point $(\bar{u}, \bar{a})$. We can strictly separate them by a closed affine hyperplane $\mathscr{H}$ of $V \times \mathbf{R}$ with equation:

$$\mathscr{H} = \{ (u, a) \in V \times \mathbf{R} \,|\, \ell(u) + \alpha a = \beta \} \tag{3.1}$$

where ℓ is a continuous non-zero linear functional over V, α and $\beta \in \mathbf{R}$. We thus have:

$$\ell(\bar{u}) + \alpha \bar{a} < \beta \tag{3.2}$$

$$(3.3) \qquad \forall (u, a) \in \text{epi } F, \qquad \ell(u) + \alpha a > \beta.$$

If $F(\bar{u}) < +\infty$, we can take $u = \bar{u}$ and $a = F(\bar{u})$ which gives $\alpha(F(\bar{u}) - \bar{a}) > 0$ where $\alpha > 0$. When (3.2) and (3.3) are divided by α, we obtain:

$$(3.4) \qquad \bar{a} < \frac{\beta}{\alpha} - \frac{1}{\alpha}\ell(\bar{u}) < F(\bar{u}).$$

The continuous affine function $\frac{\beta}{\alpha} - \frac{1}{\alpha}\ell(.)$ (whose graph is nothing other than $\mathscr{H}$) therefore answers the problem.

If $F(\bar{u}) = +\infty$, either $\alpha \neq 0$ and we are back with the preceding case, or $\alpha = 0$. In this case, (3.2) and (3.3) mean that the continuous affine function $\beta - \ell(.)$ is >0 at $\bar{u}$ and <0 over dom F. The above case allows us to construct a continuous affine function everywhere less than F, e.g. $\gamma - m(.)$. Then for every $c > 0$, $\gamma - m(.) + c(\beta - \ell(.))$ is always a continuous affine function everywhere less than F, and it only remains to choose c sufficiently large so that

$$(3.5) \qquad \gamma - m(\bar{u}) + c(\beta - \ell(\bar{u})) > \bar{a}. \ \blacksquare$$

3.2. Γ-regularization

Definition 3.2. Let F and G be two functions of V into $\overline{\mathbf{R}}$. The following are equivalent to each other:

(i) G is the pointwise supremum of the continuous affine functions everywhere less than F;

(ii) G is the largest minorant of F in $\Gamma(V)$. G is then called the Γ-regularization of F.

We shall now show the equivalence of (i) and (ii). Let us call G_1 (or G_2) the pointwise supremum of continuous affine functions (or functions of $\Gamma(V)$) everywhere less than F. Then G_1 and G_2 belong to $\Gamma(V)$, as the pointwise supremum of functions of $\Gamma(V)$, and $G_1 \leqslant G_2$. Conversely, every continuous affine minorant of G_2 is a function of $\Gamma(V)$ everywhere less than F. By definition it must be less than G_1 everywhere. The functions G_1 and G_2 have the same set of affine continuous minorants. As they belong both to $\Gamma(V)$, they must coincide. $\blacksquare$

In particular, if $F \in \Gamma(V)$, it coincides with its Γ-regularization. In general, we can construct the epigraph of the Γ-regularization as the closed convex hull of the epigraph of the function.

Proposition 3.2. *Let $F: V \to \bar{\mathbf{R}}$ and G be its Γ-regularization. If there exists a continuous affine function everywhere less than F, we have:*

$$\text{epi } G = \overline{\text{co}} \text{ epi } F.$$

Proof. Let ϕ be a continuous affine function everywhere less than F. It is easy to see that the closed convex set co epi F is the epigraph of a convex l.s.c. function G. Since

$$\text{epi F} \subset \overline{\text{co}} \text{ epi } F \subset \text{epi } \phi$$

we have $F \geqslant G \geqslant \phi$, and therefore $G \in \Gamma(V)$. If, finally, $G' \leqslant F$ where $G' \in \Gamma(V)$ then epi G' is a closed convex set containing epi F, and hence containing $\overline{\text{co}}$ epi $F =$ epi G, which means that $G' \leqslant G$. ∎

Thus, for example, if $\mathscr{A} \subset V$, the Γ-regularization of its indicator function $\chi_{\mathscr{A}}$ is none other than the indicator function of its closed convex envelope.

We may wonder what ordering relations exist between F, its Γ-regularization G and its l.s.c. regularization $\bar{F}$ as defined in Section 2. Corollary 2.1 and Proposition 3.2 immediately give us the result:

Proposition 3.3. *Let $F: V \to \bar{\mathbf{R}}$, and G be its Γ-regularization.*

(i) $G \leqslant \bar{F} \leqslant F$;

(ii) *if F is convex and admits a continuous affine minorant, $\bar{F} = G$.*

4. POLAR FUNCTIONS

4.1. Definition

In this paragraph, as in those which follow, we shall designate by V and V^* two vector spaces placed in duality by a bilinear pairing denoted by $\langle .,. \rangle$. The spaces V and V^* will be supplied with topologies $\sigma(V, V^*)$ and $\sigma(V^*, V)$ which render them l.c.s. and Hausdorff.

Let F be a function of V into $\bar{\mathbf{R}}$. If $u^* \in V^*$ and $\alpha \in \mathbf{R}$, the continuous affine function $u \mapsto \langle u, u^* \rangle - \alpha$ is everywhere less than F if and only if

$$\forall u \in V, \qquad \alpha \geqslant \langle u, u^* \rangle - F(u),$$

or again:

$$\alpha \geqslant F^*(u^*) \tag{4.1}$$

if we agree to set

$$F^*(u^*) = \sup_{u \in V} \{ \langle u, u^* \rangle - F(u) \}. \tag{4.2}$$

The consideration of the continuous affine minorants of F thus leads us to define by (4.2) a function $F^*: V^* \to \bar{\mathbf{R}}$.

Definition 4.1. *If $F: V \to \bar{\mathbf{R}}$, formula* (4.2) *defines a function from V^* into $\bar{\mathbf{R}}$, denoted by F^*, and called the polar* (*or conjugate*) *function of F.*

It is obvious that in (4.2) we can confine ourselves to those u in dom F:

(4.3) $$F^*(u^*) = \sup_{u \in \operatorname{dom} F} \{ \langle u, u^* \rangle - F(u) \}$$

which enables us to see that F^* is the pointwise supremum of the family of continuous affine functions $\langle u, . \rangle - F(u)$, for $u \in \operatorname{dom} F$, of V^* into $\bar{\mathbf{R}}$. We therefore conclude that $F^* \in \Gamma(V^*)$, and in particular that F is l.s.c. and convex. Note that if F is the constant $+\infty$, dom $F = \varnothing$ and F^* is the constant $-\infty$.

This immediately results in the following properties:

(4.4) $$F^*(0) = -\inf_{u \in V} F(u);$$

(4.5) *if $F \leqslant G$, we have $F^* \geqslant G^*$;*

(4.6) $$(\inf_{i \in I} F_i)^* = \sup_{i \in I} F_i^*,$$
$$(\sup_{i \in I} F_i)^* \leqslant \inf_{i \in I} F_i^*,$$

for every family $(F_i)_{i \in I}$ of functions over V;

(4.7) $$(\lambda F)^*(u^*) = \lambda F^*(u^*/\lambda),$$
for every $\lambda > 0$;

(4.8) $$(F + \alpha)^* = F^* - \alpha,$$
for every $\alpha \in \mathbf{R}$;

(4.9) *for every $a \in V$, we denote by F_a the translated function $F_a(v) = F(v - a)$. Then*

$$(F_a)^*(u^*) = F^*(u^*) + \langle a, u^* \rangle.$$

4.2. Bipolars. Dual convex functions

We can repeat the process, thereby leading to the bipolar F^{**}, which is now a function of V into $\bar{\mathbf{R}}$:

(4.10) $$F^{**}(u) = \sup_{u^* \in V^*} \{ \langle u, u^* \rangle - F^*(u^*) \}.$$

From the above, $F^{**} \in \Gamma(V)$, and we can compare F and F^{**}, which are defined over the same space. The result is as follows:

Proposition 4.1. *Let F be a function of V into $\overline{\mathbf{R}}$. Then its bipolar F^{**} is none other than its Γ-regularization. In particular, if $F \in \Gamma(V)$, $F^{**} = F$.*

Proof. By definition, the Γ-regularization of F is the pointwise supremum of all continuous affine minorants of F. We can restrict ourselves to those which are maximal, i.e., from (4.1), those functions:

$$u \mapsto \langle u, u^* \rangle - F^*(u^*). \tag{4.11}$$

But their pointwise supremum is none other than F^{**}, from (4.10). Whence the result. ∎

The repetition of this process is limited:

Corollary 4.1. *For every $F\colon V \to \overline{\mathbf{R}}$, we have $F^* = F^{***}$.*

Proof. As F^{**} is the Γ-regularization of F, we have $F^{**} \leqslant F$, and so from (4.5):

$$F^* \leqslant F^{***}.$$

Alternatively, from (4.10), for every $u \in C$:

$$\langle u^*, u \rangle - F^{**}(u) \leqslant F^*(u^*)$$

whence

$$F^{***}(u^*) = \sup_{u \in V} \{ \langle u^*, u \rangle - F^{**}(u) \} \leqslant F^*(u^*).$$

We have seen that $F \in \Gamma(V)$ if and only if $F^{**} = F$. We thus arrive at the following definition:

Definition 4.2. The polarity establishes a bijection between $\Gamma(V)$ and $\Gamma(V^*)$. $F \in \Gamma(V)$ and $G \in \Gamma(V^*)$ are said to be *in duality* if they correspond in the bijection:

$$F = G^* \quad \textit{and} \quad G = F^*. \tag{4.12}$$

The constants $\pm\infty$ on V and $\pm\infty$ on V^* are in duality. Thus $F \in \Gamma_0(V)$ if and only if $F^* \in \Gamma_0(V^*)$: the polarity establishes a one-to-one correspondence between $\Gamma_0(V)$ and $\Gamma_0(V^*)$.

4.3. Examples

Let $\mathscr{A}$ be a subset of V and $\chi_{\mathscr{A}}$ its indicator function. Let us seek its polar:

$$\chi_{\mathscr{A}}^*(u^*) = \sup_{u \in V} \{ \langle u, u^* \rangle - \chi_a(u) \}$$

$$\chi_{\mathscr{A}}^*(u^*) = \sup_{u \in \mathscr{A}} \langle u, u^* \rangle.$$

It is a convex function, l.s.c. and positively homogeneous on V^*, termed the *support function* of $\mathscr{A}$. We have seen in Section 3 that $\chi_{\mathscr{A}}^{**} = \chi_{\overline{\mathrm{co}}\,\mathscr{A}}$. In particular, $\mathscr{A}$ and $\overline{\mathrm{co}}\,\mathscr{A}$ will have the same support function.

Here is an extremely useful example of dual convex functions. We will take for V a normed space, for V^* its topological dual, and denote by $\|.\|$ the norm of V and by $\|.\|_*$ the norm of V^*. Then V and V^* are in duality and we supply them with the weak topologies $\sigma(V, V^*)$ and $\sigma(V^*, V)$.

We take an *even* function $\varphi \in \Gamma_0(\mathbf{R})$ and we call φ^* its conjugate convex function which also belongs to $\Gamma_0(\mathbf{R})$. Then we define $F: V \to \mathbf{R}$ and $G: V^* \to \mathbf{R}$ by:

$$F(u) = \varphi(\|u\|)$$
$$G(u^*) = \varphi^*(\|u^*\|_*).$$

Proposition 4.2. *Under the above hypotheses, F and G are dual.*

Proof. It is obvious that $F \in \Gamma_0(V)$ and $G \in \Gamma_0(V^*)$. It is thus sufficient to prove that $F^* = G$. To do this, we write:

$$\begin{aligned} F^*(u^*) &= \operatorname*{Sup}_{u \in V} \{ \langle u^*, u \rangle - \varphi(\|u\|) \} \\ &= \operatorname*{Sup}_{t \geq 0} \operatorname*{Sup}_{\substack{u \in V \\ \|u\| = t}} \{ \langle u^*, u \rangle - \varphi(\|u\|) \} \\ &= \operatorname*{Sup}_{t \geq 0} \{ t\,\|u^*\|_* - \varphi(t) \} \\ &= (\text{as } \varphi \text{ is even}) \\ &= \operatorname*{Sup}_{t \in \mathbf{R}} \{ t\,\|u^*\|_* - \varphi(t) \} = \varphi^*(\|u^*\|_*). \end{aligned}$$

Remark 4.1. Let $\alpha, \alpha^* \in \,]1, \infty[$, satisfying

$$1/\alpha + 1/\alpha^* = 1.$$

It is easy to verify that the functions

$$\varphi(t) = \frac{1}{\alpha}|t|^\alpha, \qquad \varphi^*(t) = \frac{1}{\alpha^*}|t|^{\alpha^*}$$

belong to $\Gamma_0(\mathbf{R})$ and are conjugate. We therefore deduce that

$$F(u) = \frac{1}{\alpha}\|u\|^{\alpha} \quad \text{and} \quad G(u^*) = \frac{1}{\alpha^*}\|u^*\|_*^{\alpha^*}$$

are conjugate convex functions. ■

Remark 4.2. Proposition 4.2 is no longer valid if φ is not even. In that case we have

$$F^*(u^*) = \varphi_1^*(\|u^*\|_*) = \varphi_2^*(\|u^*\|_*),$$

where φ_i^* is the conjugate function of the function φ_i defined by ($i = 1, 2$):

$$\begin{aligned} &\varphi_1(t) = \varphi(|t|), && t \in \mathbf{R}, \\ &\varphi_2(t) = +\infty, && t < 0, \qquad \varphi_2(t) = \varphi(t), \qquad t \geqslant 0. \end{aligned}$$ ■

Remark 4.3. More generally, Proposition 4.2 is valid under the following hypothesis:

> for all $m > 0$, the function $t \to \varphi(t) - mt$ attains its minimum at a point $t \geqslant 0$.

This condition is satisfied if φ is even or equally if $\varphi \geqslant 0$ and $\varphi(0) = 0$. ■

5. SUBDIFFERENTIABILITY

5.1. Definition

Henceforth V will designate a l.c.s., V^* its topological dual, $\langle .,. \rangle$ the bilinear canonical pairing over $V \times V^*$ and F a mapping of V into $\bar{\mathbf{R}}$. We shall say that a continuous affine function ℓ everywhere less than F is exact at the point $u \in V$ if $\ell(u) = F(u)$. Necessarily, $F(u)$ will be finite and ℓ will have the form:

$$\begin{aligned} \ell(v) &= \langle v - u, u^* \rangle + F(u) \\ &= \langle v, u^* \rangle + F(u) - \langle u, u^* \rangle. \end{aligned}$$

Necessarily ℓ is maximal: its constant term is the greatest possible, whence:

$$F(u) - \langle u, u^* \rangle = -F^*(u^*). \tag{5.1}$$

Definition 5.1. A function F of V into $\bar{\mathbf{R}}$ is said to be *subdifferentiable* at the point $u \in V$ if it has a continuous affine minorant which is exact at u. The slope $u^* \in V^*$ of such a minorant is called a *subgradient of F* at u, and the set of subgradients at u is called the *subdifferential* at u and is denoted $\partial F(u)$.

If F is not subdifferentiable at u, we have $\partial F(u) = \varnothing$. We have the following characterization:

$$\text{(5.2)} \qquad u^* \in \partial F(u) \quad \text{if and only if} \quad F(u) \quad \text{is finite and} \\ \langle v - u, u^* \rangle + F(u) \leqslant F(v), \qquad \forall v \in V$$

If ℓ is a continuous affine function bounded above by F, ℓ is everywhere less than F^{**}, the Γ-regularization. If, furthermore, $\ell(u) = F(u)$, we obtain $\ell(u) \leqslant F^{**}(u) \leqslant F(u)$, whence $\ell(u) = F^{**}(u)$, from which these results follow:

$$\text{(5.3)} \qquad \textit{if} \quad \partial F(u) \neq \varnothing, \qquad F(u) = F^{**}(u),$$

$$\text{(5.4)} \qquad \textit{if} \quad F(u) = F^{**}(u), \qquad \partial F(u) = \partial F^{**}(u).$$

We now have a direct consequence of this definition; it already makes us anticipate the role of subdifferentiation in optimization problems:

$$\text{(5.5)} \qquad F(u) = \min_{v \in V} F(v) \quad \text{if and only if} \quad 0 \in \partial F(u).$$

We will now bring in polar functions. We take from (5.1) the following characterization:

Proposition 5.1. *Let F be a function of V into $\bar{\mathbf{R}}$ and F^* its polar. Then $u^* \in \partial F(u)$ if and only if:*

$$\text{(5.6)} \qquad F(u) + F^*(u^*) = \langle u, u^* \rangle.$$

Proof. The necessary condition has been established in (5.1). Conversely, if (5.6) is satisfied, the continuous affine function

$$\langle \cdot, u^* \rangle + F(u) - \langle u, u^* \rangle$$

is everywhere less than F (since its constant term is equal to $-F^*(u^*)$) and is exact at u. ■

Corollary 5.1. *The set $\partial F(u)$ (possibly empty) is convex and $\sigma(V^*, V)$-closed in V^*.*

Proof. By the definition (4.2) of F^*, we always have

$$F^*(u^*) - \langle u, u^* \rangle \geqslant -F(u).$$

Then Proposition 5.1 can be written:

$$\partial F(u) = \{ u^* \in V^* \mid F^*(u^*) - \langle u, u^* \rangle \leqslant -F(u) \}$$

and the second term is closed and convex since $F^* \in \Gamma(V^*)$. ■

Corollary 5.2. *For every function F of V into* $\bar{\mathbf{R}}$, *we have*

(5.7) $$u^* \in \partial F(u) \Rightarrow u \in \partial F^*(u^*).$$

If, furthermore, $F \in \Gamma(V)$ *we have*

(5.8) $$u^* \in \partial F(u) \Leftrightarrow u \in \partial F^*(u^*).$$

Proof. Since $F^{**} \leqslant F$, if $u^* \in \partial F(u)$, (5.6) entails that:

(5.9) $$F^{**}(u) + F^*(u^*) \leqslant \langle u, u^* \rangle.$$

Since the inverse inequality is always satisfied, (5.9) is in fact an equality, which means that $u \in \partial F^*(u^*)$.

If $F \in \Gamma(V)$, $F^{**} = F$, and (5.8) therefore is a result of (5.7). ∎

For convex functions, we have at our disposal a very simple criterion for subdifferentiability:

Proposition 5.2. *Let F be a convex function of V into* $\bar{\mathbf{R}}$, *finite and continuous at the point $u \in V$. Then $\partial F(v) \neq \varnothing$ for all $v \in \widehat{\mathring{\operatorname{dom}} F}$, and in particular $\partial F(u) \neq \varnothing$.*

Proof. Since F is finite and continuous at u, it is bounded above in a neighbourhood of u and so is finite and continuous at each point of $\widehat{\mathring{\operatorname{dom}} F}$ (Prop. 2.5). Thus it is sufficient to show that $\partial F(u) \neq \varnothing$.

Since F is convex, epi F is a convex subset of $V \times \mathbf{R}$. Since F is continuous, the interior of epi F is non-empty. To understand this, we need only take an open neighbourhood $\mathcal{O}$ of u over which F is bounded above by the constant $c \in \mathbf{R}$: the set $\mathcal{O} \times]c, +\infty[$ is an open subset of $V \times \mathbf{R}$ contained in epi F:

Since $(u, F(u))$ belongs to the boundary of epi F, we can separate it from $\widehat{\mathring{\operatorname{epi}} F}$ by a closed affine hyperplane (Cor. 1.1). We thus obtain a supporting hyperplane $\mathscr{H}$ of epi F, containing $(u, F(u))$. Let us write its equation:

$$\mathscr{H} = \{ (v, a) \in V \times \mathbf{R} \mid \langle v, u^* \rangle + \alpha a = \beta \} \text{ where } u^* \in V^*, \ \alpha \text{ and } \beta \in \mathbf{R}$$

where the coefficients, not all zero, are linked by:

$$\forall (v, a) \in \operatorname{epi} F, \qquad \langle v, u^* \rangle + \alpha a \geqslant \beta$$
$$\langle u, u^* \rangle + \alpha F(u) = \beta.$$

If $\alpha = 0$, we will have $\langle v - u, u^* \rangle \geqslant 0$ for all v in dom F, whence $u^* = 0$ since dom F is a neighbourhood of u. We thus have $\alpha > 0$, and dividing through by α:

$$\forall v \in \operatorname{dom} F, \qquad \frac{\beta}{\alpha} - \langle v, u^*/\alpha \rangle \leqslant F(v)$$

$$\frac{\beta}{\alpha} - \langle u, u^*/\alpha \rangle = F(u).$$

Finally,

$$\forall v \in V, \qquad \langle v - u, - u^*/\alpha \rangle + F(u) \leqslant F(v)$$

which proves that $-u^*/\alpha \in \partial F(u)$, which is therefore non-empty. ■

Remark 5.1. Corollary 6.1 below shows that a proper l.s.c. convex function F defined on a complete normed space is subdifferentiable "almost everywhere" (more precisely, over a dense subset) inside dom F.

5.2. Relation with Gâteaux-differentiability

We shall now complete our demonstration that, at least in the context of convex functions, subdifferentiability constitutes a generalization of differentiability.

Definition 5.2. Let F be a function of V into $\bar{\mathbf{R}}$. We call the limit as $\lambda \to 0_+$, if it exists, of

$$\frac{F(u + \lambda v) - F(u)}{\lambda} \tag{5.10}$$

the *directional derivative of F at u in the direction v* and denote it by $F'(u;v)$. If there exists $u^* \in V^*$ such that:

$$\forall v \in V, \qquad F'(u\,;v) = \langle v, u^* \rangle$$

we say that F is *Gâteaux-differentiable at u*, call u^* the *Gâteaux-differential at u of F*, and denote it by $F'(u)$.

The uniqueness of the Gâteaux-differential follows directly; it is characterized by:

$$\forall v \in V, \qquad \lim_{\lambda \to 0_+} \frac{F(u + \lambda v) - F(u)}{\lambda} = \langle v, F'(u) \rangle. \tag{5.11}$$

The case of convex functions is particularly interesting since the expression (5.10) is in that instance an increasing function of λ. Thus, when $\lambda \to 0_+$, *this expression always possesses a limit*, which, however, can be $\pm\infty$. We will show that essentially the case of Gâteaux-differentiability is the same as that of the uniqueness of the subgradient.

Proposition 5.3. *Let F be a convex function of V into $\bar{\mathbf{R}}$. If F is Gâteaux-differentiable at $u \in V$, it is subdifferentiable at u and $\partial F(u) = \{F'(u)\}$. Conversely, if at the point $u \in V$, F is continuous and finite and has only one subgradient, then F is Gâteaux-differentiable at u and $\partial F(u) = \{F'(u)\}$.*

Proof. If F is Gâteaux-differentiable at u, it is obvious that $F'(u) \in \partial F(u)$; if indeed $v \in V$ and $w = v - u$, we have

$$F(u + w) - F(u) \geqslant F'(u; w) = \langle w, F'(u) \rangle$$
$$F(v) - F(u) \geqslant \langle v - u, F'(u) \rangle.$$

Alternatively if $u^* \in \partial F(u)$, we will have for all $w \in V$ and all $\lambda > 0$:

$$F(u + \lambda w) - F(u) \geqslant \lambda \langle w, u^* \rangle$$

on dividing through by λ and passing to the limit we get:

$$\langle w, F'(u) \rangle \geqslant \langle w, u^* \rangle$$
$$\langle w, F'(u) - u^* \rangle \geqslant 0.$$

Since w is any point of V, $u^* = F'(u)$.

Let us turn to the converse. Since F is convex, we have for all v in V:

$$\forall \lambda \in \mathbf{R}, \qquad F(u) + \lambda F'(u; v) \leqslant F(u + \lambda v).$$

Geometrically, this means that in $V \times \mathbf{R}$, the straight line:

$$\mathscr{L} = \{ (u + \lambda v, F(u) + \lambda F'(u; v)) \mid \lambda \in \mathbf{R} \},$$

does not pass through the interior of epi F. But $\overset{\circ}{\widehat{\text{epi } F}}$ is an open convex set since epi F is convex, and it is non-empty since F is continuous and finite. From the Hahn-Banach theorem, there is a closed affine hyperplane $\mathscr{H}$ containing $\mathscr{L}$ which does not intersect $\overset{\circ}{\widehat{\text{epi } F}}$. It is easy to see that $\mathscr{H}$ is the graph of a continuous affine function everywhere less than F and exact at u. Since the subgradient u^* of F at u has been supposed unique, the slope of $\mathscr{H}$ is u^* and since $\mathscr{H}$ contains $\mathscr{L}$:

$$F'(u; v) = \langle v, u^* \rangle$$

which proves that F is Gâteaux-differentiable at u with differential u^*. ∎

The convexity of a Gâteaux-differentiable function may be characterized in the following way:

Proposition 5.4. *Let F be a Gâteaux-differentiable function of $\mathscr{A} \subset V$, $\mathscr{A}$ convex, into* $\mathbf{R}$. *Then the following are equivalent to each other:*

$$\text{(5.12)} \qquad F \text{ is convex over } \mathscr{A}$$

$$\text{(5.13)} \qquad F(v) \geqslant F(u) + \langle F'(u), v - u \rangle, \qquad \forall u, v \in \mathscr{A}.$$

Similarly, the following are equivalent to each other:

(5.14) $\qquad F$ *is strictly convex over* $\mathscr{A}$

(5.15) $\qquad F(v) > F(u) + \langle F'(u), v - u \rangle \qquad \forall u, v \in \mathscr{A}, u \neq v.$

Proof. From the preceding proposition, (5,12) implies (5.13). Conversely, the inequality (5.13) written with u and $(1-\lambda)u + \lambda v (u, v \in \mathscr{A}, \lambda \in]0,1[)$ necessitates

(5.18) $\qquad F(u) \geqslant F(u + \lambda(v - u)) + \lambda \langle F'(u + \lambda(v - u)), u - v \rangle.$

Similarly,

(5.19) $\quad F(v) \geqslant F(u + \lambda(v - u)) + (1 - \lambda) \langle F'(u + \lambda(v - u)), v - u \rangle.$

By multiplying (5.18) by $(1 - \lambda)$, (5.19) by λ and adding the inequalities, we obtain:

$$F((1 - \lambda)u + \lambda v) \leqslant (1 - \lambda)F(u) + \lambda F(v).$$

To show that (5.14) implies (5.15) we note that, as in Proposition 5.3,

$$F(u + \lambda(v - u)) < (1 - \lambda)F(u) + \lambda F(v),$$

if $u, v \in \mathscr{A}$ and $\lambda \in]0,1[$. Whence, F being convex:

$$\langle F'(u), v - u \rangle \leqslant \frac{F(u + \lambda(v - u)) - F(u)}{\lambda} < F(v) - F(u).$$

To show that (5.15) implies (5.14) we proceed as for the convexity, noting that (5.18) and (5.19) are strict inequalities if $u \neq v$. ∎

The convexity of a function is expressed by the monotonicity of its Gâteaux-differential:

Proposition 5.5. *Let F be a Gâteaux-differentiable function of $\mathscr{A} \subset V$, $\mathscr{A}$ convex, into* $\mathbf{R}$. *It will be convex if and only if its differential F' is a monotone mapping of V into V^*, i.e. if:*

(5.20) $\qquad \forall u_1$ *and* $u_2 \in V, \qquad \langle u_1 - u_2, F'(u_1) - F'(u_2) \rangle \geqslant 0.$

Proof. From Proposition 5.3, if F is convex, $F'(u_1)$ and $F'(u_2)$ are the subgradients of F at u_1 and at u_2:

$$\langle u_2 - u_1, F'(u_1) \rangle + F(u_1) \leqslant F(u_2)$$
$$\langle u_1 - u_2, F'(u_2) \rangle + F(u_2) \leqslant F(u_1).$$

Adding these terms together, we obtain (5.20).

Conversely, if F is Gâteaux-differentiable and if F' is monotone, for all u and v in V, the function ϕ of $[0,1]$ into $\mathbf{R}$ defined by

$$\phi(\lambda) = F(u + \lambda(v - u))$$

is differentiable with derivative:

$$\phi'(\lambda) = \langle v - u, F'(u + \lambda(v - u)) \rangle.$$

But, because of (5.20), ϕ' is increasing: thus ϕ is convex over $[0,1]$ and in particular

$$\phi(\lambda) \leqslant (1 - \lambda)\phi(0) + \lambda\phi(1), \qquad \forall\lambda \in [0, 1],$$

which is the desired inequality:

$$F((1 - \lambda)u + \lambda v) \leqslant (1 - \lambda)F(u) + \lambda F(v), \qquad \forall\lambda \in [0, 1]. \quad \blacksquare$$

5.3. Subdifferential calculus

It only remains for us to examine to what degree the ordinary differential calculus can be extended into a subdifferential calculus. Some results follow directly:

(5.21) *let* $F: V \to \overline{\mathbf{R}}$ *and* $\lambda > 0$. *At every point* $u \in V$, *we have* $\partial(\lambda F)(u) = \lambda\, \partial F(u)$,

(5.22) *let* F_1 *and* F_2: $V \to \mathbf{R}$. *At every point* $u \in V$, *we have* $\partial(F_1 + F_2)(u) \supset \partial F_1(u) + \partial F_2(u)$.

The equality in (5.22) is far from being always realized. Here, however, is a simple case where it holds:

Proposition 5.6. *If* F_1 *and* $F_2 \in \Gamma(V)$, *and if there is a point* $\bar{u} \in \operatorname{dom} F_1 \cap \operatorname{dom} F_2$ *where* F_1 *is continuous, we have*:

(5.23) $$\forall u \in V, \qquad \partial(F_1 + F_2)(u) = \partial F_1(u) + \partial F_2(u).$$

Proof. We have to show that the inverse inclusion of (5.22) is true, that is, that each $u^* \in \partial(F_1 + F_2)(u)$ can be decomposed into $u_1^* + u_2^*$, with $u_1^* \in \partial F_1(u)$ and $u_2^* \in \partial F_2(u)$. Our hypothesis is that F_1 and F_2 have finite values at u and that

(5.24) $$\forall v \in V, \qquad F_1(v) + F_2(v) \geqslant \langle v - u, u^* \rangle + F_1(u) + F_2(u).$$

Consider the convex sets in $V \times \mathbf{R}$:

$$C_1 = \{ (v, a) \mid F_1(v) - \langle v - u, u^* \rangle - F_1(u) \leqslant a \}$$
$$C_2 = \{ (v, a) \mid a \leqslant F_2(u) - F_2(v) \}.$$

The inequality (5.24) implies that they only have boundary points in common. But C_1 is the epigraph of the function G:

$$G(v) = F_1(v) - \langle v, u^* \rangle - F_1(u) + \langle u, u^* \rangle$$

which is convex and continuous at $\bar{u}$. Thus C_1 is a convex set with a non-empty interior. We can thus separate $\mathring{C}_1$ and C_2 by a closed affine hyperplane $\mathscr{H}$ (Cor. 1.1). It is easily verified that $\mathscr{H}$ is "non-vertical" (see Prop. 5.2) and is thus the graph of a continuous affine function.

$$v \to \langle v, v^* \rangle + \alpha, \quad \text{where} \quad v^* \in V^* \quad \text{and} \quad \alpha \in \mathbf{R}.$$

The separation can be written:

$$F_2(u) - F_2(v) \leqslant \langle v, v^* \rangle + \alpha \leqslant F_1(v) - \langle v - u, u^* \rangle - F_1(u), \qquad \forall v \in V.$$

In setting $v = u$, we obtain $\alpha = -\langle u, v^* \rangle$, and therefore:

$$\forall v \in V, \qquad \langle v - u, -v^* \rangle + F_2(u) \leqslant F_2(v)$$
$$\forall v \in V, \qquad \langle v - u, u^* + v^* \rangle + F_1(u) \leqslant F_1(v).$$

Hence $-v^* \in \partial F_2(u)$ and $u^* + v^* \in \partial F_1(u)$. Whence we have the desired decomposition, $u^* = u_1^* + u_2^*$, with $u_1^* = u^* + v^*$ and $u_2^* = -v^*$. ∎

Finally, after considering the subdifferential of a sum of functions, let us examine the subdifferential of a composite function. We take two l.c.s. V and Y with topological duals V^* and Y^* and a continuous linear mapping $\Lambda: V \to Y$ with transpose $\Lambda^*: Y^* \to V^*$. Let $F \in \Gamma(Y)$; the function $F \circ \Lambda: V \to \mathbf{R}$ belong to $\Gamma(V)$.

Proposition 5.7. *Let there be a point $\Lambda\bar{u}$ where F is continuous and finite. Then for all points u of V, we have*

$$\partial(F \circ \Lambda)(u) = \Lambda^* \partial F(\Lambda u). \tag{5.25}$$

Proof. Let $p^* \in \partial F(\Lambda u)$. By definition:

$$\forall p \in Y, \qquad \langle p - \Lambda u, p^* \rangle + F(\Lambda u) \leqslant F(p)$$

and *a fortiori*:

$$\forall v \in V, \qquad \langle \Lambda v - \Lambda u, p^* \rangle + F \circ \Lambda(u) \leqslant F \circ \Lambda(v)$$

$$\forall v \in V, \qquad \langle v - u, \Lambda^* p^* \rangle + F \circ \Lambda(u) \leqslant F \circ \Lambda(v).$$

Thus $\Lambda^* p^* \in \partial(F \circ \Lambda)(u)$ which proves that:

$$\Lambda^* \, \partial F(\Lambda u) \subset \partial(F \circ \Lambda)(u). \tag{5.26}$$

Conversely, take $u^* \in \partial(F \circ \Lambda)(u)$.

$$\forall v \in V, \qquad \langle v - u, u^* \rangle + F \circ \Lambda(u) \leqslant F \circ \Lambda(v). \tag{5.27}$$

Let us consider the affine subspace in $Y \times \mathbf{R}$.

$$\mathscr{L} = \{ (\Lambda v, \langle v - u, u^* \rangle + F \circ \Lambda(u)) \mid v \in V \}.$$

The inequality (5.27) shows that $\mathscr{L}$ and epi F only have boundary points in common. Since F has been assumed to be convex and continuous at $\Lambda\bar{u}$, $\overset{\circ}{\widehat{\text{epi}\, F}} \neq \varnothing$, and there is a closed affine hyperplane $\mathscr{H}$ containing $\mathscr{L}$ which does not intersect $\overset{\circ}{\widehat{\text{epi}\, F}}$. As usual, $\mathscr{H}$ is non-vertical, and therefore is the graph of a continuous affine function of Y into $\mathbf{R}$:

$$p \to \langle p, p^* \rangle + \alpha, \quad \text{where} \quad p^* \in Y^* \quad \text{and} \quad \alpha \in \mathbf{R}.$$

Since $\mathscr{H}$ contains $\mathscr{L}$:

$$\forall v \in V, \qquad \langle \Lambda v, p^* \rangle + \alpha = \langle v - u, u^* \rangle + F \circ \Lambda(u)$$

$$\alpha = F \circ \Lambda(u) - \langle u, u^* \rangle \tag{5.28}$$

$$\forall v \in V, \qquad \langle \Lambda v, p^* \rangle = \langle v, u^* \rangle. \tag{5.29}$$

Thus $u^* = \Lambda^* p^*$. Finally, since $\mathscr{H}$ does not intersect $\overset{\circ}{\widehat{\text{epi}\, F}}$:

$$\forall p \in Y, \qquad \langle p, p^* \rangle + F \circ \Lambda(u) - \langle u, \Lambda^* p^* \rangle \leqslant F(p)$$

$$\forall p \in Y, \qquad \langle p - \Lambda u, p^* \rangle + F \circ \Lambda(u) \leqslant F(p). \tag{5.30}$$

Thus $p^* \in \partial F(\Lambda u)$. Whence the desired result: $u^* = \Lambda^* p^* \in \Lambda^* \partial F(\Lambda u)$, and

$$\partial(F \circ \Lambda)(u) \subset \Lambda^* \, \partial F(\Lambda u). \quad \blacksquare \tag{5.31}$$

6. ε-SUBDIFFERENTIABILITY

Henceforth we shall assume that *V is a Banach space*.

6.1. An ordering relation over $V \times R$

Let us take a number $m > 0$, and consider in $V \times \mathbf{R}$ the closed convex cone $\mathscr{C}(m)$, with non-empty interior, defined by:

(6.1) $$\mathscr{C}(m) = \{ (u, a) \in V \times \mathbf{R} \mid a + m \|u\| \leqslant 0 \}.$$

We can associate with it an ordering relation over $V \times \mathbf{R}$ for which it will be the cone of positive elements. We shall denote if by $\leqslant$. By definition:

(6.2) $$(u, a) \leqslant (v, b) \Leftrightarrow (v - u, b - a) \in \mathscr{C}(m).$$

Proposition 6.1. *Let S be a closed subset of $V \times \mathbf{R}$ such that*

(6.3) $$\inf \{ a \mid (u, a) \in S \} > -\infty.$$

Then S has a maximal element for the given ordering.

Proof. It is sufficient to show that every totally ordered family of S has an upper bound in S; the conclusion will then follow from Zorn's lemma. Hence let $(a_i, u_i)_{i \in I}$ be a totally ordered family of S, and $\mathscr{F}$ the filter over S formed by its starting sections:

(6.4) $$A \in \mathscr{F} \Leftrightarrow \exists i \in I: \quad A \supset \{ (a_j, u_j) \mid j \geqslant i \}.$$

The family $(a_i)_{i \in I}$ is decreasing by (6.1) and, from (6.3), it is bounded. It is thus convergent in $\mathbf{R}$, and from the inequality which defines the given ordering:

(6.5) $$m \|u_i - u_j\| \leqslant |a_i - a_j| \qquad \forall i, j \in I$$

we deduce that $\mathscr{F}$ is a Cauchy filter over $V \times \mathbf{R}$. It therefore converges to a limit $(\bar{a}, \bar{u})$ which belongs to S since the latter is closed.

It only remains to show that $(\bar{a}, \bar{u})$ is the required maximal element. For this we take any $i \in I$ and write the inequality

(6.6) $$(a_j - a_i) + m \|u_j - u_i\| \leqslant 0 \qquad \forall j \geqslant i$$

and, passing to the limit in j:

(6.7) $$(\bar{a} - a_i) + m \|\bar{u} - u_i\| \leqslant 0.$$

Thus $(\bar{a}, \bar{u}) \geqslant (a_i, u_i)$ for all i, which establishes the result. ∎

6.2. Application to non-convex functions

Let F be a mapping of V into $\bar{\mathbf{R}}$ with $-\infty < \inf F < +\infty$. To say that $F(u) = \inf F$ amounts to saying that 0 is the subgradient of F at u. We can enquire what property of differentiability is related to the fact that $F(u) \leqslant \inf F + \varepsilon$.

Theorem 6.1. *Let F be a l.s.c. function of V into $\bar{\mathbf{R}}$, with $-\infty < \inf F < +\infty$, and let there be a point u where*

(6.8) $$F(u) \leqslant \inf F + \varepsilon.$$

For all $\lambda > 0$, there is a $u_\lambda \in V$ such that

$$\|u - u_\lambda\| \leqslant \lambda \text{ and } F(u_\lambda) \leqslant F(u) \tag{6.9}$$

$$\text{épi } F \cap \{ (u_\lambda, F(u_\lambda)) + \mathscr{C}(\varepsilon/\lambda) \} = (u_\lambda, F(u_\lambda)). \tag{6.10}$$

Proof. Let us apply Proposition 6.1 to the closed set $S = \text{epi } F$, for the ordering relation associated with the cone $\mathscr{C}(\varepsilon/\lambda)$. There is a maximal element (a_λ, u_λ) which is greater than $(F(u), u)$. Since (a_λ, u_λ) is maximal, $a_\lambda = F(u_\lambda)$ and (6.10) is satisfied.

To verify (6.9), we simply write $(F(u), u) \leqslant (F(u_\lambda), u_\lambda)$:

$$\frac{\varepsilon}{\lambda} \|u - u_\lambda\| \leqslant F(u) - F(u_\lambda).$$

The second term is bounded by ε due to (6.8). Whence we obtain (6.9). ∎

To understand this proposition more clearly, we may note that if F is furthermore Gâteaux-differentiable, condition (6.10) implies that, for all $v \in V$:

$$\forall t \in [0, 1], \qquad F(u_\lambda) - \frac{\varepsilon}{\lambda} t \|v\| \leqslant F(u_\lambda + tv) \tag{6.11}$$

$$-\frac{\varepsilon}{\lambda} \|v\| \leqslant \langle F'(u_\lambda), v \rangle \tag{6.12}$$

whence it immediately follows that:

$$\|F'(u_\lambda)\|_* \leqslant \varepsilon/\lambda. \tag{6.13}$$

If we further specify that λ be $\sqrt{\varepsilon}$, we obtain the following more striking corollary:

Corollary 6.1. *Let F be a Gâteaux-differentiable l.s.c. function of V into* **R** *and u a point where:*

$$F(u) \leqslant \inf F + \varepsilon. \tag{6.14}$$

Then there is a point v such that:

$$F(v) \leqslant F(u) \tag{6.15}$$

$$\|u - v\| \leqslant \sqrt{\varepsilon} \tag{6.16}$$

$$\|F'(v)\|_* \leqslant \sqrt{\varepsilon}. \tag{6.17}$$

6.3. Application to convex functions

Let V still be a Banach space, V^* its dual, $F \in \Gamma_0(V)$ and F^* its polar. We know (definition of F^*) that $\inf \{F(u) + F^*(u^*) - \langle u, u^* \rangle\} \geqslant 0$. If the minimum is attained for a pair (u, u^*), then $u \in \partial F^*(u^*)$ and $u^* \in \partial F(u)$. Otherwise, we arrive at the concept of ε-subdifferentiation.

Definition 6.1. *We call the set of $u^* \in V^*$ such that:*

(6.18) $$0 \leqslant F(u) + F^*(u^*) - \langle u, u^* \rangle \leqslant \varepsilon$$

the ε-subdifferential of F at the point $u \in V$, and denote it by $\partial_\varepsilon F(u)$.

The relation (6.18) being symmetrical in u and u^*

$$u^* \in \partial_\varepsilon F(u) \Leftrightarrow u \in \partial_\varepsilon F^*(u^*).$$

This also implies that the function $v \to \langle v - u, u^* \rangle + F(u) - \varepsilon$ is bounded above by F. We immediately deduce from this that, for all $\varepsilon > 0$, $\partial_\varepsilon F(u)$ is non-empty if and only if $F(u)$ is finite. The sets $\partial_\varepsilon F(u)$ decrease with ε and their intersection for $\varepsilon > 0$ is the subdifferential $\partial F(u)$.

The principal result concerning ε-subdifferentials is the following:

Theorem 6.2. *Let V be a Banach space and V^* its topological dual. Let $F \in \Gamma_0(V)$, F^* its polar, $u \in V$, $u^* \in V^*$ with $u^* \in \partial_\varepsilon F(u)$. For all $\lambda > 0$, there exists $u_\lambda \in V$ and $u^* \in V^*$ such that*

(6.19) $$\|u - u_\lambda\| \leqslant \lambda, \qquad \|u^* - u^*_\lambda\| \leqslant \varepsilon/\lambda,$$

(6.20) $$u^*_\lambda \in \partial F(u_\lambda).$$

In particular ($\lambda = \sqrt{\varepsilon}$), if $u^* \in \partial_\varepsilon F(u)$, there exists $u_\varepsilon \in V$ and $u^*_\varepsilon \in V^*$ such that

(6.21) $$\|u - u_\varepsilon\| \leqslant \sqrt{\varepsilon}, \qquad \|u^* - u^*_\varepsilon\| \leqslant \sqrt{\varepsilon},$$
$$u^*_\varepsilon \in \partial F(u_\varepsilon).$$

Proof. Consider the function over V

(6.22) $$G(v) = F(v) - \langle v, u^* \rangle + F^*(u^*).$$

By hypothesis (6.18), we have $G(u) \leqslant \inf G + \varepsilon$.
We can thus apply Theorem 6.1 to G.
There exists $u_\lambda \in V$ such that

(6.23) $$\|u - u_\lambda\| \leqslant \lambda$$

(6.24) $$\text{epi } G \cap \{ (u_\lambda, G(u_\lambda)) + \mathscr{C}(\varepsilon/\lambda) \} = (u_\lambda, G(u_\lambda)).$$

For greater simplicity, we shall denote by $\mathscr{C}_\lambda$ the cone $(u_\lambda, G(u_\lambda)) + \mathscr{C}(\varepsilon/\lambda)$ in $V \times \mathbf{R}$. It is a closed convex set with non-empty interior and epi G is a closed convex set. From (6.24) and the Hahn-Banach theorem, we can separate $\mathring{\mathscr{C}}_\lambda$ and epi G by a closed affine hyperplane $\mathscr{H}$ of $V \times \mathbf{R}$ with equation:

(6.25) $$\langle v, h^* \rangle + ar + b = 0, \qquad h^* \in V^*, \quad a \quad \text{and} \quad b \in \mathbf{R}$$

Since $\mathscr{C}_\lambda$ is closed and convex, it is the closure of its (non-empty) interior and so $\mathscr{H}$ also separates $\mathscr{C}_\lambda$ and epi G. We cannot allow $a = 0$ in (6.25) otherwise $\langle v, h^* \rangle + b$ would keep a constant sign over $\mathscr{C}_\lambda$ from which we would conclude that $h^* = 0$ and $b = 0$. We can thus restrict ourselves to the case where $a = 1$, which gives us

(6.26) $$\langle v, h^* \rangle + r + b \geqslant 0 \qquad \forall (v, r) \in \text{epi } G$$

(6.27) $$\langle v, h^* \rangle + r + b \leqslant 0 \qquad \forall (v, r) \in \mathscr{C}_\lambda.$$

Because of (6.24), the inequalities (6.26) and (6.27) are equalities at the point $(u_\lambda, G(u_\lambda))$, which gives us:

(6.28) $$b = -\langle u_\lambda, h^* \rangle - G(u_\lambda).$$

The relation (6.27) can then be written:

(6.29) $$\langle v - u_\lambda, h^* \rangle + r - G(u_\lambda) \leqslant 0 \qquad \forall (v, r) \in \mathscr{C}_\lambda.$$

Returning to the definition of $\mathscr{C}_\lambda$:

(6.30) $$\langle w, h^* \rangle + s \leqslant 0 \qquad \forall (w, s) \in \mathscr{C}(\varepsilon/\lambda).$$

Thus, taking $s = -\varepsilon/\lambda$, and returning to the definition (6.1) of $\mathscr{C}(\varepsilon/\lambda)$, we get:

(6.31) $$\sup_{\|w\| \leqslant 1} \langle w, h^* \rangle \leqslant \varepsilon/\lambda$$

(6.32) $$\|h^*\|_* \leqslant \varepsilon/\lambda.$$

It only remains to put $u_\lambda^* = u^* - h^* \in V^*$. From (6.32) we obtain:

(6.33) $$\|u^* - u_\lambda^*\|_* \leqslant \varepsilon/\lambda$$

and by substituting (6.22) and (6.28) into (6.26):

(6.34) $$\forall v \in V, \qquad \langle v, h^* \rangle + G(v) - \langle u_\lambda, h^* \rangle - G(u_\lambda) \geqslant 0$$

(6.35) $$\forall v \in V, \qquad F(v) - F(u_\lambda) - \langle v - u_\lambda, u_\lambda^* \rangle \geqslant 0.$$

But the latter demonstrates that $u_\lambda^* \in \partial F(u_\lambda)$, and so concludes the proof. ■

Corollary 6.2. *Let V be a Banach space, and $F \in \Gamma_0(V)$. The set of points where F is subdifferentiable is dense in* dom F.

Proof. Let $u_0 \in \text{dom } F$ and $\varepsilon > 0$ be fixed; since $F = F^{**}$,

$$F(u_0) = \operatorname*{Sup}_{u^* \in V^*} [\langle u_0, u^* \rangle - F^*(u^*)],$$

and there exists $u_0^* \in V^*$ such that:

$$\langle u_0, u_0^* \rangle - F^*(u_0^*) \geqslant F(u_0) - \varepsilon,$$

that is $u_0^* \in \partial_\varepsilon F(u_0)$. Applying Theorem 6.2 with $\lambda = \sqrt{\varepsilon}$, we deduce that F is subdifferentiable at a point $u_\varepsilon \in V$ such that $\|u_0 - u_\varepsilon\| \leqslant \sqrt{\varepsilon}$.

CHAPTER II

Minimization of Convex Functions and Variational Inequalities

Orientation

In this chapter we give some simple results related to the minimization of convex functions: the existence of the minimum, the characterization of solutions, etc. We also give analogous results for variational inequalities.

1. A RESULT CONCERNING EXISTENCE

We recall that a Banach space is *reflexive* if its unit ball is compact in the weak topology. This implies that every bounded sequence admits a weakly converging subsequence. Hilbert spaces and LP spaces ($1 < p < \infty$) are reflexive.

Let V be a *reflexive* Banach space (with norm $\|\ \ \|$) and $\mathscr{C}$ a non-empty *closed convex* subset of V. We take a function F of $\mathscr{C}$ into $\mathbf{R}$ and we assume that

(1.1) $\qquad F$ is convex, l.s.c. and proper.

We are concerned with the minimization problem

$$\operatorname*{Inf}_{u\in\mathscr{C}} F(u). \tag{1.2}$$

Any element $u \in \mathscr{C}$ such that:

$$F(u) = \operatorname*{Inf}_{v\in\mathscr{C}} F(v) \tag{1.3}$$

is termed a *solution* of the problem.

In certain cases it is preferable to replace the problem (1.2) by a minimization problem throughout the space V; for this the function $\hat{F}\colon V \to \overline{\mathbf{R}}$ is associated with F and with the convex set $\mathscr{C}$:

$$\hat{F}(u) = \begin{vmatrix} F(u) & \text{if} & u \in \mathscr{C}, \\ +\infty & \text{if} & u \notin \mathscr{C}. \end{vmatrix} \tag{1.4}$$

It is obvious from I(2.1) and I(2.9) that $\hat{F}$ is convex and l.s.c.; moreover, the problem:

$$\operatorname*{Inf}_{u\in V} \hat{F}(u) \tag{1.5}$$

is identical with the problem (1.2): the infimum is the same, as well as the set of solutions.

Proposition 1.1. *The set of solutions of* (1.2) *is a closed convex set* ($\subset \mathscr{C}$) *which is possibly empty.*

Proof. Apart from the trivial cases where the infimum in (1.2) (denoted by α) is equal to $\pm\infty$, we notice that the set of solutions of (1.2) is:

$$\{ u \in V \mid \hat{F}(u) \leqslant \alpha \}$$

and the result follows from I(2.3) and I(2.8). ∎

A simple criterion for the existence of solutions is the following:

Proposition 1.2. *Let us assume, in addition to* (1.1), *that*

(1.6) *the set* $\mathscr{C}$ *is bounded,*

or that the function F is coercive over $\mathscr{C}$, i.e. that:

(1.7) $$\lim F(u) = +\infty, \quad \textit{for} \quad u \in \mathscr{C}, \quad \|u\| \to \infty.$$

Then the problem (1.2) *has at least one solution. It has a unique solution if the function F is strictly convex over $\mathscr{C}$.*

Proof. Let u_n be a minimizing sequence of (1.2), that is, a sequence of elements of $\mathscr{C}$ such that:

$$F(u_n) \to \inf_{v \in \mathscr{C}} F(v) = \alpha.$$

Note that α belongs *a priori* to $[-\infty, +\infty[$; we will see from what follows that $\alpha \neq -\infty$. The sequence u_n is bounded in V: in the case of (1.6) this results from the fact that the set $\mathscr{C}$ is itself bounded, and in the case of the hypothesis (1.7) this results from the fact that the sequence $F(u_n)$ is bounded above. Thus we can extract from u_n a subsequence u_{n_i}, which converges weakly in V to an element u belonging to $\mathscr{C}$. From Corollary I.2.2 F is l.s.c. on $\mathscr{C}$ for the weak topology of V, and hence:

$$F(u) \leqslant \varliminf_{n_i \to \infty} F(u_{n_i}) = \alpha,$$

u is a solution of (1.2) and $\alpha \neq -\infty$.

If two different solutions u_1 and u_2 exist, then from Proposition 1.1, $(u_1 + u_2)/2$ is also a solution; if F is strictly convex this is impossible as in that case:

$$F\left(\frac{u_1 + u_2}{2}\right) < \frac{1}{2}(F(u_1) + F(u_2)) = \alpha. \quad ∎$$

Remark 1.1. Let $a(u,v)$ be a continuous bilinear form over V, coercive in the sense that

$$a(u, u) \geqslant \alpha \|u\|^2 \quad \forall u \in V, \quad \text{with} \quad \alpha > 0. \tag{1.8}$$

Given $\ell \in V^*$, there exists a unique u which achieves the minimum over $\mathscr{C}$ of the functional:

$$F(v) = a(v, v) - 2 \langle \ell, v \rangle. \tag{1.9}$$

This is a direct consequence of the proposition; we need to verify that the function

$$v \mapsto a(v, v)$$

is strictly convex. Now, from (1.8), $a(v - w, v - w) \geqslant 0$, whence

$$2a(v, w) \leqslant a(v, v) + a(w, w), \tag{1.10}$$

and for $\lambda \in]0, 1[$:

$$\begin{aligned} & a(\lambda v + (1 - \lambda)w, \lambda v + (1 - \lambda)w) \\ & = \lambda^2 a(v, v) + 2\lambda(1 - \lambda)a(v, w) + (1 - \lambda)^2 a(w, w) \\ & \leqslant \lambda a(v, v) + (1 - \lambda)a(w, w), \end{aligned}$$

and the equality is only possible if there is equality in (1.10), that is, if $v = w$.

For the coerciveness (1.7), we write:

$$\begin{aligned} F(v) = a(v, v) - 2 \langle \ell, v \rangle & \geqslant \alpha \|v\|^2 - 2 \|\ell\|_* \|v\| \\ & \geqslant \frac{\alpha}{2} \|v\|^2 - \frac{4}{\alpha} \|\ell\|_*^2 \end{aligned} \tag{1.11}$$

($\|.\|_*$ = the norm in V^*).

When $\mathscr{C}$ is bounded, (1.8) can be replaced by

$$a(u, u) \geqslant 0, \qquad \forall u \in V. \quad \blacksquare \tag{1.12}$$

2. CHARACTERIZATION OF SOLUTIONS

We wish to characterize the (or a) solution of (1.2) when the function F is differentiable or is the sum of a differentiable function and a non-differentiable function.

Proposition 2.1. *We assume that the function F satisfies* (1.1) *and is Gâteaux-differentiable with continuous derivative F'.*

Then if $u \in \mathscr{C}$, the following three conditions are equivalent to each other:

$$u \text{ is a solution of } (1.2), \tag{2.1}$$

$$\langle F'(u), v - u \rangle \geqslant 0 \qquad \forall v \in \mathscr{C}, \tag{2.2}$$

$$\langle F'(v), v - u \rangle \geqslant 0 \qquad \forall v \in \mathscr{C}. \tag{2.3}$$

Proof. (*a*) if u achieves the minimum in (1.2), then for any $v \in \mathscr{C}$ and $\lambda \in]0,1[$,

$$F(u) \leqslant F((1 - \lambda)u + \lambda v),$$

whence

$$\frac{1}{\lambda}[F(u + \lambda(v - u)) - F(u)] \geqslant 0. \tag{2.4}$$

When we take the limit as $\lambda \to 0$, the first term in (2.4) converges to

$$\langle F'(u), v - u \rangle$$

and we indeed obtain (2.2).

Conversely if u satisfies (2.2), then for $v \in \mathscr{C}$ and $\lambda \in]0,1[$

$$F(v) - F(u) \geqslant \frac{1}{\lambda}[F((1 - \lambda)u + \lambda v) - F(u)].$$

Taking the limit as $\lambda \to 0$, we find:

$$F(v) - F(u) \geqslant \langle F'(u), v - u \rangle \geqslant 0, \tag{2.5}$$

which shows that u is a solution of (1.2).

(*b*) We shall now show that (2.2) is equivalent to (2.3). F' is a monotone operator of V into V' (Prop. I.5.5):

$$\langle F'(v) - F'(u), v - u \rangle \geqslant 0, \qquad \forall u, v \in V. \tag{2.6}$$

If in particular u satisfies (2.2), by adding (2.2) and (2.6) we see that u satisfies (2.3).

If u now satisfies (2.3), by taking $v = (1 - \lambda)u + \lambda w$, $w \in \mathscr{C}$, $\lambda \in]0,1[$, we find that:

$$\lambda \langle F'((1 - \lambda)u + \lambda w), w - u \rangle \geqslant 0$$

$$\langle F'(u + \lambda(w - u)), w - u \rangle \geqslant 0. \tag{2.7}$$

The function $\lambda \mapsto \langle F'(u + \lambda(w - u)), w - u \rangle$ is the derivative of the scalar function $\lambda \mapsto F(u + \lambda(w - u))$; it is continuous and hence, when $\lambda \to 0$, (2.7) requires (2.2):

$$\langle F'(u), w - u \rangle \geqslant 0, \qquad \forall w \in \mathscr{C}. \ \blacksquare$$

Remark 2.1. For the situation described in Remark 1.1, the differential of the function F is given by:

$$\langle F'(u), v \rangle = 2 \{ a(u, v) - \langle \ell, v \rangle \}, \qquad \forall v \in V,$$

and thus the fact that u achieves the minimum of (1.9) in $\mathscr{C}$ is an equivalent condition to each of the following: $u \in \mathscr{C}$ and

$$a(u, v - u) - \langle \ell, v - u \rangle \geqslant 0, \qquad \forall v \in \mathscr{C}, \tag{2.8}$$

or

$$a(v, v - u) - \langle \ell, v - u \rangle \geqslant 0, \qquad \forall v \in \mathscr{C}. \tag{2.9}$$

Proposition 2.2. *Let us assume that $F = F_1 + F_2$, F_1 and F_2 being l.s.c. convex functions of $\mathscr{C}$ into $\mathbf{R}$, F_1 being Gâteaux-differentiable with differential F_1'.*

Then if $u \in \mathscr{C}$, the following three conditions are equivalent to each other:

$$u \text{ is a solution of } (1.2), \tag{2.10}$$

$$\langle F_1'(u), v - u \rangle + F_2(v) - F_2(u) \geqslant 0, \qquad \forall v \in \mathscr{C}, \tag{2.11}$$

$$\langle F_1'(v), v - u \rangle + F_2(v) - F_2(u) \geqslant 0, \qquad \forall v \in \mathscr{C}. \tag{2.12}$$

Proof. (*a*) If u achieves the minimum in (1.2), then:

$$F(u) \leqslant F((1 - \lambda)u + \lambda v), \qquad \forall v \in \mathscr{C}, \quad \forall \lambda \in]0, 1[,$$

and using the convexity of F_2,

$$F_1(u) + F_2(u) \leqslant F_1((1 - \lambda)u + \lambda v) + (1 - \lambda)F_2(u) + \lambda F_2(v),$$

$$\frac{1}{\lambda}[F_1((1 - \lambda)u + \lambda v) - F_1(u)] + F_2(v) - F_2(u) \geqslant 0;$$

making λ tend to zero, we obtain precisely (2.11).

Conversely if u satisfies (2.11), then by the convexity of F_1 (*cf.* (2.5)), we find that:

$$F_1(v) - F_1(u) - \langle F_1'(u), v - u \rangle \geqslant 0, \qquad \forall v \in \mathscr{C},$$

which when combined with (2.11) yields:

$$F(v) - F(u) \geqslant 0, \qquad \forall v \in \mathscr{C},$$

and u is a solution of (1.2).

(*b*) From the monotonicity of F_1':

$$\langle F_1'(v) - F_1'(u), v - u \rangle \geqslant 0,$$

and adding this inequality to (2.11) we obtain (2.12). Conversely if u satisfies (2.12), then as in Proposition 2.1, taking $v = (1-\lambda)u + \lambda w$, $w \in \mathscr{C}$, $\lambda \in]0,1[$, we find that:

$$\lambda \langle F_1'((1-\lambda)u + \lambda w), w - u \rangle + F_2((1-\lambda)u + \lambda w) - F_2(u) \geqslant 0;$$

by the convexity of F_2:

$$\lambda \langle F_1'((1-\lambda)u + \lambda w), w - u \rangle + \lambda F_2(w) - \lambda F_2(u) \geqslant 0.$$

Dividing by λ and letting $\lambda \to 0$, we obtain (2.11) with v replaced by w. ■

Remark 2.2. Proposition 2.2 clearly contains Proposition 2.1.

Example: Proximity mappings

We shall assume for this example that V is a Hilbert space (with scalar product $((.,.))$), and that φ is a proper l.s.c. convex function of V into $\mathbf{R}$.[1] Let:

$$F = F_1 + F_2, \qquad F_2(u) = \varphi(u),$$
$$F_1(u) = \tfrac{1}{2}\|u - x\|^2, \qquad x \in V \text{ given.}$$

Since the function F_1 is strictly convex and l.s.c. (Remark 1.1), it is obvious that F is strictly convex l.s.c. Moreover, the function F is coercive (property (1.7)) over V: indeed, since $\varphi \in \Gamma_0(V)$, φ is bounded from below by a continuous affine function which can be written:

$$((y, u)) + \alpha, \qquad \alpha \in \mathbf{R},$$

whence

$$F(u) \geqslant \tfrac{1}{2}\|u - x\|^2 + ((y, u)) + \alpha$$

$$F(u) \geqslant \tfrac{1}{2}\|u + y - x\|^2 - \tfrac{1}{2}\|y - x\|^2 + \tfrac{1}{2}\|x\|^2 + \alpha,$$

and we deduce that $F(u) \to +\infty$ for $\|u\| \to \infty$, $u \in V$.

Thus we can apply Proposition 1.2 with $\mathscr{C} = V$: there is a unique element u of V which achieves the minimum of:

$$F(v) = \tfrac{1}{2}\|v - x\|^2 + \varphi(v).$$

Proposition 2.2 implies that u is characterized by one or other of the following conditions:

[1] $\varphi \in \Gamma_0(V)$ in the terminology of Chapter I.

(2.13) $\quad ((u - x, v - u)) + \varphi(v) - \varphi(u) \geqslant 0, \qquad \forall v \in V,$

(2.14) $\quad ((v - x, v - u)) + \varphi(v) - \varphi(u) \geqslant 0, \qquad \forall v \in V.$

The mapping $x \to u = u(x)$ of V into itself, introduced by Moreau [2], is termed a *proximity mapping* (with respect to φ) and we write

(2.15) $\quad u = \operatorname{prox} x \quad \text{or} \quad u = \operatorname{prox}_\varphi x.$

As a special case of proximity mapping, when

(2.16) $\quad \varphi = \chi_{\mathscr{C}} =$ indicator function of a closed convex subset $\mathscr{C}$ of V,

we find the projection operator over $\mathscr{C}$. The equivalent conditions (2.13) to (2.15) then become

(2.17) $\quad u \in \mathscr{C}$ and $((u - x, v - u)) \geqslant 0, \qquad \forall v \in \mathscr{C},$

(2.18) $\quad u \in \mathscr{C}$ and $((v - x, v - u)) \geqslant 0, \qquad \forall v \in \mathscr{C},$

(2.19) $\quad u = \Pi_{\mathscr{C}} x =$ projection of x onto $\mathscr{C}$.

Remark 2.3. Inequalities such as (2.2), (2.3), (2.11) and (2.12) are called *variational inequalities* and they arise naturally in problems of minimization of convex functions. Propositions 1.2, 2.1 and 2.2, when combined, provide existence theorems for the solution of certain variational inequalities (e.g. proximity mappings). In the following section we shall give a direct solution of certain variational inequalities of a different type.

3. DIRECT STUDY OF CERTAIN VARIATIONAL INEQUALITIES

Let us again assume that V is a reflexive Banach space (with norm $\|.\|$) and let V^* be its dual (with norm $\|.\|_*$).

We take a mapping A of V into V^* and a function φ of V into $\overline{\mathbf{R}}$ which satisfies

(3.1) $\quad \varphi$ is a proper convex l.s.c. function.

We are now concerned with the existence of elements $u \in V$ which are solutions of the variational inequality:

(3.2) $\quad \langle Au - f, v - u \rangle + \varphi(v) - \varphi(u) \geqslant 0, \qquad \forall v \in V$

for f given in V^*.

The hypotheses about A are as follows:

(3.3) $\quad A$ is weakly continuous over the subspaces of finite dimension of V.

(3.4) A is monotone: $\langle Au - Av, u - v \rangle \geqslant 0, \qquad \forall u, v \in V.$

(3.5) There exists $v_0 \in \operatorname{dom} \varphi$ such that

$$\frac{\langle Av, v - v_0 \rangle + \varphi(v)}{\|v\|} \to +\infty \quad \text{if} \quad \|v\| \to \infty.$$

We then have:

Theorem 3.1. *Let A and φ satisfy* (3.1) *and* (3.3)–(3.5). *For f given in V^*, there is at least one $u \in V$ satisfying* (3.2).

The theorem will be proved in several stages.

Lemma 3.1. *Theorem* 3.1 *is true if we assume in addition that V is of finite dimension and that* dom φ *is bounded.*

Proof. Since V is of finite dimension, we can provide this space with a Hilbert structure with scalar product $((.,.))$ and *identify* V and V^* in such a way that the duality pairing can be identified with the scalar product of V. If u is a solution of (3.2), we then have:

(3.6) $$((Au - f, v - u)) + \varphi(v) - \varphi(u) \geqslant 0, \qquad \forall v \in V$$

or alternatively:

(3.7) $$((u - (u + f - Au), v - u)) + \varphi(v) - \varphi(u) \geqslant 0, \qquad \forall v \in V,$$

which, from (2.13) and (2.15) is equivalent to

(3.8) $$u = \operatorname{prox}_\varphi (u + f - Au).$$

Since the mapping $\operatorname{prox}_\varphi$ is valued in dom φ which is a closed convex set bounded in V, the existence of a solution of (3.8) is an immediate consequence of Brouwer's fixed-point theorem provided that we can show that the mapping

(3.9) $$u \mapsto \operatorname{prox}_\varphi (u + f - Au)$$

is continuous. Now from (3.3) the mapping A and hence the mapping $u \to u + f - Au$ is continuous in V.

It is therefore sufficient to show that the mapping $\operatorname{prox}_\varphi$ is continuous. But if $f_1, f_2 \in V$, $u_1 = \operatorname{prox}_\varphi f_1$, $u_2 = \operatorname{prox}_\varphi f_2$, the relations (2.13) require that:

$$((u_1 - f_1, u_2 - u_1)) + \varphi(u_2) - \varphi(u_1) \geqslant 0,$$
$$((u_2 - f_2, u_1 - u_2)) + \varphi(u_1) - \varphi(u_2) \geqslant 0,$$

and by addition:

(3.10) $$\|u_1 - u_2\|^2 \leqslant ((f_1 - f_2, u_1 - u_2)) \leqslant \|f_1 - f_2\| \, \|u_1 - u_2\|,$$

$$\|u_1 - u_2\| \leqslant \|f_1 - f_2\|, \tag{3.11}$$

which proves the continuity of prox_φ, and concludes the proof of the lemma. ∎

Lemma 3.2. *Theorem* 3.1 *is valid if we make the single additional hypothesis that V is of finite dimension.*

Proof. Let $r > 0$ and φ_r be a function of V into $\overline{\mathbf{R}}$ defined by:

$$\varphi_r(u) = \begin{vmatrix} \varphi(u) & \text{if} & \|u\| \leqslant r \\ +\infty & \text{if} & \|u\| > r. \end{vmatrix} \tag{3.12}$$

Since φ_r is the sum of φ and of the indicator function of the ball $\{\|u\| \leqslant r\}$, it is clear that φ_r is convex and l.s.c. and that it is proper for r sufficiently large (for dom φ_r not to be empty).

Lemma 3.1 implies the existence of $u_r \in V$, $\|u_r\| \leqslant r$, satisfying:

$$\langle Au_r - f, v - u_r \rangle + \varphi_r(v) - \varphi_r(u_r) \geqslant 0, \qquad \forall v \in V. \tag{3.13}$$

For r sufficiently large ($r \geqslant \|v_0\|$), we can put $v = v_0$ in (3.13) and we find that

$$\langle Au_r - f, v_0 - u_r \rangle + \varphi(v_0) - \varphi(u_r) \geqslant 0.$$

$$\frac{1}{\|u_r\|}\{\langle Au_r, u_r - v_0 \rangle + \varphi(u_r)\} \leqslant \frac{1}{\|u_r\|}\{-\langle f, v_0 - u_r \rangle + \varphi(v_0)\}. \tag{3.14}$$

As the right-hand side of (3.14) is bounded independently of r, it follows from hypothesis (3.5) that the u_r's are bounded in V. Thus we can extract from u_r a sequence u_{r_i}, $r_i \to +\infty$, which converges in V to an element u (V is of finite dimension). Using (3.3) it is easy to take the limit in (3.13) (V being of finite dimension) and this shows that u is a solution of (3.2) (note that $\varphi_{r_i}(v) = \varphi(v)$ since $r_i \geqslant \|v\|$ and that $\varphi_r(u_r) = \varphi(u_r)$).

Proof of Theorem 3.1. Let $\mathscr{V}$ be the family of finite-dimensional subspaces of V which contain v_0.

For every $V_m \in \mathscr{V}$, the application of Lemma 3.2 shows that there exists $u_m \in V_m$ such that

$$\langle Au_m - f, v - u_m \rangle + \varphi(v) - \varphi(u_m) \geqslant 0, \qquad \forall v \in V_m. \tag{3.15}$$

Making $v = v_0 \in V$, we establish an inequality analogous to (3.14) which requires that the family u_m is bounded in V. Under these conditions, we have, for an ultrafilter $\mathscr{U}$ which is finer than $\mathscr{V}$,

$$u_m \to u \qquad \text{weakly in } V. \tag{3.16}$$

Writing (3.15) with $v = u$ and taking the upper limit, we find

$$\overline{\lim} \langle Au_m, u_m - u \rangle \leqslant \overline{\lim} [\varphi(u) - \varphi(u_m)].$$

Since the function φ is convex and l.s.c.,

$$\overline{\lim} [\varphi(u) - \varphi(u_m)] = \varphi(u) - \underline{\lim}\, \varphi(u_m) \leqslant 0,$$

which implies that:

$$\overline{\lim} \langle Au_m, u_m - u \rangle \leqslant 0. \tag{3.17}$$

Let us admit for a moment the following lemma

Lemma 3.3. *If u_m converges to u weakly in V and satisfies* (3.17), *then*:

$$\underline{\lim} \langle Au_m, u_m - v \rangle \geqslant \langle Au, u - v \rangle, \qquad \forall v \in V. \quad \blacksquare \tag{3.18}$$

With this result we can take the lower limit in (3.15) ($v \in V$ fixed):

$$\underline{\lim} \langle Au_m - f, u_m - v \rangle \leqslant \overline{\lim} [\varphi(v) - \varphi(u_m)] \leqslant \varphi(v) - \underline{\lim}\, \varphi(u_m).$$

The lower semi-continuity of φ and (3.18) yield the inequality

$$\langle Au - f, u - v \rangle \leqslant \varphi(v) - \varphi(u),$$

which shows that u is a solution of (3.2). $\blacksquare$

Proof of Lemma 3.3. This has been given by Brezis [1] and exploits an idea of G. J. Minty [1]. Since A is monotone

$$\langle Au_m, u_m - u \rangle \geqslant \langle Au, u_m - u \rangle$$

and we find that $\underline{\lim} \langle Au_m, u_m - u \rangle \geqslant 0$ and hence

$$\lim \langle Au_m, u_m - u \rangle = 0.$$

To establish (3.18), we consider $w = (1 - \lambda)u + \lambda v$, $\lambda \in]0, 1[$ and, because of the monotonicity, we write

$$\langle Au_m - Aw, u_m - w \rangle \geqslant 0,$$

therefore

$$\lambda \langle Au_m, u - v \rangle \geqslant - \langle Au_m, u_m - u \rangle + \langle Aw, u_m - u \rangle - \lambda \langle Aw, v - u \rangle$$

and taking the lower limit, it follows that:

$$\lambda\, \underline{\lim} \langle Au_m, u - v \rangle \geqslant \lambda \langle Aw, u - v \rangle.$$

Dividing by λ and letting $\lambda \to 0$, from (3.3) we obtain

$$\langle Aw, v - u \rangle = \langle A((1 - \lambda)u + \lambda v), v - u \rangle \to \langle Au, v - u \rangle,$$

and we thus have (3.18) in the limit. ∎

Remark 3.1. Property (3.4) was only used to prove Lemma 3.3. Having noted this, we can replace Hypothesis (3.4) by

$$(3.19) \qquad \left| \begin{array}{l} \text{If } u_m \text{ converges weakly to } u \text{ in } V \text{ with} \\ \overline{\lim} \langle Au_m, u_m - u \rangle \leqslant 0, \\ \text{we have} \\ \underline{\lim} \langle Au_m, u_m - v \rangle \geqslant \langle Au, u - v \rangle, \qquad \forall v \in V. \end{array} \right.$$

An operator A of V into V^* which satisfies (3.19) is called *pseudo-monotone*; *cf.* H. Brezis [1], J. L. Lions [3].

Remark 3.2. We find in Lions [3] many existence theorems for more general variational inequalities than the foregoing.

Special cases.

From Proposition 3.1 we can deduce several special cases.

Proposition 3.1. *Let $\mathscr{C}$ be a closed convex set of V, and let A satisfy* (3.3)–(3.5). *For all f of V^*, there is a $u \in \mathscr{C}$ such that*

$$(3.20) \qquad \langle Au - f, v - u \rangle \geqslant 0, \qquad \forall v \in \mathscr{C}.$$

Proof. We apply Theorem 3.1, with $\varphi =$ the indicator function of $\mathscr{C}$.

Proposition 3.2. *Let A be an operator of V into V^* which satisfies* (3.3)–(3.5). *For all f of V^* there is a $u \in V$ such that*

$$(3.21) \qquad Au = f.$$

Proof. We apply Proposition 3.1 with $\mathscr{C} = V$; setting $v = u + w$ and $v = u - w$, $w \in V$, in (3.20), we find

$$\langle Au - f, w \rangle = 0, \qquad \forall w \in V,$$

and (3.21) follows directly.

Proposition 3.3. *Let V be a Hilbert space, $A \in \mathscr{L}(V, V^*)$ satisfying*

$$(3.22) \qquad \langle Au, u \rangle \geqslant \alpha \|u\|^2, \qquad \forall u \in V, \quad \alpha > 0.$$

For all $f \in V^*$ there is a unique $u \in V$ satisfying

$$Au = f. \tag{3.23}$$

Proof. We apply Proposition 3.2; A clearly satisfies (3.3) and (3.4); (3.5) results from (3.22). The existence being thus obtained, the uniqueness follows easily from (3.22). ■

Remark 3.2. (Interpretation of variational inequalities with sub-differentials.)

If u satisfies (3.2) then by definition (*cf.* I(5.1))

$$f - Au \in \partial\varphi(u). \tag{3.24}$$

In particular, if u satisfies (3.20),

$$f - Au \in \partial\chi_{\mathscr{C}}(u), \tag{3.25}$$

$\chi_{\mathscr{C}}$ = the indicator function of $\mathscr{C}$.

Remark 3.3. If $a(u,v) = \langle Au, v\rangle$ is moreover a symmetric bilinear form over $V \times V$, the problems (1.9) and (3.23) are equivalent; this is an immediate result of (2.2).

Proposition 3.3 is the classic Lax-Milgram lemma, known also as the projection lemma.

Remark 3.4. The operator A is said to be strictly monotone if it satisfies, instead of (3.4),

$$\langle Au - Av, u - v \rangle > 0, \qquad \forall u, v \in V, \quad u \neq v. \tag{3.26}$$

In this case, there is at most one u which satisfies (3.2) and exactly one under the hypotheses of Theorem 3.1.

CHAPTER III

Duality in Convex Optimization

Orientation

In this chapter we shall associate to a minimization problem $(\mathscr{P})$ a maximization problem $(\mathscr{P}^*)$ termed the dual problem of $\mathscr{P}$ and we shall examine the relationship between these two problems (the comparison of the infimum with the supremum, and the relationship between the solutions in particular). ■

1. THE PRIMAL PROBLEM AND THE DUAL PROBLEM

Let V and V^* be two topological vector spaces placed in duality by the bilinear pairing $\langle .,. \rangle_V$. The elements of V will be denoted by $u, v, w, \ldots$, and those of V^* will be denoted by $u^*, v^*, w^*, \ldots$

Taking a function F of V into $\overline{\mathbf{R}}$, we are concerned with the minimization problem

$$(\mathscr{P}) \qquad \operatorname*{Inf}_{u \in V} F(u).$$

Problem $\mathscr{P}$ will be termed the primal problem. The infimum for problem $\mathscr{P}$ will be denoted by *inf* $\mathscr{P}$ and, as in Chapter II, every element u of V such that

$$F(u) = \inf \mathscr{P} \tag{1.1}$$

will be termed a *solution* of $\mathscr{P}$.

Problem $\mathscr{P}$ will be said to be *non-trivial* if there exists $u_0 \in V$ such that

$$F(u_0) < +\infty. \tag{1.2}$$

Note that the function F is arbitrary at the moment; later we will pay particular attention to the case where $F \in \Gamma_0(V)$; in this case problem $\mathscr{P}$ is clearly non-trivial.

Perturbed problems

Let us assume to be given, in the following way, a family of perturbations of problem $\mathscr{P}$.

Let there be given two other Hausdorff topological vector spaces Y and Y^* placed in duality by a bilinear pairing $\langle .,. \rangle$. Since in general there is no possibility of ambiguity, the pairing between V and V^* and the pairing between Y and Y^* will both be denoted by $\langle .,. \rangle$. The elements of Y will be denoted by p, q, r, and those of Y^* by $p^*, q^*, r^*, \ldots$

We shall also consider a function (denoted by Φ) of $V \times Y$ into $\bar{\mathbf{R}}$ such that

$$\Phi(u, 0) = F(u), \tag{1.3}$$

and for every $p \in Y$ we shall consider the minimization problem

$$(\mathscr{P}_p) \qquad \operatorname*{Inf}_{u \in V} \Phi(u, p).$$

Clearly for $p = 0$, $\mathscr{P}_0$ is none other than problem $\mathscr{P}$. The problems $\mathscr{P}_p$ will be said to be perturbed problems of $\mathscr{P}$ (with respect to the given perturbations).

We shall describe in Sections 4 and 5 the main types of useful perturbations.

The dual problem with respect to the given perturbations

Given problem $\mathscr{P}$ and the perturbed problems $\mathscr{P}_p$, we are now able to define a dual problem. For this let $\Phi^* \in \Gamma(V^* \times Y^*)$,

$$\Phi^* : V^* \times Y^* \to \bar{\mathbf{R}}, \tag{1.4}$$

be the conjugate function of Φ in the duality between $V \times Y$ and $V^* \times Y^*$ (*cf.* Chap. I).[1] The problem:

$$(\mathscr{P}^*) \qquad \operatorname*{Sup}_{p \in Y^*} \{ -\Phi^*(0, p^*) \}$$

is termed the *dual problem* of $\mathscr{P}$ with respect to Φ (or with respect to the given perturbations). The supremum for problem $\mathscr{P}^*$ is denoted sup $\mathscr{P}^*$ and any element of p^* of Y^* such that

$$-\Phi^*(0, p^*) = \operatorname{Sup} \mathscr{P}^* \tag{1.5}$$

is termed a solution of $\mathscr{P}^*$.

The first relationship between $\mathscr{P}$ and $\mathscr{P}^*$ is the following.

[1] The pairing between $V \times Y$ and $V^* \times Y^*$ is written, classically, as:

$$\langle (u^*, p^*), (u, p) \rangle_{V \times Y} = \langle u^*, u \rangle_V + \langle p^*, p \rangle_Y.$$

Proposition 1.1.

$$-\infty \leqslant \operatorname{Sup} \mathscr{P}^* \leqslant \operatorname{Inf} \mathscr{P} \leqslant +\infty. \tag{1.6}$$

Proof. Let $p^* \in Y^*$; by definition (*cf.* I(4.2))

$$\Phi^*(0, p^*) = \operatorname*{Sup}_{\substack{u \in V \\ p \in Y}} [\langle p^*, p \rangle - \Phi(u, p)];$$

in particular for every $u \in V$:

$$\Phi^*(0, p^*) \geqslant \langle p^*, 0 \rangle - \Phi(u, 0),$$

$$-\Phi^*(0, p^*) \leqslant \Phi(u, 0). \tag{1.7}$$

The relationship (1.7) which is valid for any $p^* \in Y^*$ and $u \in V$ implies $\sup \mathscr{P}^* \leqslant \inf \mathscr{P}$. ∎

Remark 1.2. Various counter examples show that all the inequalities appearing in (1.6) can effectively be strict inequalities or equalities. *Cf.* Rockafellar [7].

Proposition 1.2. *If problem $\mathscr{P}$ is non-trivial, then*

$$\sup \mathscr{P}^* \leqslant \inf \mathscr{P} < +\infty. \tag{1.8}$$

If problem $\mathscr{P}^*$ is non-trivial, then

$$-\infty < \sup \mathscr{P}^* \leqslant \inf \mathscr{P}. \tag{1.9}$$

If $\mathscr{P}$ and $\mathscr{P}^*$ are non-trivial, $\inf \mathscr{P}$ and $\sup \mathscr{P}^*$ are finite

$$-\infty < \sup \mathscr{P}^* \leqslant \inf \mathscr{P} < +\infty. \tag{1.10}$$

Proof. If $\mathscr{P}$ is non-trivial, there exists $u_0 \in V$ such that $F(u_0) = \Phi(u_0, 0) < +\infty$ and thus with (1.6):

$$\sup \mathscr{P}^* \leqslant \inf \mathscr{P} \leqslant \Phi(u_0, 0) < +\infty.$$

The proof for (1.9) is analogous; (1.10) results from (1.8) and (1.9).

Reiteration of duality

The technique used to form the dual of a minimization problem can easily be extended to a maximization problem, if we note that:

$$\sup_{u \in V} [-G(u)] = -\inf_{u \in V} G(u).$$

In particular it is natural to associate the perturbed problems ($u^* \in V^*$)

$$(\mathscr{P}^*_{u^*}) \qquad \operatorname*{Sup}_{p^* \in Y^*} \{-\Phi^*(u^*, p^*)\},$$

with the dual problem $\mathscr{P}^*$ and to determine the dual problem of $\mathscr{P}^*$ with respect to these perturbations; we easily arrive at the following problem which we will term the bidual problem of $\mathscr{P}$:

$$(\mathscr{P}^{**}) \qquad \operatorname*{Inf}_{u \in V} \{ \Phi^{**}(u, 0) \},$$

where Φ^{**} is the conjugate function of Φ^*, i.e. the Γ-regularization of $\Phi (\Phi^{**} \in \Gamma(V \times Y))$.

After this, we can no longer repeat the dualization process; indeed the natural perturbations of $\mathscr{P}^{**}$ are the problems

$$\operatorname*{Inf}_{u \in V} \Phi^{**}(u, p)$$

and the dual problem of $\mathscr{P}^{**}$ with respect to these perturbations is the problem

$$\operatorname{Sup} \{ -\Phi^{***}(0, p^*) \}$$

and the latter problem is identical to problem $\mathscr{P}^*$ since $\Phi^{***} = \Phi^*$ (Cor. I.4.1).

If problem $\mathscr{P}^{**}$ is identical to problem $\mathscr{P} (\Phi^{**}(u,0) = \Phi(u,0),\ \forall u \in V)$, each of the problems $\mathscr{P}$ and $\mathscr{P}^*$ is found to be the dual of the other and there is thus complete symmetry between primal and dual problems. This will certainly be the case if $\Phi^{**} = \Phi$ which is the same as

$$(1.11) \qquad \Phi \in \Gamma_0(V \times Y).$$

In particular, this implies that if F is not identically equal to $+\infty$,

$$(1.12) \qquad F \in \Gamma_0(V)$$

(in fact, if $\Phi \in \Gamma_0(V \times Y)$, $u \mapsto \Phi(u,0) = F(u)$ is convex and l.s.c. with values in $]-\infty, +\infty]$, thus it belongs to $\Gamma_0(V)$; Prop. I.3.1).

This is one of the reasons why situations where hypothesis (1.11) is satisfied are of particular interest to us. This symmetry between $\mathscr{P}$ and $\mathscr{P}^*$ does not however imply that $\inf \mathscr{P} = \sup \mathscr{P}^*$ and the only general conclusion remains (1.10); the equality $\inf \mathscr{P} = \sup \mathscr{P}^*$ will appear in Section 2 as a kind of regularity property of the problem under consideration.

2. NORMAL PROBLEMS AND STABLE PROBLEMS

In the following sections we shall assume in general that

$$(2.1) \qquad \Phi \in \Gamma_0(V \times Y).$$

For $p \in Y$ let

$$(2.2) \qquad h(p) = \inf \mathscr{P}_p = \operatorname*{Inf}_{u \in V} \Phi(u, p).$$

Lemma 2.1. *Under hypothesis* (2.1), *the function* $h\colon Y \mapsto \overline{\mathbf{R}}$ *is convex.*

Proof. Let $p, q \in Y$ and $\lambda \in]0,1[$. We have to show that

$$h(\lambda p + (1 - \lambda)q) \leqslant \lambda h(p) + (1 - \lambda)h(q), \tag{2.3}$$

whenever the right-hand side is defined (*cf.* Definition I.2.1). Thus (2.3) is obvious (or there is nothing to prove) if $h(p)$ or $h(q) = +\infty$. Let us therefore assume that $h(p) < +\infty$ and $h(q) < +\infty$. For every $a > h(p)$ (resp. for every $b > h(q)$) there is a $u \in V$ (resp. $v \in V$) such that

$$h(p) \leqslant \Phi(u, p) \leqslant a,$$
$$h(q) \leqslant \Phi(v, q) \leqslant b.$$

Then:

$$\begin{aligned} h(\lambda p + (1 - \lambda)q) &= \operatorname*{Inf}_{w \in V} \Phi(w, \lambda p + (1 - \lambda)q) \\ &\leqslant \Phi(\lambda u + (1 - \lambda)v, \lambda p + (1 - \lambda)q) \\ &\leqslant \text{(by the convexity of } \Phi) \\ &\leqslant \lambda\Phi(u, p) + (1 - \lambda)\Phi(v, q) \\ &\leqslant \lambda a + (1 - \lambda)b. \end{aligned}$$

If we let a decrease towards $h(p)$ and b decrease towards $h(q)$, the inequality

$$h(\lambda p + (1 - \lambda)q) \leqslant \lambda a + (1 - \lambda)b$$

yields (2.3) in the limit. ■

Remark 2.1. Lemma 2.1 remains valid under the single hypothesis that Φ is convex.

Remark 2.2. In general $h \notin \Gamma_0(Y)$ (in spite of hypothesis (2.1)). ■

We can associate with the function $h \in \overline{\mathbf{R}}^Y$ its conjugate function $h^* \in \Gamma_0(Y^*)$; we have:

Lemma 2.2.

$$h^*(p^*) = \Phi^*(0, p^*), \qquad \forall p^* \in Y^*. \tag{2.4}$$

Proof. By definition:

$$\begin{aligned} h^*(p^*) &= \operatorname*{Sup}_{p \in Y} [\langle p^*, p \rangle - h(p)] \\ &= \operatorname*{Sup}_{p \in Y} [\langle p^*, p \rangle - \inf_{u \in V} \Phi(u, p)] \\ &= \operatorname*{Sup}_{p \in Y} \operatorname*{Sup}_{u \in V} [\langle p^*, p \rangle - \Phi(u, p)] \\ &= \Phi^*(0, p^*). \quad ■ \end{aligned}$$

It follows that

Lemma 2.3.

$$\sup \mathcal{P}^* = \sup_{p^* \in Y^*} [-h^*(p^*)] = h^{**}(0). \tag{2.5}$$

Inequality (1.6): $\sup \mathcal{P}^* \leqslant \inf \mathcal{P}$, is thus equivalent to the well-known inequality $h^{**}(0) \leqslant h(0)$.

Definition 2.1. Problem $\mathcal{P}$ is said to be *normal* if $h(0)$ is finite and h is l.s.c. at 0.

Proposition 2.1. *Under hypotheses* (2.1), *the three following conditions are equivalent to each other*

(i) $\mathcal{P}$ *is normal,*
(ii) $\mathcal{P}^*$ *is normal,*
(iii) $\inf \mathcal{P} = \sup \mathcal{P}^*$ *and this number is finite.*

Proof. We shall demonstrate the equivalence of (i) and (iii); the equivalence of (ii) and (iii) follows directly from the fact that $\mathcal{P}^{**} = \mathcal{P}$.

Assuming $\mathcal{P}$ to be normal and letting $\bar{h}$ be the l.s.c. regularization of h, we have

$$h^{**} \leqslant \bar{h} \leqslant h \tag{2.6}$$

(*cf.* Prop. I.3.3). By hypothesis $\bar{h}(0) = h(0) \in \mathbf{R}$. Since h is convex, $\bar{h}$ is l.s.c. and convex, assumes a finite value at 0 and hence cannot take the value $-\infty$ (Prop. I.2.4). It follows that $\bar{h} \in \Gamma_0(Y)$ (Prop. I.3.1) and thus $\bar{h}^{**} = \bar{h}$. The inequality (2.6) implies by duality that

$$h^* = h^{***} \geqslant \bar{h}^* \geqslant h^*$$

so that $h^* = \bar{h}^*$ and $h^{**} = \bar{h}^{**} = \bar{h}$; whence $\bar{h}(0) = h(0) = h^{**}(0)$ and the latter equality is none other than (iii) by virtue of Lemma 2.3.

Conversely (iii) implies that $h(0) = h^{**}(0) \in \mathbf{R}$ and with (2.6), $h(0) = \bar{h}(0) \in \mathbf{R}$ which means that the problem $\mathcal{P}$ is normal. ∎

Remark 2.3. The equivalence between (i) and (iii) is true under the single condition that h is convex, which holds if Φ is convex (*cf.* Remark 2.1).

Definition 2.2. Problem $\mathcal{P}$ is said to be *stable* if $h(0)$ is finite and h is subdifferentiable at 0.

Proposition 2.2. *The following two conditions are equivalent to each other*:

(i) $\mathcal{P}$ *is stable,*
(ii) $\mathcal{P}$ *is normal and* $\mathcal{P}^*$ *has at least one solution.*

Proof. If $\mathscr{P}$ is stable, $h(0)$ is finite and $\partial h(0)$ is non-empty; we know that in this case $h(0) = h^{**}(0)$ ($\in \mathbf{R}$) which implies that problem $\mathscr{P}$ is normal and in addition $\partial h^{**}(0) = \partial h(0) \neq \varnothing$ (*cf.* I(5.4)). Thus (i) implies (ii) by means of Lemma 2.4 below.

Conversely if $\mathscr{P}$ is normal, $h(0) = h^{**}(0) \in \mathbf{R}$ and if $\mathscr{P}^*$ has some solutions, $\partial h^{**}(0) = \partial h(0) \neq \varnothing$ (*cf.* I(5.4)) which implies that $\mathscr{P}$ is stable. ∎

Lemma 2.4. *The set of solutions of $\mathscr{P}^*$ is identical to $\partial h^{**}(0)$.*

Proof. If $p^* \in Y^*$ is a solution of $\mathscr{P}^*$, we have

$$- \Phi^*(0, p^*) \geqslant - \Phi^*(0, q^*), \qquad \forall q^* \in Y^*,$$

that is, from (2.4):

$$- h^*(p^*) \geqslant - h^*(q^*),$$
$$- h^*(p^*) = \sup_{q^* \in Y^*} [\langle 0, q^* \rangle - h^*(q^*)]$$
$$- h^*(p^*) = h^{**}(0)$$

and this is equivalent to

$$p^* \in \partial h^{**}(0). \quad ∎$$

When problem $\mathscr{P}$ is stable, Lemma 2.4 shows that the solutions of $\mathscr{P}^*$ are subgradients of h at 0, and therefore, in certain cases, the derivatives of inf $\mathscr{P}$ with respect to the given perturbations (*cf.* Chap. VII, Section 5).

Corollary 2.1. *Under hypothesis* (2.1) *the three following conditions are equivalent to each other*:

(i) $\mathscr{P}$ *and* $\mathscr{P}^*$ *are normal and have some solutions,*
(ii) $\mathscr{P}$ *and* $\mathscr{P}^*$ *are stable,*
(iii) $\mathscr{P}$ *is stable and has some solutions.*

Proof. The equivalence between (i) and (ii) follows from Proposition 2.2 and from the fact that $\mathscr{P}$ is the dual problem of $\mathscr{P}^*$. The equivalence between (i) and (iii) follows directly from Proposition 2.2. ∎

Proposition 2.3. *A stability criterion.*
Let us assume that Φ is convex, that inf $\mathscr{P}$ *is finite and that*

$$(2.7) \qquad \begin{cases} \text{There exists } u_0 \in V \text{ such that } p \mapsto \Phi(u_0, p) \text{ is finite} \\ \text{and continuous at } 0 (\in Y). \end{cases}$$

Then problem $\mathscr{P}$ is stable.

Proof. Because of Lemma 2.1 and Remark 2.1, h is convex and by hypothesis $h(0)$ is finite.

As the function $p \mapsto \Phi(u_0, p)$ is convex and continuous at $0(\in Y)$, there exists a neighbourhood $\mathscr{V}$ of 0 in Y, on which this function is bounded above:

$$\Phi(u_0, p) \leqslant M < +\infty, \qquad \forall p \in \mathscr{V}.$$

But

$$h(p) = \operatorname*{Inf}_{u \in V} \Phi(u, p) \leqslant \Phi(u_0, p) \leqslant M, \qquad \forall p \in \mathscr{V}$$

and Proposition I.2.5 thus implies that h is continuous at 0. Proposition I.5.2 then implies that h is subdifferentiable at 0.

Remark 2.4.

Hypothesis (2.7) which implies the existence of solutions for the dual problem $\mathscr{P}^*$ (for convex Φ and finite $\inf \mathscr{P}$), can be compared with the classical constraint qualification hypothesis in Operations Research (*cf.* M. Slater [1], H. P. Kunzi and W. Krelle [1], J. Cea [1]). Sometimes we will refer to it as the qualification hypothesis for problem $\mathscr{P}$.

Existence of solutions and extremal relations

Proposition 2.4. *If $\mathscr{P}$ and $\mathscr{P}^*$ possess solutions and if*

(2.8) $\inf \mathscr{P} = \sup \mathscr{P}^*$, *and this number is finite,*

all solutions $\bar{u}$ of $\mathscr{P}$ and all solutions $\bar{p}^*$ of $\mathscr{P}^*$ are linked by the extremality relation

$$\Phi(\bar{u}, 0) + \Phi^*(0, \bar{p}^*) = 0 \tag{2.9}$$

or

$$(0, \bar{p}^*) \in \partial\Phi(\bar{u}, 0). \tag{2.10}$$

Conversely if $\bar{u} \in V$ and $\bar{p}^* \in Y^*$ satisfy the extremality relation (2.9), *then $\bar{u}$ is a solution of $\mathscr{P}$, $\bar{p}^*$ is a solution of $\mathscr{P}^*$ and we have* (2.8).

Proof. We have in fact:

$$\inf \mathscr{P} = \Phi(\bar{u}, 0) = \sup \mathscr{P}^* = -\Phi^*(0, \bar{p}^*);$$

(2.9) can also be written

$$\Phi(\bar{u}, 0) + \Phi^*(0, \bar{p}^*) = \langle (\bar{u}, 0), (0, \bar{p}^*) \rangle,$$

which is identical to (2.10) from I(5.6).

If conversely $\bar{u}$ and $\bar{p}^*$ satisfy (2.9), it follows from (1.7) that

$$\Phi(\bar{u}, 0) = \operatorname*{Inf}_{u \in V} \Phi(u, 0)$$

$$-\Phi^*(0, \bar{p}^*) = \operatorname*{Sup}_{p^* \in Y^*} [-\Phi^*(0, p^*)]$$

which establishes the required properties.

Remark 2.5. By definition of a polar function

$$\Phi(u, 0) + \Phi^*(0, p^*) \geqslant \langle (u, 0), (0, p^*) \rangle = 0, \qquad \forall u \in V, \quad \forall p^* \in Y^*,$$

and for this reason (2.9) is called an extremality relation.

Remark 2.6. Relation (2.8) holds if problem $\mathscr{P}$ is normal or stable or if (2.7) is satified and inf $\mathscr{P}$ is finite (Φ being convex). Proposition 2.4 applies in particular to the situation of Corollary 2.1 ((i), (ii), or (iii) being satisfied). ■

The following proposition, by regrouping the previously proved results, provides an illustration.

Proposition 2.5. *We shall assume that V is a reflexive Banach space, that $\Phi \in \Gamma_0(V \times Y)$, that condition* (2.7) *is satisfied and that*:

$$\text{(2.11)} \qquad \lim \Phi(u, 0) = +\infty \quad \textit{if} \quad u \in V, \qquad \|u\| \to \infty.$$

Under these conditions, $\mathscr{P}$ and $\mathscr{P}^*$ each have at least one solution,

$$\text{(2.12)} \qquad \inf \mathscr{P} = \sup \mathscr{P}^*,$$

and the extremality relation (2.9) (2.10) *is satisfied.*

Proof. The function $F(u) = \Phi(u, 0)$ satisfies the hypotheses of Proposition II.1.2 and so we deduce the existence of a solution of $\mathscr{P}$.

From (2.7), problem $\mathscr{P}$ is stable (Prop. 2.3), which implies (2.12) and the existence of a solution for $\mathscr{P}^*$. Finally the extremality relations (2.9) (2.10) follow from the preceding and from Proposition 2.4 above.

3. LAGRANGIANS AND SADDLE POINTS

Definition 3.1. The function denoted by L such that

$$L : V \times Y^* \mapsto \bar{\mathbf{R}},$$

$$\text{(3.1)} \quad -L(u, p^*) = \operatorname*{Sup}_{p \in Y} [\langle p^*, p \rangle - \Phi(u, p)], \qquad \forall u \in V, \quad \forall p^* \in Y^*.$$

will be called the *Lagrangian* of problem $\mathscr{P}$ relative to the given perturbations.

We can write

$$- L(u, p^*) = \Phi_u^*(p^*), \tag{3.2}$$

where Φ_u denotes for fixed $u \in V$ the function $p \mapsto \Phi(u, p)$, and $\Phi_u^* \in \Gamma(Y^*)$ denotes the conjugate function of $\Phi_u \in \bar{\mathbf{R}}^Y$.

The properties of L are given by

Lemma 3.1. *For all u of V, L_u:*

$$L_u : p^* \mapsto L(u, p^*), \tag{3.3}$$

is a concave u.s.c. function of Y^* into $\bar{\mathbf{R}}$.

If Φ is convex, then for all $p^* \in Y^*$, the function L_{p^*}:

$$L_{p^*} : u \mapsto L(u, p^*) \tag{3.4}$$

is convex from V into $\bar{\mathbf{R}}$.

Proof. We have $-L_u = \Phi_u^* \in \Gamma(Y^*)$ and so L_u is concave and u.s.c.[1]
Let us now demonstrate the convexity of L_{p^*}; we have

$$L(u, p^*) = \operatorname*{Inf}_{p \in Y} [\Phi(u, p) - \langle p^*, p \rangle]. \tag{3.5}$$

Let $u, v \in V$ and $\lambda \in]0, 1[$; the inequality

$$L(\lambda u + (1 - \lambda)v, p^*) \leqslant \lambda L(u, p^*) + (1 - \lambda)L(v, p^*), \tag{3.6}$$

is obvious if $L(u, p^*)$ or $L(v, p^*) = +\infty$. Hence let us assume that $L(u, p^*) < +\infty$ and that $L(v, p^*) < +\infty$ and let $a > L(u, p^*)$ and $b > L(v, p^*)$ be fixed. Due to (3.5) there exists $p \in Y$, $q \in Y$, such that

$$L(u, p^*) \leqslant \Phi(u, p) - \langle p^*, p \rangle \leqslant a, \tag{3.7}$$

$$L(v, p^*) \leqslant \Phi(v, q) - \langle p^*, q \rangle \leqslant b; \tag{3.8}$$

but:

$$\begin{aligned} L(\lambda u + (1 - \lambda)v, p^*) &\leqslant \Phi(\lambda u + (1 - \lambda)v, \lambda p + (1 - \lambda)q) \\ &\qquad\qquad - \langle p^*, \lambda p + (1 - \lambda)q \rangle \\ &\leqslant \text{(by the convexity of } \Phi) \\ &\leqslant \lambda[\Phi(u, p) - \langle p^*, p \rangle] \\ &\qquad\qquad + (1 - \lambda)[\Phi(v, q) - \langle p^*, q \rangle] \\ &\leqslant \text{(by (3.7) and (3.8))} \\ &\leqslant \lambda a + (1 - \lambda)b. \end{aligned}$$

(1) A function F is concave if $-F$ is convex.

Letting a decrease towards $L(u, p^*)$ and b decrease towards $L(v, p^*)$, we obtain (3.6).

Remark 3.1. We cannot assert that L_{p^*} is l.s.c. even if we assume that $\Phi \in \Gamma_0(V \times Y)$. ■

It will be useful to express problems $\mathscr{P}$ and $\mathscr{P}^*$ in terms of the function L. Without assuming anything about Φ, we have

$$\begin{aligned}\Phi^*(u^*, p^*) &= \operatorname*{Sup}_{\substack{u \in V\\ p \in Y}} [\langle u^*, u \rangle + \langle p^*, p \rangle - \Phi(u, p)]\\ &= \operatorname*{Sup}_{u \in V} \{ \langle u^*, u \rangle + \operatorname*{Sup}_{p \in Y} [\langle p^*, p \rangle - \Phi(u, p)] \}\\ &= \operatorname*{Sup}_{u \in V} [\langle u^*, u \rangle - L(u, p^*)].\end{aligned}$$

Whence:

$$-\Phi^*(0, p^*) = \operatorname*{Inf}_{u \in V} L(u, p^*) \tag{3.9}$$

and problem $\mathscr{P}^*$,

$$\operatorname*{Sup}_{p^* \in Y^*} [-\Phi^*(0, p^*)]$$

can be written as:

(3.10) $\quad(\mathscr{P}^*)\qquad \operatorname*{Sup}_{p^* \in Y}\operatorname*{Inf}_{u \in V} L(u, p^*) \qquad$ (*for an arbitrary function* Φ).

Similarly, if we assume that $\Phi \in \Gamma_0(V \times Y)$ then $\forall u \in V$, the function $\Phi_u : p \to \Phi(u, p)$, belongs to $\Gamma(Y)$ and thus $\Phi_u^{**} = \Phi_u$, whence:

$$\begin{aligned}\Phi(u, p) &= \Phi_u^{**}(p)\\ &= \operatorname*{Sup}_{p^* \in Y^*} [\langle p, p^* \rangle - \Phi_u^*(p^*)]\\ &= \operatorname*{Sup}_{p^* \in Y^*} [\langle p, p^* \rangle + L(u, p^*)].\end{aligned}$$

Hence:

$$\Phi(u, 0) = \operatorname*{Sup}_{p^* \in Y^*} L(u, p^*) \tag{3.11}$$

and problem $\mathscr{P}$ can be written as

(3.12) $\quad(\mathscr{P})\qquad \operatorname*{Inf}_{u \in V}\operatorname*{Sup}_{p^* \in Y^*} L(u, p^*) \quad (\Phi \in \Gamma_0(V \times Y))$ [1].

[1] It is sufficient for this that $\Phi_u \in \Gamma(Y)$, $\forall u \in V$.

Remark 3.2. By introducing the Lagrangian L, the problems $\mathscr{P}$ and $\mathscr{P}^*$ are shown to be related to the min-max problems which arise in games theory (*cf.* Chap. VI). Note that (1.6) is exactly the inequality Sup Inf $L \leqslant$ Inf Sup L, well known in games theory. The point of view adopted here and that of games theory lead in parallel fashion to similar results; therefore we will just give a few remarks and comments related to this second approach. See also Chapter VI.

Definition 3.2. $(\bar{u},\bar{p}^*) \in V \times Y^*$ is called a *saddle point of* L if

$$L(\bar{u}, p^*) \leqslant L(\bar{u}, \bar{p}^*) \leqslant L(u, \bar{p}^*), \qquad \forall u \in V, \quad \forall p^* \in Y^*. \tag{3.13}$$

Proposition 3.1. *Under the hypothesis that* $\Phi \in \Gamma_0(V \times Y)$, *the following two conditions are equivalent to each other:*

(i) $(\bar{u}, \bar{p}^*)$ *is a saddle point of* L,
(ii) $\bar{u}$ *is a solution of* $\mathscr{P}$, $\bar{p}^*$ *is a solution of* $\mathscr{P}^*$, *and* $\inf \mathscr{P} = \sup \mathscr{P}^*$.

Proof. From (3.12) and (3.9)

$$L(\bar{u}, \bar{p}^*) = \operatorname*{Inf}_{u \in V} L(u, \bar{p}^*) = -\Phi^*(0, \bar{p}^*),$$

and from (3.12) and (3.11)

$$L(\bar{u}, \bar{p}^*) = \operatorname*{Sup}_{p^* \in Y^*} L(\bar{u}, p^*) = \Phi(\bar{u}, 0).$$

Thus:

$$\Phi(\bar{u}, 0) + \Phi^*(0, \bar{p}^*) = 0,$$

and due to Proposition 2.4, $\bar{u}$ is a solution of $\mathscr{P}$, $\bar{p}^*$ is a solution of $\mathscr{P}^*$ and $\inf \mathscr{P} = \sup \mathscr{P}^*$.

Let us now show that (ii) implies (i); using (3.9) and (3.11) we obtain

$$-\Phi^*(0, \bar{p}^*) = \operatorname*{Inf}_{u \in V} L(u, \bar{p}^*) \leqslant L(\bar{u}, \bar{p}^*)$$

$$\Phi(\bar{u}, 0) = \operatorname*{Sup}_{p^* \in Y^*} L(\bar{u}, p^*) \geqslant L(\bar{u}, \bar{p}^*),$$

and as $\Phi(\bar{u}, 0) = -\Phi^*(0, \bar{p}^*)$, this implies

$$\operatorname*{Sup}_{p^* \in Y^*} L(\bar{u}, p^*) = L(\bar{u}, \bar{p}^*) = \operatorname*{Inf}_{u \in V} L(u, p^*),$$

which is equivalent to (3.13). ∎

Proposition 3.2. *Let us assume that* $\Phi \in \Gamma_0(V \times Y)$ *and that problem* $\mathscr{P}$ *is stable.*

Then $\bar{u} \in V$ is a solution of $\mathscr{P}$ if and only if there exists $\bar{p}^* \in \bar{Y}^*$ such that $(\bar{u}, \bar{p}^*)$ is a saddle point of L.

Proof. Clearly if $\bar{u} \in V$ and if such a $\bar{p}^*$ exists, Proposition 3.1 implies that $\bar{u}$ is a solution of $\mathscr{P}$ (and $\bar{p}^*$ a solution of $\mathscr{P}^*$).

Now if $\bar{u}$ is a solution of $\mathscr{P}$, since the problem is stable, problem $\mathscr{P}^*$ has at least one solution $\bar{p}^*$ and $\inf \mathscr{P} = \sup \mathscr{P}^*$; Proposition 3.1 then implies that $(\bar{u}, \bar{p}^*)$ is a saddle point of L. ∎

Remark 3.3. We defer until Chapter VI a more precise comparison between the optimization problems and the saddle point ones. ∎

4. IMPORTANT SPECIAL CASES (I)

Orientation

In this section and the next one, we shall apply the preceding results to two important particular cases:

a general frame work, based on Fenchel [1]–[2] and Rockafellar [5]–[7]–[8] (*cf.* also Temam [1]) and especially adapted to certain problems in the calculus of variations;

duality according to Arrow–Hurwicz [1] which permits us to recover the duality results in convex programming (Kuhn–Tucker theorem [1]).

In this section we shall consider the "calculus of variations" setting. ∎

Given the paired spaces V and V^*, Y and Y^*, we shall assume the existence of a continuous linear operator Λ of V into Y, $\Lambda \in \mathscr{L}(V, Y)$, with transpose $\Lambda^* \in \mathscr{L}(Y^*, V^*)$. We shall also assume that the function F to be minimized can be written as:

$$F(u) = J(u, \Lambda u), \tag{4.1}$$

where J is a function of $V \times Y$ into $\bar{\mathbf{R}}$. Problem $\mathscr{P}$ takes the form

$$\inf_{u \in V} J(u, \Lambda u). \tag{4.2}$$

In this case the function Φ will be

$$\Phi(u, p) = J(u, \Lambda u - p). \tag{4.3}$$

It is easy to determine the dual problem; if $J^* \in \Gamma(V^* \times Y^*)$ denotes the conjugate function of J, we have

$$\Phi^*(0, p^*) = \operatorname*{Sup}_{\substack{u \in V \\ p \in Y}} [\langle p^*, p \rangle - J(u, \Lambda u - p)]$$

$$= \operatorname*{Sup}_{u \in V} \operatorname*{Sup}_{p \in Y} [\langle p^*, p \rangle - J(u, \Lambda u - p)];$$

setting, for fixed u, $q = \Lambda u - p$, we find that

$$\Phi^*(0, p^*) = \operatorname*{Sup}_{u \in V} \operatorname*{Sup}_{q \in Y} [\langle p^*, \Lambda u \rangle - \langle p^*, q \rangle - J(u, q)],$$

$$\Phi^*(0, p^*) = \operatorname*{Sup}_{\substack{u \in V \\ q \in Y}} [\langle \Lambda^* p^*, u \rangle - \langle p^*, q \rangle - J(u, q)]$$

whence

$$\Phi^*(0, p^*) = J^*(\Lambda^* p^*, - p^*). \tag{4.4}$$

Thus problem $\mathscr{P}^*$ can be written as

$$\operatorname*{Sup}_{p^* \in Y^*} [- J^*(\Lambda^* p^*, - p^*)]. \tag{4.5}$$

The relationship between the properties of Φ and those of J is rendered more precise by the following two obvious remarks:

(4.6) If J is convex, Φ is convex.

(4.7) If $J \in \Gamma_0(V \times Y)$, $\Phi \in \Gamma_0(V \times Y)$.

Theorem 4.1. *Let us assume that J is convex, that* inf $\mathscr{P}$ *is finite and that*

(4.8) *There exists $u_0 \in V$ such that $J(u_0, \Lambda u_0) < +\infty$, the function $p \mapsto J(u_0, p)$ being continuous at Λu_0.*

Then problem (4.2) *is stable*:

$$\inf \mathscr{P} = \sup \mathscr{P}^*, \tag{4.9}$$

and $\mathscr{P}^*$ has at least one solution $\bar{p}^*$.

Proof. We apply Proposition 2.3 having noted that (4.8) implies (2.7).

Proposition 4.1. *The following two conditions are equivalent to each other:*

(i) *$\bar{u}$ is a solution of* (4.2), *$\bar{p}^*$ is a solution of* (4.5) *and*

$$\inf \mathscr{P} = \sup \mathscr{P}^*.$$

(ii) *$\bar{u} \in V$ and $\bar{p}^* \in Y^*$ satisfy the extremality relations*

$$J(\bar{u}, \Lambda\bar{u}) + J^*(\Lambda^*\bar{p}^*, -\bar{p}^*) = 0 \tag{4.10}$$

which is equivalent to

$$(\Lambda^*\bar{p}^*, -\bar{p}^*) \in \partial J(\bar{u}, \Lambda\bar{u}). \tag{4.11}$$

We simply apply Proposition 2.4.

Theorem 4.2. *Let us assume that V is a reflexive Banach space, $J \in \Gamma_0(V \times Y)$, that condition* (4.8) *is satisfied and that*

$$\lim J(u, \Lambda u) = +\infty, \quad \text{if } u \in V, \qquad \|u\| \to \infty. \tag{4.12}$$

Under these conditions, (4.2) *and* (4.5) *each have (at least) one solution,*

$$\inf \mathscr{P} = \sup \mathscr{P}^*, \tag{4.13}$$

and the extremality relation (4.10)–(4.11) *is satisfied.*

Apply Proposition 2.5.

Remark 4.1. The Lagrangian associated with this problem can be written as:

$$-L(u, p^*) = \operatorname*{Sup}_{p \in Y} [\langle p^*, p \rangle - J(u, \Lambda u - p)]$$

or putting $q = \Lambda u - p$

$$L(u, p^*) = -\langle p^*, \Lambda u \rangle + J_u^*(-p^*), \tag{4.14}$$

where J_u is the function $p \mapsto J(u, p)$ and $J_u^* \in \Gamma(Y^*)$ is its conjugate function.

Remark 4.2. The above can also be usefully particularized in the case where, with the same spaces and the same operator Λ, the function J can be *decomposed* into the form

$$J(u, \Lambda u) = F(u) + G(\Lambda u), \tag{4.15}$$

with:

$$F \in \mathbf{R}^V, \qquad G \in \mathbf{R}^Y.\text{[1]}$$

Problem $\mathscr{P}$ can be written as:

$$\operatorname*{Inf}_{u \in V} [F(u) + G(\Lambda u)]; \tag{4.16}$$

[1] This function F must not be confused with the "initial" function F introduced in Section 1.

it is easy to verify that:

$$J^*(u^*, p^*) = F^*(u^*) + G^*(p^*), \tag{4.17}$$

where $F^* \in \Gamma(V^*)$ and $G^* \in \Gamma^*(Y^*)$ are the conjugate functions of F and G respectively.

Problem $\mathscr{P}^*$ can be written as:

$$\operatorname*{Sup}_{p^* \in Y^*} [- F^*(\Lambda^* p^*) - G^*(- p^*)]. \tag{4.18}$$

Let us note here that

(4.19) If F and G are convex, J (and hence Φ) is convex,

(4.20) If $F \in \Gamma_0(V)$ and $G \in \Gamma_0(Y)$ then

$$J \in \Gamma_0(V \times Y), \qquad \Phi \in \Gamma_0(V \times Y).$$

Proposition 4.1 and Theorems 4.1 and 4.2 apply; condition (4.8) can be written as:

(4.21) *There exists $u_0 \in V$ such that $F(u_0) < +\infty$, $G(\Lambda u_0) < +\infty$, G being continuous at Λu_0.*

The extremality relation (4.10) can be decoupled into:

$$\begin{aligned} 0 &= J(\bar{u}, \Lambda\bar{u}) + J^*(\Lambda^*\bar{p}^*, - \bar{p}^*) \\ &= F(\bar{u}) + F^*(\Lambda^*\bar{p}^*) + G(\Lambda\bar{u}) + G^*(- \bar{p}^*) \\ &= [F(\bar{u}) + F^*(\Lambda^*\bar{p}^*) - \langle \Lambda^*\bar{p}^*, \bar{u} \rangle] \\ &\quad + [G(\Lambda\bar{u}) + G^*(- \bar{p}^*) - \langle - \bar{p}^*, \Lambda\bar{u} \rangle]. \end{aligned}$$

As each expression in square brackets is positive or zero, this implies

$$F(\bar{u}) + F^*(\Lambda^*\bar{p}^*) - \langle \Lambda^*\bar{p}^*, \bar{u} \rangle = 0, \tag{4.22}$$

$$G(\Lambda\bar{u}) + G^*(- \bar{p}^*) + \langle \bar{p}^*, \Lambda\bar{u} \rangle = 0, \tag{4.23}$$

and these conditions amount to saying that

$$\Lambda^*\bar{p}^* \in \partial F(\bar{u}), \tag{4.24}$$

$$- \bar{p}^* \in \partial G(\Lambda\bar{u}). \quad \blacksquare \tag{4.25}$$

Remark 4.3. The above can be further particularized to the case where Y is a product

$$Y = \prod_{i=1}^{m} Y_i, \qquad Y^* = \prod_{i=1}^{m} Y_i^*, \tag{4.26}$$

the spaces Y_i and Y_i^* being mutually dual and the function G being decomposed into

$$G(p) = \sum_{i=1}^{m} G_i(p_i),$$
$$p = (p_1, ..., p_m) \in Y, \; G_i \in \bar{\mathbf{R}}^{Y_i}.$$

We have:

$$G^*(p^*) = \sum_{i=1}^{m} G_i^*(p_i^*), \qquad \forall p^* = (p_1^*, ..., p_m^*) \in Y^*,$$

where $G_i^* \in \Gamma(Y_i^*)$ is the conjugate function of G_i.

In this case the extremality relation (4.23) can be decomposed into the m relations ($1 \leqslant i \leqslant m$):

$$G_i(\Lambda_i u) + G_i^*(-p_i^*) + \langle p_i^*, \Lambda_i u \rangle = 0, \tag{4.27}$$

(where $\Lambda_i \bar{u}$ is the ith component of Λu), which is equivalent to

$$-\bar{p}_i^* \in \partial G_i(\Lambda_i \bar{u}). \quad \blacksquare \tag{4.28}$$

5. IMPORTANT SPECIAL CASES (II)

Ordering relations associated with cones

We recall[1] that in a vector space Y, a subset $\mathscr{C}$ of Y such that $\lambda\mathscr{C} \subset \mathscr{C}$, $\forall \lambda > 0$ is called a cone with vertex O. The cone is *pointed* or *unpointed* according to whether $O \in \mathscr{C}$ or $O \notin \mathscr{C}$; a pointed cone with vertex O is *salient* if $\mathscr{C} \cap \{-\mathscr{C}\} = \{0\}$.

We can associate a partial ordering relation denoted by $\leqslant$ or $\geqslant$ with a pointed cone $\mathscr{C}$ by setting

$$p \leqslant q \Leftrightarrow q - p \in \mathscr{C}. \tag{5.1}$$

Obviously $p \leqslant p$, $\forall p \in Y$; if $p \leqslant q$, and $q \leqslant r$, then $p \leqslant r$; the partial ordering relation is compatible with the structure of a vector space in the sense that

$$p \geqslant 0 \Rightarrow \lambda p \geqslant 0, \qquad \forall \lambda > 0, \tag{5.2}$$

$$p \geqslant q \Rightarrow p + r \geqslant q + r, \qquad \forall r. \tag{5.3}$$

[1] For these questions *cf.* Bourbaki [2].

The cone $\mathscr{C}$ is the set of *positive elements* for this ordering relation

$$\mathscr{C} = \{ p \in Y \mid p \geqslant 0 \}. \tag{5.4}$$

The set $\{-\mathscr{C}\}$ is the set of negative elements

$$\{ -\mathscr{C} \} = \{ p \in Y \mid p \leqslant 0 \}. \tag{5.5}$$

If, moreover, the cone $\mathscr{C}$ is salient, the relation $\leqslant$ is an ordering relation:

$$p \leqslant q, q \leqslant p \Rightarrow p = q. \tag{5.6}$$

Conversely, if we are given a partial ordering relation over Y, $\leqslant$ (or an ordering relation), compatible with the structure of the vector space Y (*cf.* (5.2) (5.3)), then the set of positive elements ($\geqslant 0$) is a pointed (or pointed salient cone) with vertex O.

If, now, Y and Y^* are two vector spaces in duality, we can associate with a cone $\mathscr{C}$ of Y its polar cone $\mathscr{C}^*$

$$\mathscr{C}^* = \{ p^* \in Y^* \mid \langle p^*, p \rangle \geqslant 0, \ \forall p \in \mathscr{C} \}. \tag{5.7}$$

Since $\mathscr{C}^*$ is a pointed cone with vertex O in Y^*, it defines a partial ordering relation denoted by $\leqslant$ or $\geqslant$:

$$p^* \leqslant q^* \Leftrightarrow q^* - p^* \in \mathscr{C}^*; \tag{5.8}$$

hence $\mathscr{C}^*$ is evidently the *cone of positive elements* in Y^*.

If, finally, Y and Y^* are two dual topological vector spaces and if $\mathscr{C}$ is a *pointed closed convex cone with vertex* O, then from the above we have the following property:

$$p \in \mathscr{C} \Leftrightarrow p \geqslant 0 \Leftrightarrow \langle p^*, p \rangle \geqslant 0, \qquad \forall p^* \in \mathscr{C}^*. \tag{5.9}$$

Indeed $\mathscr{C}^{**} = \mathscr{C}$, $\mathscr{C}^{**}$ being the polar cone of $\mathscr{C}^*$ and so $p \in \mathscr{C} \Leftrightarrow p \in \mathscr{C}^{**} \Leftrightarrow$ (by definition) $\Leftrightarrow \langle p, p^* \rangle \geqslant 0$, $\forall p^* \in \mathscr{C}^*$.

Primal problem and dual problem

As in the preceding sections, we take the two pairs of locally convex topological vector spaces in duality V and V^*, Y and Y^*.

Let $\mathscr{A}$ be a non-empty closed convex subset of V, and take J such that

$$J \text{ is a convex l.s.c. function of } \mathscr{A} \text{ into } \mathbf{R}. \tag{5.10}$$

Let $\mathscr{C}$ be a closed convex cone of Y defining a partial ordering relation $\leqslant$

and let $\mathscr{C}^*$ be its polar cone. Finally we consider a mapping B (possibly non-linear) of $\mathscr{A}$ into Y such that:

(5.11) B is convex with respect to the relation $\leqslant$:

$$B(\lambda u + (1 - \lambda)v) \leqslant \lambda B(u) + (1 - \lambda)B(v) \quad \forall u, v \in \mathscr{A}, \forall \lambda \in]0, 1[.$$

(5.12) For each $p^* \in Y^*$, $p^* \geqslant 0$, the mapping $u \mapsto \langle p^*, Bu \rangle$ of $\mathscr{A}$ into $\mathbf{R}$ is l.s.c.

(5.13) $$\{ u \in \mathscr{A} \mid Bu \leqslant 0 \} \neq \varnothing.$$

The primal problem $\mathscr{P}$ which now concerns us is the following:

(5.14) $$\operatorname*{Inf}_{\substack{u \in \mathscr{A} \\ Bu \leqslant 0}} J(u).$$

This can be written as

(5.15) $$\operatorname*{Inf}_{u \in V} F(u)$$

provided we set

(5.16) $$F(u) = \begin{vmatrix} Ju & \text{if } u \in \mathscr{A} \text{ and } Bu \leqslant 0, \\ +\infty & \text{otherwise.} \end{vmatrix}$$

The perturbation function Φ is thus chosen to be

(5.17) $$\Phi(u, p) = \begin{vmatrix} Ju & \text{if } u \in \mathscr{A} \text{ and } Bu \leqslant p, \\ +\infty & \text{otherwise.} \end{vmatrix}$$

This function Φ cannot take the value $-\infty$ and from (5.12) is not identically equal to $+\infty$: thus it is *proper*.

We note that Φ can be written as

(5.18) $$\Phi(u, p) = \hat{J}(u) + \chi_{\mathscr{E}_p}(u),$$

where

(5.19) $$\hat{J}(u) = \begin{vmatrix} J(u) & \text{if } u \in \mathscr{A}, \\ +\infty & \text{otherwise,} \end{vmatrix}$$

and $\chi_{\mathscr{E}_p}$ is the indicator function of the set

(5.20) $$\mathscr{E}_p = \{ u \in V \mid u \in \mathscr{A} \text{ and } Bu \leqslant p \}.$$

The following lemmas will specify more exactly the properties of Φ.

Lemma 5.1. (i) *The set $\mathscr{E}_p$ is closed and convex in V, $\forall p \in Y$.*
(ii) *The set $\mathscr{E} = \{(u, p) \in V \times Y \mid u \in \mathscr{A},\ Bu \leqslant p\}$ is closed and convex in $V \times Y$.*

Proof. (i) From (5.9) and (5.20), for $u \in \mathscr{A}$,

$$u \in \mathscr{E}_p \Leftrightarrow Bu - p \leqslant 0 \Leftrightarrow \langle Bu - p, p^* \rangle \geqslant 0, \quad \forall p^* \geqslant 0.$$

Due to (5.12) and (5.13), the function

$$u \mapsto \langle Bu - p, p^* \rangle,$$

of $\mathscr{A}$ into $\mathbf{R}$ is l.s.c. and convex for fixed $p^* \geqslant 0$, and hence the set

$$\{ u \in \mathscr{A} | \langle Bu - p, p^* \rangle \leqslant 0 \}$$

is closed and convex; the same holds good *a fortiori* for the intersection of all such sets which correspond to all the $p^* \geqslant 0$.

(ii) The same reasoning applies to $\mathscr{E}$, if we observe simply that the function

$$\{ u, p \} \mapsto \langle Bu - p, p^* \rangle$$

is l.s.c. and convex of $\mathscr{A} \times Y$ into $\mathbf{R}$, $\forall p^* \geqslant 0$ fixed. ■

Lemma 5.2.

(5.21) $$\Phi \in \Gamma_0(V \times Y).$$

Proof. Let us write Φ in the form (5.18) or, more accurately, in the form

(5.22) $$\Phi(u, p) = \hat{J}(u) + \chi_{\mathscr{E}}(\{ u, p \}), \quad \chi_{\mathscr{E}} \text{ is the indicator function of } \mathscr{E}.$$

Clearly $\hat{J}$ is l.s.c. and convex and so is $\chi_{\mathscr{E}}$ ($\mathscr{E}$ is a closed convex set); thus the same is true for Φ and since Φ is proper, Proposition I.3.1 implies (5.21).

The dual problem

By definition, for $p^* \in Y^*$:

$$\Phi^*(0, p^*) = \operatorname*{Sup}_{\substack{u \in V \\ p \in Y}} \{ \langle p^*, p \rangle - \Phi(u, p) \} = \operatorname*{Sup}_{\substack{u \in \mathscr{A} \\ p \in Y \\ Bu \leqslant p}} \{ \langle p^*, p \rangle - J(u) \};$$

putting $p = Bu + q$, we have

$$\Phi^*(0, p^*) = \operatorname*{Sup}_{u \in \mathscr{A}} \operatorname*{Sup}_{\substack{q \in Y \\ q \geqslant 0}} \{ \langle p^*, Bu \rangle + \langle p^*, q \rangle - J(u) \}$$

$$\Phi^*(0, p^*) = \chi_{\mathscr{C}^*}(-p^*) + \operatorname*{Sup}_{u \in \mathscr{A}} \{ \langle p^*, Bu \rangle - J(u) \}$$

$$-\Phi^*(0, p^*) = -\chi_{\mathscr{C}^*}(-p^*) + \operatorname*{Inf}_{u \in \mathscr{A}} \{ -\langle p^*, Bu \rangle + J(u) \}.$$

Thus the dual problem is

(5.23) $$\operatorname*{Sup}_{p^* \leqslant 0} \operatorname*{Inf}_{u \in \mathscr{A}} \{ - \langle p^*, Bu \rangle + J(u) \}.$$

As an application of Propositions 2.3 and 2.5 we have:

Proposition 5.1. *In addition to hypotheses* (5.10) *to* (5.13), *we assume that the infimum in* (5.14) *is finite and we make the following hypothesis:*(1)

(5.24) *There exists $u_0 \in \mathscr{A}$ such that $-Bu_0 < 0$, that is $-Bu_0 \in \mathring{\mathscr{C}}$ = the interior of $\mathscr{C}$.*

Then problem (5.14) *is stable.*

Proof. From (5.24) there exists a neighbourhood $\mathscr{V}$ of 0 in Y such that $-Bu_0 + p \in \mathscr{C}$, $\forall p \in \mathscr{V}$. Then $\Phi(u_0, p) = J(u_0)$, $\forall p \in \mathscr{V}$, and $p \mapsto \Phi(u_0, p)$ is finite and continuous at $0 (\in Y)$; condition (2.7) is thus satisfied.

Proposition 5.2. *Let us assume that V is a reflexive Banach space, that conditions* (5.10), (5.11), (5.13) *and* (5.24) *are satisfied and additionally that*

(5.25) $$\lim J(u) = +\infty, \quad \text{if} \quad u \in \mathscr{A}, \qquad \|u\|_V \to \infty.$$

Under these conditions $\mathscr{P}$ and $\mathscr{P}^*$ each possess at least one solution,

(5.26) $$\inf \mathscr{P} = \sup \mathscr{P}^*,$$

and we have the extremality relation

(5.27) $$\langle \bar{p}^*, B\bar{u} \rangle = 0$$ (2).

Proof. Lemma 5.2 and the proof of Proposition 5.1 demonstrate that the hypothesis of Proposition 2.5 are satisfied. The extremality relation (2.9) gives, taking (5.17) and (5.23) into account,

$$J(\bar{u}) = \operatorname*{Inf}_{u \in \mathscr{A}} \{ - \langle \bar{p}^*, Bu \rangle + J(u) \} \leqslant - \langle \bar{p}^*, B\bar{u} \rangle + J(\bar{u});$$

but $\bar{p}^* \leqslant 0$ and $B\bar{u} \leqslant 0$ since $\bar{p}^*$ is a solution of $\mathscr{P}^*$ and u a solution of $\mathscr{P}$; hence $\langle \bar{p}^*, B\bar{u} \rangle \geqslant 0$ and the above inequality implies that $\langle \bar{p}^*, B\bar{u} \rangle \leqslant 0$ and thus we have (5.27).

(1) Called the constraint qualification hypothesis (*cf.* Arrow and Hurwicz [1], Slater [1]).
(2) Note that $\bar{p}^* \leqslant 0$, $B\bar{u} \leqslant 0$ and hence, *a priori*, $\langle \bar{p}^*, B\bar{u} \rangle \geqslant 0$.

Computation of the Lagrangian

In the present case, the results take on a more interesting form using the Lagrangian which we shall compute:

$$
\begin{aligned}
-L(u,p^*) &= \operatorname*{Sup}_{p\in Y}\{\langle p^*,p\rangle - \Phi(u,p)\}\\
&= -\hat{J}(u) + \operatorname*{Sup}_{\substack{p\in Y\\ p\geqslant Bu}}\langle p^*,p\rangle\\
&= -\hat{J}(u) + \langle p^*, Bu\rangle + \operatorname*{Sup}_{\substack{q\in Y\\ q\geqslant 0}}\langle p^*,q\rangle\\
&= -\hat{J}(u) + \langle p^*, Bu\rangle + \chi_{\mathscr{C}^*}(-p^*)
\end{aligned}
$$

(5.28) $$L(u,p^*) = \hat{J}(u) - \langle p^*, Bu\rangle \dot{-} \chi_{\mathscr{C}^*}(-p^*).^{(1)}$$

From Definition 3.2 and (5.28) a pair $\{\bar{u},\bar{p}^*\}\in V\times V^*$ is a saddle point of L if and only if

(5.29) $$\bar{u}\in\mathscr{A},\qquad \bar{p}^*\leqslant 0,$$

and

(5.30) $$J(\bar{u}) - \langle p^*, B\bar{u}\rangle \leqslant J(\bar{u}) - \langle\bar{p}^*, B\bar{u}\rangle \leqslant J(u) - \langle \bar{p}^*, Bu\rangle,\quad \forall u\in\mathscr{A},\quad \forall p^*\leqslant 0.$$

Indeed if $\{\bar{u},\bar{p}^*\}$ is a saddle point of L, then by writing (3.13) with $u\in\mathscr{A}$ and $p^*\leqslant 0$, we find that

(5.31) $$\begin{aligned}\hat{J}(\bar{u}) - \langle p^*, B\bar{u}\rangle - \chi_{\mathscr{C}^*}(-p^*) &\leqslant \hat{J}(\bar{u}) - \langle\bar{p}^*, B\bar{u}\rangle \dot{-} \chi_{\mathscr{C}^*}(-\bar{p}^*)\\ &\leqslant \hat{J}(u) - \langle\bar{p}^*, Bu\rangle - \chi_{\mathscr{C}^*}(-\bar{p}^*).\end{aligned}$$

We successively deduce from these inequalities that we cannot have $\hat{J}(u) = +\infty$, nor $\chi_{\mathscr{C}^*}(-\bar{p}^*) = +\infty$, and (5.29) follows; (5.30) is then an obvious consequence of (5.31). Conversely if (5.29) and (5.30) are true, (5.31) is true for $u\in\mathscr{A}$ and $p^*\leqslant 0$; but if $u\notin\mathscr{A}$, $L(u,\bar{p}^*) = +\infty$, and if $p^*\notin\mathscr{C}^*$, $L(\bar{u},p^*) = -\infty$ and (3.13) is clearly true.

(1) The addition and the subtraction operations in $\bar{\mathbf{R}}$ are specified as follows:

$$(+\infty) \dot{+} (-\infty) = (+\infty) \dot{-} (+\infty) = +\infty$$
$$(+\infty) + (-\infty) = (+\infty) \underset{\cdot}{-} (+\infty) = -\infty$$

With this remark, Proposition 3.2 yields (1)

Theorem 5.1. *Let us assume that conditions* (5.10)–(5.11)–(5.13) *and* (5.24) *are satisfied and that the infimum in* (5.14) *is finite.*(2)

Then $\bar{u} \in \mathscr{A}$ is a solution of (5.14) *if and only if there exists $\bar{p}^* \in Y$, $\bar{p}^* \leqslant 0$ such that $\{\bar{u}, \bar{p}^*\}$ is a saddle point of L:*

$$J(\bar{u}) - \langle p^*, B\bar{u}\rangle \leqslant J(\bar{u}) - \langle \bar{p}^*, B\bar{u}\rangle \leqslant J(u) - \langle \bar{p}^*, Bu\rangle, \quad \forall \bar{u} \in \mathscr{A}, \quad \forall p^* \leqslant 0. \tag{5.30}$$

In this case

$$\langle \bar{p}^*, B\bar{u} \rangle = 0. \tag{5.32}$$

Proof. From Proposition 5.1, problem (5.14) is stable and everything is thus a result of Proposition 3.2. Setting $p^* = 0$ in (5.31), we find that $\langle \bar{p}^*, B\bar{u}\rangle \geqslant 0$, whence (5.32), the inequality $\langle \bar{p}^*, B\bar{u}\rangle \leqslant 0$ following from the properties of $\bar{p}^*$ and $\bar{u}$. Furthermore, relation (5.32) is the extremality relation of the problem. ∎

Remark 5.1. The above results, particularly in the shape of Theorem 5.1, are given in Arrow and Hurwicz [1].

Orientation

In the rest of this section we shall apply Proposition 5.3 to the specific case of a classical problem, that of finite dimensional convex programming. ∎

Finite dimensional convex programming: the Kuhn–Tucker theorem

Here $V = V^* = \mathbf{R}^n$, $Y = Y^* = \mathbf{R}^m$; $u \in V$ will be written as $u = (u_1, \ldots, u_n)$, and $p \in Y$ will be written as $p = (p_1, \ldots, p_m)$. We shall take a closed convex set $\mathscr{A}$ of $\mathbf{R}^n$, and an l.s.c. convex function J of $\mathscr{A}$ into $\mathbf{R}$. The cone $\mathscr{C}$ of $Y = \mathbf{R}^m$ is as follows

$$\{\, p \mid p_i \geqslant 0, \ 1 \leqslant i \leqslant m \,\},$$

and the function B of $\mathscr{A}$ into Y is defined by its components $B_1, \ldots, B_m$, which are functions of $\mathscr{A}$ into $\mathbf{R}$; to satisfy (5.11) and (5.13), the B_is are assumed to be convex and l.s.c.

(1) In this section we should replace p^* by $-p^*$ in order to work with elements $p^*, \bar{p}^* \geqslant 0$, as usual in convex programming.

(2) Which follows e.g. from (5.25) when V is a reflexive Banach space.

The primal problem is

$$\text{(5.33)} \qquad \operatorname*{Inf}_{\substack{u \in \mathscr{A} \\ B_i u \leqslant 0,\ 1 \leqslant i \leqslant m}} J(u)$$

The qualification hypothesis (5.24) takes the form

(5.34) There exists $u_0 \in \mathscr{A}$, such that $B_i u_0 < 0, \quad 1 \leqslant i \leqslant m.$

Let us write

$$\text{(5.35)} \qquad L(u, p) = J(u) - \sum_{i=1}^{m} p_i B_i u, \quad \text{if} \quad p \leqslant 0 \quad \text{and} \quad u \in \mathscr{A}.$$

Theorem 5.1 then gives the Kuhn–Tucker theorem (*cf.* Kuhn and Tucker [1]).

Theorem 5.2. *With the above hypotheses, $\bar{u} \in \mathscr{A}$ is a solution of problem* (5.33) *if and only if there exists $\bar{p} \in \mathbf{R}^m$, $\bar{p} \leqslant 0$, such that*

$$\text{(5.36)} \qquad L(\bar{u}, p) \leqslant L(\bar{u}, \bar{p}) \leqslant L(u, \bar{p}), \qquad \forall u \in \mathscr{A}, \quad \forall p \geqslant 0.$$

In this case $\sum_{i=1}^{m} \langle \bar{p}_i, B_i \bar{u} \rangle = 0$, which implies that for all i, $1 \leqslant i \leqslant m$,

$$\text{(5.37)} \qquad \begin{cases} \textit{either} \quad B_i \bar{u} < 0 \quad \textit{and} \quad p_i = 0, \\ \textit{or} \quad B_i \bar{u} = 0 \quad \textit{and} \quad p_i \leqslant 0. \end{cases}$$

6. MISCELLANEOUS REMARKS

Orientation

In this paragraph we shall make two remarks, first on the bidual problem of a given problem and then on duality in variational inequalities.

6.1. The bidual problem and the generalized solution

In addition to the hypotheses of Section 1, and $\Phi \in \Gamma_0(V \times Y)$, we take two topological vector spaces V^{**} and Y^{**}, such that V^{**} and V^* (resp. Y^{**} and Y^*) are in duality and

(6.1) $V \subset V^{**}$, V dense in V^{**}, the injection being continuous

(6.2) $Y \subset Y^{**}$, Y dense in Y^{**}, the injection being continuous.

Let Φ^{**} be the conjugate functional of Φ^* in the duality between $V^* \times Y^*$ and $V^{**} \times Y^{**}$. Clearly since $\Phi \in \Gamma_0(V \times Y)$, $\Phi^{**} \in \Gamma_0(V^{**} \times Y^{**})$ is a continuation of Φ.

We now associate with the problems $(\mathscr{P})$ and $(\mathscr{P}^*)$ the bidual problem $(\mathscr{P}^{**})$ of $(\mathscr{P})$

$$(\mathscr{P}^{**}) \qquad \operatorname*{Inf}_{u \in V^{**}} \Phi^{**}(u, 0).$$

Problem $\mathscr{P}^{**}$ is the dual of $\mathscr{P}^*$ in the duality between $V^* \times Y^*$ and $V^{**} \times Y^{**}$. Using Proposition 1.1, it is easy to verify that

$$(6.3) \qquad -\infty \leqslant \sup \mathscr{P}^* \leqslant \inf \mathscr{P}^{**} \leqslant \inf \mathscr{P} \leqslant +\infty.$$

If now $\inf \mathscr{P} > \sup \mathscr{P}^{**}$, in principle there is no relation between the possible solutions of the problems $\mathscr{P}$ and $\mathscr{P}^{**}$. On the other hand, if $\inf \mathscr{P} = \inf \mathscr{P}^{**}$ then every solution of $\mathscr{P}$ is a solution of $\mathscr{P}^{**}$ and every solution of $\mathscr{P}^{**}$ which belongs to V is a solution of $\mathscr{P}$.

It is thus convenient to say that if $\inf \mathscr{P} = \inf \mathscr{P}^{**}$, problem $\mathscr{P}^{**}$ is a weak formulation of problem $\mathscr{P}$ and each solution of $\mathscr{P}^{**}$ which does not belong to V is called a weak solution of $\mathscr{P}$.

Hence it is important to have a simple criterion which determines whether

$$(6.4) \qquad \inf \mathscr{P} = \inf \mathscr{P}^{**}$$

when only problem $\mathscr{P}$ is given. But if (2.7) is true, Proposition 2.3 implies that

$$\inf \mathscr{P} = \sup \mathscr{P}^*$$

and *a fortiori* because of (6.3),

$$(6.5) \qquad \inf \mathscr{P} = \inf \mathscr{P}^{**} = \sup \mathscr{P}^*.$$

Proposition 6.1. *Let us assume that* V (resp. Y) *is a non-reflexive Banach space,* V^* (resp. Y^*) *its dual and* V^{**} (resp. Y^{**}) *its bidual. Then* (6.1) *and* (6.2) *are true. If* $\Phi \in \Gamma_0(V \times Y)$ *satisfies* (2.7), *if* dom $\Phi(., 0)$ *is a non-empty subset of* V *and*

$$(6.6) \qquad \lim_{\substack{u \in \operatorname{dom}\Phi(.,0) \\ \|u\|_V \to +\infty}} \Phi(u, 0) = +\infty,$$

then problem $\mathscr{P}$ *possesses weak solutions and all cluster points of a minimizing sequence of* $\mathscr{P}$ *are weak solutions of* $\mathscr{P}$.

Proof. A minimizing sequence of $\mathscr{P}$ is bounded in V and therefore contains a subsequence $\{u_m\}$ which converges to a limit $u \in V^{**}$ for the topology $\sigma(V^{**}, V^*)$. From the above,

$$\inf \mathscr{P}^{**} \leqslant \Phi^{**}(u, 0) \leqslant \liminf_{m' \to \infty} \Phi^{**}(u_{m'}, 0)$$
$$= \lim_{m' \to \infty} \Phi(u_{m'}, 0) = \inf \mathscr{P},$$

and thus u is a solution of $\mathscr{P}^{**}$, that is, a weak solution of $\mathscr{P}$. ■

Remark 6.1. An example where this situation arises will be given in Chapter IV, Section 3.4.

6.2. Duality in variational inequalities

We shall describe a procedure which allows us to apply certain concepts of duality to variational inequalities, *even when these do not arise from an optimization problem.*

The notation used will be that of Section II.3, with the assumption that u is a solution of the variational inequality (3.2). Then setting

(6.7) $$\xi = Au - f,$$

we have

(6.8) $$\langle \xi, v \rangle + \varphi(v) \geqslant \langle \xi, u \rangle + \varphi(u), \qquad \forall v \in V.$$

Of course (6.8) is not an optimization problem since ξ depends on u, but once u is known, ξ can be considered as a known element of V^* and (6.8) then becomes an optimization problem for which u *is* one solution

(6.9) $$\operatorname*{Inf}_{v \in V} \{ \langle \xi, v \rangle + \varphi(v) \}.$$

According to the exact form of φ, we can perturb this problem in different ways to obtain a dual problem to (6.9). ■

Example. Let us suppose for instance that

$$\varphi(v) = G(\Lambda v),$$

where $\Lambda \in \mathscr{L}(V, Y)$, Y is a Banach space, and $G \in \Gamma_0(Y)$. Setting

$$F(v) = \langle \xi, v \rangle,$$

we are in the situation of Section 4, problem (4.16).

It is easily found that

$$F^*(v^*) = \begin{vmatrix} 0 & \text{if} \quad v^* = \xi \\ +\infty & \text{otherwise.} \end{vmatrix}$$

and the dual problem of (6.9) therefore becomes

$$\text{(6.10)} \qquad \operatorname*{Sup}_{\Lambda^* p^* = \xi} \{ - G^*(- p^*) \}.$$

The relations between (6.9) and (6.10) (extremality relations, etc.) will still depend on the supplementary properties of G.

Remark 6.2. For a different approach to duality in variational inequalities, see J. Cea [2] and U. Mosco [2] [5].

PART TWO

Duality and Convex Variational Problems

This page intentionally left blank

CHAPTER IV

Applications of Duality to the Calculus of Variations (I)

Orientation

In this chapter and the next, we shall apply the results of duality to various problems of the calculus of variations arising from mechanics, physics, filtering and optimal control theory. In certain cases we shall also demonstrate some existence results for the solution of the primal problem using the results obtained in Chapter II (especially Prop. 1.2 and Theor. 3.1).

With a view to applying the concepts of duality, we shall set the problem under consideration within the framework of problem III(4.1) or more often within that of problem III(4.16): the spaces V and Y will be functional spaces of Sobolev type and Λ a differential operator. We shall make problem $\mathscr{P}^*$ [III(4.5) or (4.18)] explicit and we shall try to apply Theorems III.4.1 and III.4.2, making particular use of the extremality relations.

1. NOTATIONS AND REMINDERS

1.1. Sobolev spaces

The variable $x = (x_1, \ldots, x_n)$ will denote a point in the space $\mathbf{R}^n$. The differential operator $\partial/\partial x_i$ is denoted D_i and if $j = (j_1, \ldots, j_n)$ is a multi-integer, we write

$$D^j = D_1^{j_1} \ldots D_n^{j_n} = \frac{\partial^{|j|}}{\partial x_n^{j_1} \ldots \partial x_n^{j_n}} \tag{1.1}$$

where $|j| = j_1 + \ldots + j^n$.

If $j = (0, \ldots, 0)$ $D^j = \mathrm{I} =$ the identity.

We term Ω an open subset of $\mathbf{R}^n$, which may satisfy a regularity property of the type

(1.2) $\left\{\begin{array}{l}\text{The boundary } \Gamma \text{ of } \Omega \text{ is an } r\text{-times continuously differential} \\ \text{manifold of dimension } n-1 \text{ and } \Omega \text{ is locally situated on one} \\ \text{side only of } \Gamma.\end{array}\right.$

If (1.2) holds, we say that Ω belongs to the class $\mathscr{C}^r$.

We denote by $L^\alpha(\Omega)$ $(1 \leqslant \alpha < +\infty)$ (or $L^\infty(\Omega)$) the space of real functions of Ω into $\mathbf{R}$ whose power α is summable for the Lebesgue measure $dx = dx_1 \ldots dx_n$ (or essentially bounded over Ω). It is a Banach space with norm

$$\|f\|_{L^\alpha(\Omega)} = \left| \int_\Omega |f(x)|^\alpha \, dx \right|^{1/\alpha}$$

(or

$$\|f\|_{L^\infty(\Omega)} = \text{Ess. sup.}\,|f(x)|).$$

For $\alpha = 2$, we denote the Hilbert scalar product of $L^2(\Omega)$ by

$$(f, g) = \int_\Omega f(x) g(x) \, dx,$$

and

$$|f| = (f, f)^{1/2} = \|f\|_{L^2(\Omega)}.$$

For m an integer and $1 \leqslant \alpha \leqslant +\infty$, we denote by $W^{m,\alpha}(\Omega)$ the Sobolev space [1] [2] (*cf.* also Lions [2]) of $u \in L^\alpha(\Omega)$, all of whose derivatives of order $\leqslant m$ are in $L^\alpha(\Omega)$. It is a Banach space for the norm

$$\|u\|_{W^{m,\alpha}(\Omega)} = \left| \sum_{|j| \leqslant m} \|D^j u\|^\alpha_{L^\alpha(\Omega)} \right|^{1/\alpha}. \tag{1.3}$$

For $\alpha = 2$, we write $H^m(\Omega) = W^{m,2}(\Omega)$, and the norm (1.3) is the Hilbert norm corresponding to the scalar product

$$((u, v))_{H^m(\Omega)} = \sum_{|j| \leqslant m} (D^j u, D^j v). \tag{1.4}$$

The closure in $W^{m,\alpha}(\Omega)$ (or $H^m(\Omega)$) of the subspace of functions with compact support in Ω is denoted by $W_0^{m,\alpha}(\Omega)$ (or $H_0^m(\Omega)$ for $\alpha = 2$). This is also the closure of $\mathscr{D}(\Omega)$, the subspace of functions of Ω into $\mathbf{R}$ which are indefinitely differentiable and of compact support in Ω.

If the open space Ω is regular [e.g. (1.2) with $r = m + 2$], we can define a trace operator

$$\gamma = (\gamma_0, \ldots, \gamma_{m-1}) \tag{1.5}$$

which is linear and (in particular) continuous from $W^{m,\alpha}(\Omega)$ into $[L^\alpha(\Gamma)]^m$ (1), and such that if u is m-times continuously differentiable in $\bar{\Omega}$,

$$\gamma_0 u = u|_\Gamma, \qquad \gamma_1 u = \left.\frac{\partial u}{\partial \nu}\right|_\Gamma, \quad \ldots, \qquad \gamma_j u = \left.\frac{\partial^j u}{\partial \nu^j}\right|_\Gamma,$$

(1) $L^\alpha(\Gamma)$ = the space of functions L^α over Γ for the surface measure $d\Gamma$.

where ν is the unit vector normal to Γ and directed towards the exterior of Ω. In this case

(1.6) $$W_0^{m,\alpha}(\Omega) = \operatorname{Ker} \gamma = \text{the kernel of the mapping } \gamma.$$

For all this and other results concerning Sobolev spaces, see Sobolev [1] [2], Lions [2], Lions and Magenes [1]. Other properties of these spaces will be recalled when needed.

1.2. Calculation of a conjugate function

Let Ω be an open subset of $\mathbf{R}^n$ and g a Carathéodory mapping of $\Omega \times \mathbf{R}^l$ into $\mathbf{R}$, i.e.:

(1.7) $$\forall \xi \in \mathbf{R}^l, \qquad x \to g(x, \xi) \text{ is a measurable function,}$$

(1.8) for almost all $x \in \Omega$, $\xi \to g(x, \xi)$ is a continuous function.

We shall frequently make use of the following result of Krasnoselski [1] (Th. 2.1, p.22); the proof we give is shorter than the original one, and is due to J. M. Lasry.

Proposition 1.1. *Let E and F be two Banach spaces, Ω a Borel subset of $\mathbf{R}^n$, and $g: \Omega \times E \to F$ a Carathéodory mapping. For each measurable function $u: \Omega \to E$, let $G(u)$ be the measurable function $\Omega \ni x \mapsto g(x, u(x)) \in F$.*

If G maps $L^p(\Omega; E)$ into $L^r(\Omega; F)$ $1 \leqslant p, r < \infty$, then G is continuous in the norm topology.

Proof. Let u_n, $n \in \mathbf{N}$, be a sequence of functions in $L^p(\Omega; E)$ such that $\|u_n - \bar{u}\|_{L^p(\Omega;E)} \to 0$. We are going to show that there is a subsequence u_{n_k}, $k \in \mathbf{N}$, such that $\|G(u_{n_k}) - G(\bar{u})\|_{L^r(\Omega;E)} \to 0$, which will prove the Theorem.

Define $h: \Omega \times E \to \mathbf{R}^+$ by:

$$h(x, \eta) = \|g(x, \eta + \bar{u}(x)) - g(x, \bar{u}(x))\|_F^r.$$

Pick a subsequence u_{n_k} such that $\|u_{n_k} - \bar{u}\|_{L^p(\Omega;E)} \leqslant 2^{-k}$, and let $v_k = u_{n_k} - \bar{u}$. It follows that $v_k(x) \to 0$ almost everywhere, and hence that $h(x, v_k(x)) \to 0$ almost everywhere. In particular, since $h \geqslant 0$ for almost every $x \in \Omega$, we can find $k(x) \in \mathbf{N}$ such that:

$$\operatorname*{Sup}_k h(x, v_k(x)) = h(x, v_{k(x)}(x)).$$

Set $v_{k(x)}(x) = \bar{v}(x)$. It is easy to verify that the function $x \in \Omega \to \bar{v}(x)$ is measurable, and since

$$\int_\Omega \|\bar{v}(x)\|_E \, dx \leqslant \int_\Omega \operatorname*{Sup}_k \|v_k(x)\|_E^p \, dx \leqslant \sum_k \|v_k(x)\|_{L^p(\Omega;E)}^p < +\infty,$$

$\bar{v}$ belongs to $L^p(\Omega; E)$. From the assumption on g, it follows that the function $x \mapsto h(x, \bar{v}(x))$ belongs to $L^1(\Omega)$. From the Lebesgue convergence theorem and the inequality:

$$h(x, v_k(x)) \leqslant h(x, \bar{v}(x))$$

it follows that:

$$\textstyle\int_\Omega h(x, v_k(x))\, \mathrm{d}x \to 0. \quad \blacksquare$$

This enables us to define the functional

$$G : u \mapsto \int_\Omega g(x, u(x))\, \mathrm{d}x \tag{1.11}$$

which is continuous from V into $\mathbf{R}$.

We may wonder what the conjugate function G^* defined over the conjugate space

$$V^* = L^{\alpha'_1}(\Omega) \times \dots \times L^{\alpha'_m}(\Omega),$$

is (α'_j is given by $1/\alpha'_j + 1/\alpha_j = 1$).

The answer is given by the following proposition.

Proposition 1.2. *With the above hypotheses:*

$$G^*(v) = \int_\Omega g^*(x, v(x))\, \mathrm{d}x, \qquad \forall v \in V^*, \tag{1.12}$$

where

$$g^*(x, y) = \operatorname*{Sup}_{\eta \in \mathbf{R}^\ell} [y \cdot \eta - g(x, \eta)], \qquad \text{a.e.} \quad x \in \Omega. \tag{1.13}$$

This result is a special case of Proposition IX.2.1 and its proof requires results which we have not yet established. However, we shall use it for several examples in Chapters IV and V. (The results we shall obtain will not in any way interfere with the proof of Proposition IX.2.1.)

2. FIRST EXAMPLES

2.1. The Dirichlet problem

Given f in $L^2(\Omega)$, we seek a function u solution of

$$\left| \begin{aligned} -\Delta u &= f \quad \text{in} \quad \Omega \\ u &= 0 \quad \text{on} \quad \Gamma. \end{aligned} \right. \tag{2.1}$$

The variational form of the problem is well known: we have to find $u \in H_0^1(\Omega)$ such that

$$a(u, v) = (f, v), \qquad \forall v \in H_0^1(\Omega) \tag{2.2}$$

where

$$a(u, v) = \sum_{i=1}^{n} (D_i u, D_i v). \tag{2.3}$$

Because of Poincaré's inequality[1] the norm $\|u\| = \sqrt{a(u,u)}$ is equivalent over $H_0^1(\Omega)$ to the norm (1.3) and the Lax Milgram Lemma (Prop. II.3.4) guarantees the existence and uniqueness of a solution $\bar{u}$ of (2.2). We also know (Remark II.3.4) that $\bar{u}$ achieves the minimum in $H_0^1(\Omega)$ of

$$\tfrac{1}{2} a(u, u) - (f, u). \tag{2.4}$$

We reduce this to the situation in Chapter III (4.15), setting

$$V = H_0^1(\Omega), \qquad Y = L^2(\Omega)^n,$$
Λ = the gradient operator,
$$V^* = H^{-1}(\Omega) = \text{the dual of } H_0^1(\Omega), \qquad Y^* = Y = L^2(\Omega)^n,$$
$$F(u) = -(f, u), \quad \forall u \in V = H_0^1(\Omega),$$
$$G(p) = \frac{1}{2} \int_\Omega |p(x)|^2 \, dx$$

Problem $\mathscr{P}$ (III(4.16)) is thus identical with the problem of minimizing (2.4) over $H_0^1(\Omega)$.

It is easily seen that

$$F^*(u^*) = \sup_{u \in V} \langle u^* + f, u \rangle,^{(2)}$$
$$= \begin{cases} 0 & \text{if } u^* + f = 0, \\ +\infty & \text{otherwise.} \end{cases}$$

From Proposition I.4.2,

$$G^*(p^*) = \frac{1}{2} \int_\Omega |p^*(x)|^2 \, dx.$$

Problem $\mathscr{P}^*$, the dual of $\mathscr{P}$ (*cf.* III(4.18)) can thus be written as

$$\sup_{p^* \in L^2(\Omega)^n} [-F^*(\Lambda^* p^*) - G^*(-p^*)], \tag{2.5}$$

where Λ^* is the divergence operator.

[1] $\|u\|_{L^\alpha(\Omega)} \leqslant c(\Omega, \alpha) \|D_i u\|_{L^\alpha(\Omega)}$, $\forall u \in W_0^{1,\alpha}(\Omega)$, Ω bounded, $1 \leqslant \alpha < +\infty$.

[2] We identify $H_0^1(\Omega)$ with a subspace of $L^2(\Omega)$ and so $(f,u) = \langle f, u \rangle$, $\langle .,. \rangle$ being the duality between $H_0^1(\Omega)$ and $H^{-1}(\Omega)$.

Eliminating those elements p^* for which $F^*(\Lambda^* p^*) = +\infty$, (2.5) becomes:

$$\sup_{\substack{p^* \in L^2(\Omega)^n \\ p^* = f}} \left[-\frac{1}{2} \int_\Omega |p^*(x)|^2 \, \mathrm{d}x \right]. \tag{2.6}$$

The hypotheses of Theorem III.4.2 are satisfied, in particular III(4.12) and III(4.8) (here in the form of III(4.21)). We already know of the existence of a unique solution for $\mathscr{P}$; we have the existence of a solution $\bar{p}^*$ for $\mathscr{P}^*$ and this is unique since the functional

$$p^* \mapsto \frac{1}{2} \int_\Omega |p^*(x)|^2 \, \mathrm{d}x,$$

is strictly convex (Prop. II.1.2). Furthermore, the extremality relations III(4.22) and III(4.23) hold; the first

$$F(\bar{u}) + F^*(\Lambda^* \bar{p}^*) = \langle \Lambda^* \bar{p}^*, \bar{u} \rangle$$

is trivial here and the second

$$G(\Lambda \bar{u}) + G^*(-\bar{p}^*) = -\langle \bar{p}^*, \Lambda \bar{u} \rangle,$$

implies that

$$\int_\Omega |\operatorname{grad} \bar{u}(x)|^2 \, \mathrm{d}x + \int_\Omega |\bar{p}^*(x)|^2 \, \mathrm{d}x = -2 \int_\Omega \operatorname{grad} \bar{u}(x) \,.\, p^*(x) \, \mathrm{d}x,$$

which is only possible if:

$$\bar{p}^*(x) = -\operatorname{grad} \bar{u}(x), \quad \text{a.e.} \quad x \in \Omega.$$

Finally we have

Proposition 2.1. *The Dirichlet problem* (2.1) *written as a minimization of the Dirichlet integral* (2.4), *admits problem* (2.6) *as dual problem.*

Both admit a unique solution ($\bar{u}$ *and* $\bar{p}^*$ *respectively*): *these solutions are linked by*

$$\bar{p}^* = -\operatorname{grad} \bar{u}, \tag{2.7}$$

and in addition

$$\max \mathscr{P}^* = \min \mathscr{P}. \tag{2.8}$$

Remark 2.1. The correlation between problems (2.1) and (2.6) is well known and already appears in the early works on the calculus of variations, where it is obtained using the Legendre transformation.

2.2. The non-linear Dirichlet problem

Let α be a real number, $1 < \alpha < +\infty$, and α' the conjugate exponent such that $1/\alpha + 1/\alpha' = 1$. Given f in $L^{\alpha'}(\Omega)$, we consider the problem

$$(2.9)\qquad \left| \begin{array}{l} -\sum_{i=1}^{n} D_i(|D_i u|^{\alpha-2} D_i u) = f \quad \text{in} \quad \Omega \\ u = 0 \quad \text{on} \quad \partial\Omega. \end{array} \right.$$

Equation (2.9) is none other than the Euler equation for the minimization problem:

$$(2.10)\qquad \operatorname*{Inf}_{u \in W_0^{1,\alpha}(\Omega)} \left[\frac{1}{\alpha} \sum_{i=1}^{n} \int_\Omega |D_i u(x)|^\alpha \, dx - \int_\Omega f(x) u(x) \, dx \right].$$

Clearly, Proposition II.1.2 applies (the coerciveness II(1.7) resulting from Poincaré's inequality) and implies the existence of a solution u for (2.10). Furthermore the solution is unique since the functional

$$u \mapsto \int_\Omega |D_i u(x)|^\alpha \, dx,$$

is strictly convex ($\alpha > 1$).

As before we reduce to the situation in Chapter III(4.16), setting

$$\begin{aligned} &V = W_0^{1,\alpha}(\Omega), \qquad Y = L^\alpha(\Omega)^n, \qquad \Lambda = \text{grad}, \\ &V^* = W^{-1,\alpha'}(\Omega) = \text{the dual of } W_0^{1,\alpha}(\Omega), \qquad Y^* = L^{\alpha'}(\Omega)^n, \\ &F(u) = -(f, u), \\ &G(p) = \frac{1}{\alpha} \int_\Omega |p(x)|^\alpha \, dx. \end{aligned}$$

Problem $\mathscr{P}$ (III(4.16)) is then identical with problem (2.10) above. As in example 2.1, we have

$$F^*(\Lambda^* p^*) = \left| \begin{array}{l} 0 \quad \text{if } \operatorname{div} p^* = f, \\ +\infty \quad \text{otherwise,} \end{array} \right.$$

and by Proposition I.4.2,

$$G^*(p^*) = \frac{1}{\alpha'} \int_\Omega \sum_{i=1}^{n} |p_i^*(x)|^{\alpha'} \, dx.$$

The dual problem to $\mathscr{P}$, $\mathscr{P}^*$ (*cf.* III(4.18) or (2.5) above) can be written as

$$\text{(2.11)} \qquad \operatorname*{Sup}_{\substack{p^*\in L^{\alpha}(\Omega)^n \\ \operatorname{div} p^* = f}} \left[-\frac{1}{\alpha'} \int_\Omega \sum_{i=1}^{n} |p_i^*(x)|^{\alpha'} \, dx \right].$$

Observing that the function

$$p^* \mapsto \frac{1}{\alpha'} \int_\Omega |p^*(x)|^{\alpha'} \, dx,$$

is strictly convex over $L^{\alpha'}(\Omega)^n$, and that the qualification hypothesis III(4.21) is satisfied, we can apply Theorem III.4.2. to obtain as in Proposition 2.1 above:

Proposition 2.2. *Problem* (2.10) *has* (2.11) *as its dual problem;* (2.10) *possesses a unique solution* $\bar{u}$, (2.11) *a unique solution* $\bar{p}^*$*; we have* (*extremality relations*)

$$\text{(2.12)} \qquad \bar{p}_i^*(x) = -|D_i\bar{u}(x)|^{\alpha-2} D_i\bar{u}(x), \qquad \text{a.e. } x \in \Omega$$

and

$$\text{(2.13)} \qquad \max \mathscr{P}^* = \min \mathscr{P}.$$

Remark 2.2. A problem of this type (with a non-linear Neumann boundary condition) arises in glaciology: the solution of the primal problem determines the main velocity of the glacier and the solution of the dual problem is related to the constraints in the ice. This is developed in M. C. Pélissier and L. Reynaud [1], and M. C. Pélissier [1].

2.3. The Neumann problem

Under variational form, for f given in $L^2(\Omega)$, we have to find $u \in H^1(\Omega)$ which satisfies

$$\text{(2.14)} \qquad a(u, v) = (f, v), \qquad \forall v \in H^1(\Omega)$$

where $a(u, v) = ((u, v))$ is the scalar product (1.4) over $H^1(\Omega)$. This is the Euler equation of the problem

$$\text{(2.15)} \qquad \operatorname*{Inf}_{u\in H^1(\Omega)} \left[\tfrac{1}{2}\|u\|^2 - (f, u)\right]$$

and the existence and uniqueness of a solution $\bar{u}$ follow directly from Proposition II.1.2.

To turn to the duality situation in Chapter III (4.16), we set

$$V = H^1(\Omega), \qquad Y = L^2(\Omega)^{n+1},$$
$$\Lambda u = (u, D_1 u, \dots, D_n u), \qquad \forall u \in V,$$
$$V^* = H^1(\Omega)' = \text{the dual of } H^1(\Omega), \qquad Y^* = Y = L^2(\Omega)^{n+1},$$
$$F(u) = -(f, u),$$
$$G(p) = \frac{1}{2}\int_\Omega |p(x)|^2 \, dx.$$

Problem $\mathscr{P}$ [*cf.* III(4.16)] is identical to problem (2.15) above. We have

$$F^*(\Lambda^* p^*) = \operatorname*{Sup}_{u \in V} \langle \Lambda^* p^* + f, u \rangle$$

$$F^*(\Lambda^* p^*) = \begin{vmatrix} 0 & \text{if } \Lambda^* p^* + f = 0 \\ +\infty & \text{otherwise;} \end{vmatrix}$$

from Proposition I.4.2

$$G^*(p^*) = \frac{1}{2}\int_\Omega |p^*(x)|^2 \, dx.$$

The dual problem $\mathscr{P}^*$ of $\mathscr{P}$ [*cf.* III(4.18)] can be written

$$\operatorname*{Sup}_{\substack{p^* \in L^2(\Omega)^{n+1} \\ \Lambda^* p^* + f = 0}} \left[-\frac{1}{2}\int_\Omega |p^*(x)|^2 \, dx \right] \tag{2.16}$$

Let us specify more exactly the constraint in (2.16):

$$(f, v) + (p^*, \Lambda v) = 0, \qquad \forall v \in H^1(\Omega),$$

$$(f, v) + (p_0^*, v) + \sum_{i=1}^{n} (p_i^*, D_i v) = 0, \qquad \forall v \in H^1(\Omega), \tag{2.17}$$

which implies

$$f + p_0^* = \sum_{i=1}^{n} D_i p_i^*, \tag{2.18}$$

in the sense of the distributions in Ω. It follows from (2.18) that $\sum_{i=1}^{n} D_i p_i^* \in L^2(\Omega)$ and from a trace theorem[1] by Lions and Magenes [1], we can define the trace on Γ of $\sum_{i=1} p^* v_i$, $v = (v_1, \dots, v^n)$, the unit vector normal to Γ; from Lions and Magenes [1], we also have the formula for integration by parts

$$\int_\Gamma \left(\sum_{i=1}^{n} p_i^* v_i \right) v \, d\Gamma = \sum_{i=1}^{n} (D_i p_i^*, v) + \sum_{i=1}^{n} (p_i^*, D_i v). \tag{2.19}$$

[1] Ω regular: condition (1.2) with $r \geq 2$.

Thus (2.17) and (2.18) entail

(2.20) $$\sum_{i=1}^{n} p_i^* \nu_i = 0 \quad \text{on } \Gamma.$$

Conversely (2.18) and (2.20) imply (2.17).

Once the condition $\Lambda^* p^* + f = 0$ has been made clear, problem (2.16) becomes

(2.21) $$\operatorname*{Sup}_{\substack{p^* \in L^2(\Omega)^{n+1} \\ p_0^* + f = \sum_{i=1}^{n} D_i p_i^* \\ \sum_{i=1}^{n} p_i^* \nu_i^* |_\Gamma = 0}} \left[- \frac{1}{2} \int_\Omega |p^*(x)|^2 \, dx \right].$$

Theorem III.4.2 applies and a solution $\bar{p}^*$ of (2.21) exists and is unique. We have the extremality relation III(4.23):

$$G(\Lambda \bar{u}) + G^*(-\bar{p}^*) = - \langle \bar{p}^*, \Lambda \bar{u} \rangle,$$

which implies that

(2.22) $$\bar{p}_0^* = -\bar{u}, \qquad \bar{p}_i^* = -D_i \bar{u}, \qquad 1 \leqslant i \leqslant n,$$

and, besides,

(2.23) $$\min \mathscr{P} = \max \mathscr{P}^*.$$

2.4. Other problems

Let $\mathscr{A}$ be a closed convex subset of $H^1(\Omega)$. We shall consider $H^1(\Omega)$ endowed with the norm (1.4) denoted by $\|\,.\,\|$, and given f in $L^2(\Omega)$, we examine the problem

(2.24) $$\operatorname*{Inf}_{u \in \mathscr{A}} \left[\frac{1}{2} \|u\|^2 - (f, u) \right].$$

Different cases are possible, depending on the set $\mathscr{A}$.

In order to write this problem as a problem $(\mathscr{P})$ of the type III(4.15), we set

$$V = H^1(\Omega), \qquad Y = L^2(\Omega)^{n+1},$$

$$\Lambda u = (u, D_1 u, \ldots, D_n u), \quad \forall u \in H^1(\Omega),$$

$$V^* = H^1(\Omega)' = \text{the dual of } H^1(\Omega), \qquad Y^* = Y = L^2(\Omega)^{n+1},$$

$$F(u) = \begin{vmatrix} -(f, u) & \text{if} \quad u \in \mathscr{A} \\ +\infty & \text{otherwise}, \end{vmatrix}$$

$$G(p) = \frac{1}{2}\int_{\Omega} |p(x)|^2 \, \mathrm{d}x, \qquad \forall p \in L^2(\Omega)^{n+1}.$$

It is obvious that F is convex, l.s.c. and proper whereas G is convex, continuous and proper. Their conjugate functions can be written:

$$F^*(u^*) = \operatorname*{Sup}_{u \in \mathscr{A}} \langle u^* + f, u \rangle = \chi^*_{\mathscr{A}}(u^* + f),$$

which is the support function of the convex set $\mathscr{A}$, at the point $u^* + f$ [*cf.* I(4.4)], and

$$G^*(p^*) = \frac{1}{2}\int_{\Omega} |p^*(x)|^2 \, \mathrm{d}x.$$

The dual problem of (2.24) can be written

$$\text{(2.25)} \qquad \operatorname*{Sup}_{p^* \in L^2(\Omega)^{n+1}} \left[-\chi^*_{\mathscr{A}}(\Lambda^* p^* + f) - \frac{1}{2}\int_{\Omega} |p^*(x)|^2 \, \mathrm{d}x \right].$$

Theorem III.4.2 applies: a solution $\bar{u}$ for (2.24) exists (and is also unique) and similarly a solution $\bar{p}^*$ for (2.25) exists (and is unique); we have

$$\text{(2.26)} \qquad \min \mathscr{P} = \max \mathscr{P}^*,$$

and $\bar{u}$ and $\bar{p}^*$ are linked by the extremality relations

$$\text{(2.27)} \qquad -(f, \bar{u}) + \chi^*_{\mathscr{A}}(\Lambda^* \bar{p}^* + f) = \langle p^*, \Lambda \bar{u} \rangle,$$

$$\text{(2.28)} \qquad G(\Lambda \bar{u}) + G^*(-\bar{p}^*) = -\langle \bar{p}^*, \Lambda \bar{u} \rangle,$$

from which it follows easily that

$$\text{(2.29)} \qquad \bar{u} = -\bar{p}^*_0, \qquad D_i \bar{u} = -\bar{p}^*_i, \qquad 1 \leqslant i \leqslant n.$$

For the set $\mathscr{A}$ we can consider in particular (*cf.* Lions [3]):

$\mathscr{A} = \varphi + H^1_0(\Omega)$, φ given in $H^1(\Omega)$, and (2.24) is then a non-homogeneous Dirichlet problem;

$\mathscr{A} = \{u \in H^1(\Omega) | u \geqslant \varphi \text{ in } \Omega\}$, or

$\mathscr{A} = \{u \in H^1_0(\Omega) | u \geqslant \varphi \text{ in } \Omega\}$, φ given in $H^1(\Omega)$ ($\varphi \leqslant 0$ on Γ in the second case);

$\mathscr{A} = \{u \in H^1(\Omega) | \gamma_0 u \geqslant 0 \text{ on } \Gamma\}$, γ_0 is the trace operator on Γ (*cf.* (1.5); Ω being regular).

In all these cases, it is easy to state explicitly the dual problem (2.25) and the extremality relations (2.27) and (2.29).

2.5. The Stokes problem

Given f in $L^2(\Omega)^n$, we have to determine a vector function $u = (u_1, \ldots, u_n)$ and a scalar function p such that

$$-\Delta u + \operatorname{grad} p = f, \tag{2.30}$$

$$\operatorname{div} u = 0, \tag{2.31}$$

$$u = 0 \text{ on } \Gamma. \tag{2.32}$$

Let

$$W = \{ v \in H_0^1(\Omega)^n, \operatorname{div} v = 0 \};$$

this is a Hilbert space for the scalar product induced by $H_0^1(\Omega)^n$,

$$((u, v)) = \sum_{1 \leqslant i,j \leqslant n} (D_i u_j, D_i v_j),$$

(Ω bounded).

The variational formulation of (2.30)–(2.32) is: To find $u \in W$ such that

$$((u, v)) = (f, v), \qquad \forall v \in W. \tag{2.33}$$

From II(2.8), this problem is equivalent to the minimization problem

$$\operatorname*{Inf}_{u \in W} \left[\tfrac{1}{2} \|u\|^2 - (f, u) \right]. \tag{2.34}$$

We set:

$$\begin{aligned}
&V = H_0^1(\Omega)^n, && V^* = H^{-1}(\Omega)^n = \text{the dual of } V,\\
&Y = L^2(\Omega) = Y^*, && \Lambda = \operatorname{div},\\
&F(v) = \tfrac{1}{2} \|v\|^2 - (f, v) &&\\
&G(p) = \begin{vmatrix} 0 & \text{if } p = 0 \\ +\infty & \text{otherwise.} \end{vmatrix} &&
\end{aligned}$$

The problem

$$\operatorname*{Inf}_{u \in V} [F(v) + G(\Lambda v)] \tag{2.35}$$

is then identical with problem (2.34).

We have

$$F^*(\Lambda^* p^*) = \operatorname*{Sup}_{v \in V} [(p^*, \operatorname{div} v) + (f, v) - \tfrac{1}{2} \|v\|^2]. \tag{2.36}$$

The supremum is attained at the point $v(p^*) \in V$ which satisfies

$$((v(p^*), w)) = (f, w) + (p^*, \operatorname{div} w), \qquad \forall w \in V,$$

that is, $v(p^*)$ is solution of the problem

$$\text{(2.37)} \qquad \left| \begin{array}{l} -\Delta v(p^*) = f - \operatorname{grad} p^* \quad \text{in } \Omega, \\ v(p^*) = 0 \quad \text{on } \Gamma, \end{array} \right.$$

and hence

$$\text{(2.38)} \qquad F^*(\Lambda^* p^*) = \tfrac{1}{2} \|v(p^*)\|^2.$$

The conjugate function of G is $G^* = 0$, and the dual problem of (2.35)

$$\text{(2.39)} \qquad \operatorname*{Sup}_{p^* \in L^2(\Omega)^n} [- F^*(\Lambda^* p^*) - G^*(- p^*)]$$

can be written as

$$\text{(2.40)} \qquad \operatorname*{Sup}_{p^* \in L^2(\Omega)} [- \tfrac{1}{2} \|v(p^*)\|^2].$$

From Remark II.1.1, problem (2.34) (or (2.35)) possesses a unique solution. Since hypothesis III(4.8) is not satisfied, Theorem III.4.1 is not directly applicable. On the other hand, since the function $F^*$is continuous at all points $\Lambda^* p^*$, the analogous hypothesis to III(4.8) for problem $\mathscr{P}^*$ is satisfied, which implies that problem $\mathscr{P}^*$ is stable and yields:[1]

$$\text{(2.41)} \qquad \inf \mathscr{P} = \sup \mathscr{P}^*.$$

The existence of a solution for $\mathscr{P}^*$ can be demonstrated directly: if p_m^* is a maximizing sequence, then

$$\|v(p_m^*)\| \leqslant c,$$

the sequence $v(p_m^*)$ is bounded in $H_0^1(\Omega)^n$, and by (2.37), grad p_m^* is bounded in $H^{-1}(\Omega)^n$, so p^* is bounded in $L^2(\Omega)/\mathbf{R}$. There exists a suitable subsequence (still denoted p_m^*), which converges in $L^2(\Omega)/\mathbf{R}$ to some limit $\bar{p}^*$ and we verify that $\bar{p}^*$ is a solution of $\mathscr{P}^*$; there is, however, no unique solution of $\mathscr{P}^*$[2]. The extremality relation is none other than (2.30).

Proposition 2.3. *Problem* (2.30)–(2.32) *possesses a solution* $\{\bar{u}, \bar{p}^*\}$ *where* $\bar{u}$ *is solution of the minimization problem* (2.34) *and* $\bar{p}^*$ *is solution of the dual problem* (2.40). *Moreover, the problems* $\mathscr{P}$ *and* $\mathscr{P}^*$ *are linked by the relation* (2.41).

(1) In addition to the existence (already known) of a solution of $\mathscr{P}$.

(2) The solution will be unique except for additive constants.

3. NON-DIFFERENTIABLE FUNCTIONALS

We shall consider in detail here two examples where the functions to be minimized are non-differentiable. In the first example (Section 3.1), duality allows us to display functions which have a fundamental physical significance (the constraints) for the problem. In the second example (Section 3.2, *cf.* Remark 3.4), duality permits the introduction of the Euler equation of the problem, which is non-standard since the functionals are non-differentiable.

3.1. Mossolov's problem

We now consider the problem (*cf.* Lions [4], Mossolov and Miasnikov [1])

$$\operatorname*{Inf}_{u\in H_0^1(\Omega)}\left[\frac{\alpha}{2}\int_\Omega |\operatorname{grad} u(x)|^2\,dx + \beta\int_\Omega |\operatorname{grad} u(x)|\,dx - \int_\Omega f(x)u(x)\,dx\right], \tag{3.1}$$

where α and β are constants >0 and $f\in L^2(\Omega)$ is given. We shall see that this problem can be put into duality in several different ways, leading us back by different routes to the situation in Chapter III (4.16).

We shall take $V = H_0^1(\Omega)$ endowed with the Hilbert scalar product (2.3) (Ω is bounded), $Y = L^2(\Omega)^n$, $\Lambda =$ the gradient operator,

$$V^* = H^{-1}(\Omega) = \text{the dual of } H_0^1(\Omega), \qquad Y^* = Y = L^2(\Omega)^n,$$

$$F(u) = \frac{\alpha}{2}\|u\|^2 - (f, u), \qquad \|\cdot\| = \text{norm of } H_0^1(\Omega),$$

$$G(p) = \beta\int_\Omega |p(x)|\,dx, \qquad \forall p\in Y.$$

Then problem $\mathscr{P}$:

$$\operatorname*{Inf}_{u\in V}\,[F(u) + G(\Lambda u)] \tag{3.2}$$

is identical with problem (3.1).

To state the dual problem explicitly, we calculate F^* and G^*:

$$F^*(u^*) = \operatorname*{Sup}_{u\in H_0^1(\Omega)}\left[\langle u^* + f, u\rangle - \frac{\alpha}{2}\|u\|^2\right]. \tag{3.3}$$

The maximum is attained at the point u which is solution of the Dirichlet problem

$$\left|\begin{array}{l} -\alpha\,\Delta u = u^* + f \\ u \in H_0^1(\Omega) \end{array}\right. \tag{3.4}$$

and the value of the maximum is

$$F^*(u^*) = \frac{1}{2\alpha}\,\|u^* + f\|^2_{H^{-1}(\Omega)} \tag{3.5}$$

where the norm of $H^{-1}(\Omega)$ is the norm dual to that of $H_0^1(\Omega)$.

$$G^*(p^*) = \operatorname*{Sup}_{p \in L^2(\Omega)^n} \int_\Omega [p^*(x)p(x) - \beta\,|p(x)|]\,\mathrm{d}x \tag{3.6}$$

$$G(p) = \left|\begin{array}{ll} 0 \text{ if } |p^*(x)| \leqslant \beta, & \text{a.e.} \\ +\infty & \text{otherwise.} \end{array}\right. \tag{3.7}$$

Problem $\mathscr{P}^*$ can then be written as

$$\operatorname*{Sup}_{p^* \in Y^*} [-F^*(\Lambda^* p^*) - G^*(-p^*)], \tag{3.8}$$

where, eliminating those values of p^* for which the functional takes the value $-\infty$,

$$\operatorname*{Sup}_{\substack{p^* \in L^2(\Omega)^n \\ |p^*(x)| \leqslant \beta \text{ a.e.}}} \left[-\frac{1}{2\alpha}\,\|f - \operatorname{div} p^*\|^2_{H^{-1}(\Omega)}\right]. \tag{3.9}$$

Theorem III.4.2 applies: there is a solution $\bar{u}$ of (3.1) (which is unique by strict convexity), a solution $\bar{p}^*$ of (3.2) (not necessarily unique), and these are linked by the extremality relations

$$\left|\begin{array}{l} F(\bar{u}) + F^*(\Lambda^*\bar{p}^*) = \langle \Lambda^*\bar{p}^*, \bar{u} \rangle \\ G(\Lambda\bar{u}) + G^*(-\bar{p}^*) = -\langle \bar{p}^*, \Lambda\bar{u} \rangle. \end{array}\right. \tag{3.10}$$

With (3.4) the first relation yields

$$-\alpha\,\Delta\bar{u} + \operatorname{div}\bar{p}^* = f. \tag{3.11}$$

The second relation means that

$$\int_\Omega [\beta\,|\Lambda\bar{u}(x)| + \bar{p}^*(x)\cdot\Lambda u(x)]\,\mathrm{d}x = 0,$$

and since the integral is $\geqslant 0$, we denote that:

$$(3.12) \qquad \beta\,|\mathrm{grad}\,\bar{u}(x)| = -\bar{p}^*(x).\mathrm{grad}\,\bar{u}(x), \quad \text{a.e.}$$

In the present case, the solution of the dual problem is used in characterizing the solution of the primal problem. Conversely let us assume that $\bar{u} \in H_0^1(\Omega)$ and that there exists $\bar{p}^* \in L^2(\Omega)^n$, such that we have (3.11)–(3.12) and

$$(3.13) \qquad |\bar{p}^*(x)| \leqslant \beta \qquad \text{a.e.};$$

then it can easily be seen that the extremality relations (3.10) are satisfied and Proposition III.4.1 implies that $\bar{u}$ is a solution of $\mathscr{P}$ and $\bar{p}^*$ a solution of $\mathscr{P}^*$.

Proposition 3.1. *Problems* 3.1 *and* 3.9 *are mutually dual,*

$$(3.14) \qquad \min \mathscr{P} = \max \mathscr{P}^*.$$

Problem (3.1) *possesses a unique solution $\bar{u}$ and problem* (3.9) *possesses at least one solution $\bar{p}^*$.*

An element $\bar{u} \in H_0^1(\Omega)$ is a solution of (3.1) *if and only if there exists $\bar{p}^* \in L^2(\Omega)^n$, which satisfies* (3.11), (3.12) *and* (3.13).

Another method of dualization

Again we set $V = H_0^1(\Omega)$, $V^* = H^{-1}(\Omega)$, $Y = Y^* = L^2(\Omega)^n$, $\Lambda = \mathrm{grad}$, but the functions F and G are chosen in a different way:

$$(3.15) \qquad F(u) = -(f, u), \qquad \forall u \in H_0^1(\Omega),$$

$$(3.16) \qquad G(p) = \int_\Omega \left[\frac{\alpha}{2}|p(x)|^2 + \beta\,|p(x)|\right] \mathrm{d}x.$$

With this choice of F and G, problem (3.2) is identical with problem (3.1). Let us calculate F^* and G^*:

$$(3.17) \qquad F^*(u^*) = \mathrm{Sup} \langle u^* + f, u \rangle = \begin{cases} 0 & \text{if } \bar{u}^* + f = 0 \\ +\infty & \text{otherwise.} \end{cases}$$

$$(3.18) \qquad G^*(p^*) = \sup_{p \in L^2(\Omega)^n} \int_\Omega \left[p^*(x)p(x) - \frac{\alpha}{2}|p(x)|^2 - \beta\,|p(x)|\right] \mathrm{d}x.$$

From Proposition 1.2,

$$(3.19) \qquad G^*(p^*) = \int_\Omega \left\{\sup_{\xi \in \mathbb{R}^n}\left[p^*(x)\xi - \frac{\alpha}{2}|\xi|^2 - \beta\,|\xi|\right]\right\} \mathrm{d}x.$$

Let us calculate the supremum contained in the latter integral. For the direction, the optimum is obtained at

$$\xi = \rho \frac{p^*(x)}{|p^*(x)|}$$

and is equal to

$$|p^*(x)|\, \rho - \frac{\alpha}{2}\rho^2 - \beta\rho\,;$$

we now have to maximize this expression for $\rho \geqslant 0$. We must distinguish two cases, according to whether $|p^*(x)| \leqslant$ or $\geqslant \beta$. In the first case, the maximum is attained at $\rho = 0$ and is equal to 0, and in the second case, it is attained at

$$\rho = \frac{1}{\alpha}(\,|p^*(x)| - \beta)$$

and is equal to

$$\frac{1}{2\alpha}(\,|p^*(x)| - \beta)^2\,;$$

finally:[1]

$$G^*(p^*) = \frac{1}{2\alpha}\int_\Omega (\,|p^*(x)| - \beta)_+^2 \, dx \,{}^{(2)}. \tag{3.20}$$

With the choice (3.15)–(3.16) of F and G, the dual problem of (3.1) is

$$\underset{\substack{p^* \in L^2(\Omega)^n \\ \operatorname{div} p^* = f}}{\operatorname{Sup}} \left[-\frac{1}{2\alpha}\int_\Omega (\,|p^*(x)| - \beta)_+^2 \, dx \right]. \tag{3.21}$$

Theorem III.4.2 again applies; the first extremality relation (3.10) is trivial and the second gives us:

$$\operatorname{grad} \bar{u}(x) = -\frac{\bar{q}^*(x)}{\alpha\,|\bar{q}^*(x)|}(\,|\bar{q}^*(x)| - \beta)_+, \qquad \text{a.e.,} \tag{3.22}$$

where $\bar{q}^*$ denotes a solution of (3.21).

[1] The maximum in (3.19) is attained at $\xi = 0$ if $|p^*(x)| \leqslant \beta$ and at

$$\xi = \frac{p^*(x)}{\alpha\,|p^*(x)|}(|p^*(x)| - \beta)$$

if $|p^*(x)| \geqslant \beta$.

[2] $s_+ = s$ if $s \geqslant 0$, 0 if $s \leqslant 0$; $s_- = (-s)_+$; s_+^2 means $(s_+)^2$.

Proposition 3.2. *Problems* (3.1) *and* (3.2) *are mutually dual,*

$$\min \mathscr{P} = \max \mathscr{P}^*. \tag{3.23}$$

Problem (3.21) *possesses at least one solution. An element* $\bar{u}$ *of* $H_0^1(\Omega)$ *is a solution of* (3.1) *if and only if there exists* $\bar{q}^* \in L^2(\Omega)^n$, *satisfying* $\operatorname{div} \bar{q}^* = f$ *and* (3.22).

Remark 3.1. Let A be a non-linear operator of $V = H_0^1(\Omega)$ into $V^* = H^{-1}(\Omega)$, satisfying the hypotheses of Theorem II.3.1 (II(3.3), II(3.4) or II(3.19), II(3.5)) and let

$$\phi(v) = \beta \int_\Omega |\operatorname{grad} v(x)| \, dx.$$

Theorem II.3.1 implies the existence of $u \in H_0^1(\Omega)$ such that

$$\langle A\bar{u} - f, v - \bar{u} \rangle + \phi(v) - \phi(\bar{u}) \geqslant 0, \qquad \forall v \in H_0^1(\Omega); \tag{3.24}$$

we set

$$\sigma = f - A\bar{u}; \tag{3.25}$$

we then have from (3.24),

$$\sigma \in \partial\phi(\bar{u}). \tag{3.26}$$

Recalling the technique of duality which led to (3.9), we set

$$F(v) = - \langle \sigma, v \rangle,$$

$$G(p) = \beta \int_\Omega |p(x)| \, dx;$$

and obtain the existence of $\bar{p}^* \in L^2(\Omega)^n$, such that

$$|\bar{p}^*(x)| \leqslant \beta \qquad \text{a.e.}, \tag{3.27}$$

$$\beta \, |\operatorname{grad} \bar{u}(x)| = - \bar{p}^*(x) . \operatorname{grad} \bar{u}(x) \qquad \text{a.e.}, \tag{3.28}$$

$$A\bar{u} + \operatorname{div} \bar{p}^* = f. \tag{3.29}$$

We note that analogous results can be obtained with other operators A and possibly with other spaces V; in particular for the time-dependent case,

$$V = L^2([0, T]; H_0^1(\Omega)), \qquad \text{and} \qquad Au = \frac{\partial u}{\partial t} - \alpha \, \Delta u.$$

Remark 3.2. A situation completely analogous to the foregoing appears in the mathematical theory of non-newtonian fluid flows of Bingham type, and the $\bar{p}^*$s are then directly related to the stress tensor (see Duvaut and Lions [1]).

3.2. A problem in filtering theory (1) (*cf.* Berkovitz and Pollard [1] [2])

Let Ω be an open subset of $\mathbf{R}^n$, *which may or may not be bounded,* and Φ a real function in Ω; the following problem arises in the theory of optimal filtering:

$$\text{(3.30)} \qquad \operatorname*{Inf}_{\substack{u=\phi, \partial u/\partial\nu = \partial\phi/\partial\nu \\ \text{on } \Gamma}} \frac{1}{2}\left[\int_\Omega |\Delta u|^2 \, dx + \left(\int_\Omega |u| \, dx\right)^2\right].$$

Berkovitz and Pollard in the cited reference considered the case $n = 1$, and $\Omega =]0, +\infty[$; we shall examine here the more general case where n is arbitrary. Before forming the dual of (3.30) we shall start by giving an exact formulation of this problem and by demonstrating the existence and uniqueness of solutions, a result which does not follow immediately from Chapter II.

We call $\mathscr{H}(\Omega)$ the space

$$\text{(3.31)} \qquad \{\, u \mid u \in L^1(\Omega), \Delta u \in L^2(\Omega) \,\}$$

which is a Banach space when equipped with the "natural" norm:

$$\text{(3.32)} \qquad \|u\|_{L^1(\Omega)} + \|\Delta u\|_{L^2(\Omega)}.$$

We call $\mathscr{H}_0(\Omega)$ the closure in $\mathscr{H}(\Omega)$ of the subspace $\mathscr{D}(\Omega)$ of indefinitely differentiable functions with a compact support in Ω; $\mathscr{H}_0(\Omega)$ is the set of $u \in \mathscr{H}(\Omega)$ which satisfy in a weak sense:[1]

$$\text{(3.33)} \qquad u = 0, \qquad \frac{\partial u}{\partial \nu} = 0 \quad \text{on} \quad \Gamma.$$

Given $\Phi \in \mathscr{H}(\Omega)$, the precise formulation of (3.30) is

$$\text{(3.34)} \qquad \operatorname*{Inf}_{u \in \phi + \mathscr{H}_0(\Omega)} \frac{1}{2}\left[\int_\Omega |\Delta u|^2 \, dx + \left(\int_\Omega |u| \, dx\right)^2\right].$$

Proposition 3.3. *Problem* (3.34) *possesses a unique solution.*

Before proving this result, we shall establish some lemmas.

Lemma 3.1. *If* $u \in \mathscr{H}_0(\Omega)$, *then* $\tilde{u} \in \mathscr{H}(R^n)$, $\tilde{u}$ *being the function u extended by* 0 *outside* Ω.

Proof. The mapping $\psi \in \mathscr{D}(\Omega) \mapsto \tilde{\psi} \in \mathscr{H}(\mathbf{R}^n)$, being an isometry ($\mathscr{D}(\Omega)$ endowed with the norm induced by $\mathscr{H}(\Omega)$, can be extended as a linear con-

[1] Cf. Lemmas 3.1, 3.2 and 3.3 hereafter.

tinuous mapping of $\mathscr{H}_0(\Omega)$ into $\mathscr{H}(\mathbf{R}^n)$. If $u \in \mathscr{H}_0(\Omega)$, the image of u under this mapping is $\tilde{u}$.

We note that:

(3.35) $$\Delta\tilde{u} = \widetilde{\Delta u}, \qquad \forall u \in \mathscr{H}_0(\Omega). \blacksquare$$

Lemma 3.2.

(3.36) $\mathscr{H}(\mathbf{R}^n) \subset H^2(\mathbf{R}^n)$ *and the embedding is continuous.*

Proof. If $u \in \mathscr{H}(\mathbf{R}^n)$, then its Fourier transform $\hat{u} \in L^\infty(\mathbf{R}^n)$ and $\xi^2\hat{u}(\xi) \in L^2_\xi(\mathbf{R}^n)$. We shall now show that $u \in L^2_\xi(\mathbf{R}^n)$:

$$\begin{aligned}
\int_{\mathbf{R}^n} |\hat{u}(\xi)|^2 \, d\xi &= \int_{|\xi| \leqslant 1} |\hat{u}(\xi)|^2 \, d\xi + \int_{|\xi|>1} |\hat{u}(\xi)|^2 \, d\xi \\
&\leqslant c \, \|\hat{u}\|^2_{L^\infty_\xi(\mathbf{R}^n)} + \int_{|\xi|>1} \xi^2 \, |\hat{u}(\xi)|^2 \, d\xi \\
&\leqslant c \, \|\hat{u}\|^2_{L^\infty_\xi(\mathbf{R}^n)} + \|\xi^2\hat{u}\|^2_{L^2(\mathbf{R}^n)} \\
&\leqslant c \, \|u\|^2_{\mathscr{H}(\mathbf{R}^n)}.
\end{aligned}$$

It follows easily that $\xi_i\hat{u} \in L^2_\xi(\mathbf{R}^n)$, $1 \leqslant i \leqslant n$ and

$$\|\xi_i\hat{u}(\xi)\|^2_{L^2(\mathbf{R}^n)} \leqslant c \, \|u\|^2_{\mathscr{H}(\mathbf{R}^n)} \; ;$$

(3.36) results from this. $\blacksquare$

Lemma 3.3. *If Ω is regular, of class $\mathscr{C}^2$, then $u \in \mathscr{H}_0(\Omega)$ if and only if $u \in \mathscr{H}(\Omega)$ and $\tilde{u} \in H^2(R^n)$.*

Proof. Lemmas 3.1 and 3.2 show that if $u \in \mathscr{H}_0(\Omega)$, $\tilde{u} \in H^2(\mathbf{R}^n)$.

Conversely, if $u \in \mathscr{H}(\Omega)$ and $\tilde{u} \in H^2(\mathbf{R}^n)$, it can easily be shown that $\tilde{u}$ is the limit in $\mathscr{H}(\mathbf{R}^n)$ of functions with a bounded support. Indeed, let $\theta \in \mathscr{D}(\mathbf{R}^n)$, $0 \leqslant \theta \leqslant 1$, $\theta(x) = 1$ for $|x| \leqslant 1$, $\theta(x) = 0$ for $|x| \geqslant 2$ and let θ_R:

(3.37) $$\theta_R(x) = \theta(x/R) \, ;$$

when $R \to \infty$, $\theta_R u \to u$ in $\mathscr{H}(\Omega)$ and $\theta_R\tilde{u} \to \tilde{u}$ in $H^2(\mathbf{R}^n)$.

We have now to prove that if $u \in \mathscr{H}(\Omega)$ is a function with a bounded support and if $\tilde{u} \in H^2(\mathbf{R}^n)$, then u is the limit in $\mathscr{H}(\Omega)$ of functions of $\mathscr{D}(\Omega)$. Since $\tilde{u} \in H^2(\mathbf{R}^n)$, $u \in H^2_0(\Omega)$ (from the trace theorem); thus u is the limit in $H^2_0(\Omega)$ of functions of $\mathscr{D}(\Omega)$. We can restrict outselves to functions with support in a bounded open set Ω', $\overline{\Omega'} \subset \Omega$, $\Omega' \supset \operatorname{supp} u$. The result follows then from the fact that $H^2_0(\Omega') \subset \mathscr{H}(\Omega')$ with a continuous imbedding, when Ω' is bounded.

$\blacksquare$

Proof of Proposition 3.3.

Let u_m be a minimizing sequence of (3.34). It will be of the form

$$u_m = \phi + v_m, \qquad v_m \in \mathscr{H}_0(\Omega), \tag{3.38}$$

and

$$J(u_m) = \frac{1}{2}\left[\int_\Omega |\Delta u_m|^2 \, dx + \left(\int_\Omega |u_m| \, dx\right)^2\right] \to \inf \mathscr{P}. \tag{3.39}$$

It follows from (3.39) that $J(u_m)$ is bounded from above. The sequence u_m and hence v_m are bounded in $\mathscr{H}(\Omega)$:

$$\|\Delta v_m\|_{L^2(\Omega)} \leqslant c, \tag{3.40}$$

$$\|v_m\|_{L^1(\Omega)} \leqslant c. \tag{3.41}$$

From Lemmas 3.1 and 3.2, the sequence $\tilde{v}_m$ is bounded in $H^2(\mathbf{R}^n)$

$$\|\tilde{v}_m\|_{H^2(\mathbf{R}^n)} \leqslant c. \tag{3.42}$$

From (3.40)–(3.42), we can extract a sequence m' from m such that

$$\tilde{v}_{m'} \to w, \quad \text{weakly in} \quad H^2(\mathbf{R}^n). \tag{3.43}$$

Since the imbedding of $H^2(\mathcal{O})$ into $H^1(\mathcal{O})$ is compact for every bounded ball $\mathcal{O}$ (*cf.* Lions and Magenes [1]), (3.43) implies that

$$\tilde{v}_m|_{\mathcal{O}} \to w|_{\mathcal{O}} \quad \text{strongly in} \quad H^1(\mathcal{O}). \tag{3.44}$$

Using the diagonal process, we can choose the sequence m' so that

$$\tilde{v}_{m'}(x) \to w(x) \tag{3.45}$$

From (3.41), (3.45) and Fatou's Lemma, we have

$$\int_{\mathbf{R}^n} |w(x)| \, dx \leqslant \varliminf_{m' \to \infty} \int_{\mathbf{R}^n} |\tilde{v}_{m'}(x)| \, dx \leqslant c < +\infty, \quad \text{a.e.} \tag{3.46}$$

which implies that $w \in \mathscr{H}(\mathbf{R}^n)$; of course $v = w|_\Omega \in \mathscr{H}(\Omega)$, and as $w = 0$ in $\complement\Omega$ (by (3.45)), $\tilde{v} = w$. Lemma 3.3 proves then that $v \in \mathscr{H}_0(\Omega)$.

Now let $\bar{u} = \phi + v \in \phi + \mathscr{H}_0(\Omega)$; by (3.46), (3.49) and the weak lower semi-continuity of the L^2 norm, we find that

$$J(\bar{u}) = \inf_{u \in \phi + \mathscr{H}_0(\Omega)} J(u),$$

which shows that $\bar{u}$ is a solution of (3.34).

To demonstrate the uniqueness of u, let us suppose that $\bar{u}_1$ and $\bar{u}_2$ are two solutions of (3.34) and let $u = \bar{u}_1 - \bar{u}_2$. From the strict convexity of the function

$$u \to \int_\Omega |\Delta u|^2 \, dx,$$

we find that $\Delta\bar{u}_1 = \Delta\bar{u}_2$. Thus $u \in \mathscr{H}_0(\Omega)$ and $\Delta u = 0$. The function $\tilde{u}$ (u extended by 0 beyond Ω) is in $\mathscr{H}(\mathbf{R}^n)$ and from (3.35) satisfies $\Delta\tilde{u} = 0$; thus this function $\tilde{u}$ is analytic and, being zero on $\complement\Omega$, is zero everywhere. ∎

Remark 3.3. The result of Proposition 3.3 is due to Berkovitz and Pollard [1] when $n = 1$ and $\Omega =]0,+\infty[$; it is new for the general case.

Duality for the problem (3.30)–(3.34)

We go back to the situation in Chapter III(4.16), by setting

$$\begin{aligned}
V &= \mathscr{H}(\Omega), & Y &= L^1(\Omega) \times L^2(\Omega),\\
\Lambda u &= \{ u, \Delta u \}, & &\forall u \in \mathscr{H}(\Omega),\\
V^* &= \mathscr{H}(\Omega)^* = \text{ the dual of } \mathscr{H}(\Omega), & &\\
Y^* &= L^\infty(\Omega) \times L^2(\Omega), & &
\end{aligned}$$

$$F(u) = \begin{cases} 0 & \text{if } u \in \phi + \mathscr{H}_0(\Omega),\\ +\infty & \text{otherwise,} \end{cases}$$

(F = the indicator function of $\phi + \mathscr{H}_0(\Omega)$),

$$G(p) = G_0(p_0) + G_1(p_1),$$

$$G_0(p_0) = \tfrac{1}{2}\left(\int_\Omega |p_0| \, dx\right)^2 = \tfrac{1}{2}\|p_0\|^2_{L^1(\Omega)},$$

$$G_1(p_1) = \tfrac{1}{2}\int_\Omega |p_1|^2 \, dx = \tfrac{1}{2}\|p_1\|^2_{L^2(\Omega)}.$$

The problem

$$\text{(3.47)} \qquad \operatorname*{Inf}_{u\in V}\left[F(u) + G(\Lambda u)\right]$$

is thus identical to (3.34). Let us determine the conjugate functions.

From Remark III.4.2,

$$\text{(3.48)} \qquad G^*(p^*) = G_0^*(p_0^*) + G_1^*(p_1^*),$$

and from Proposition I.4.2,

(3.49) $$G_0^*(p_0^*) = \frac{1}{2} \|p_0^*\|^2_{L^\infty(\Omega)}, \qquad \forall p_0^* \in L^\infty(\Omega),$$

(3.50) $$G_1^*(p_1^*) = \frac{1}{2} \|p_1^*\|^2_{L^2(\Omega)}, \qquad \forall p_1^* \in L^2(\Omega),$$

$$F^*(\Lambda^* p^*) = \operatorname*{Sup}_{u \in \phi + \mathscr{H}_0(\Omega)} \langle p^*, \Lambda u \rangle$$
$$= \langle p^*, \Lambda\phi \rangle + \operatorname*{Sup}_{v \in \mathscr{H}_0(\Omega)} \int_\Omega (p_0^* v + p_1^* \Delta v) \mathrm{d}x.$$

Since $\mathscr{D}(\Omega)$ is dense in $\mathscr{H}_0(\Omega)$, we have

$$F^*(\Lambda^* p^*) = \langle p^*, \Lambda\phi \rangle + \operatorname*{Sup}_{v \in \mathscr{D}(\Omega)} \int_\Omega (p_0^* v + p_1 \Delta v) \, \mathrm{d}x,$$

and the supremum is equal to 0 or $+\infty$ according to whether $\Delta p_1^* + p_0^* = 0$ (in the distribution sense) or not:

(3.51) $$F^*(\Lambda^0 p^*) = \begin{cases} \langle p^*, \Lambda\phi \rangle & \text{if } \Delta p_1^* + p_0^* = 0 \\ +\infty & \text{otherwise.} \end{cases}$$

Problem $\mathscr{P}^*$, the dual of (3.34), can be written

$$\operatorname*{Sup}_{p^* \in Y^*} [-F^*(\Lambda^* p^*) - G^*(-p^*)].$$

Eliminating those values of p^* for which the functional is equal to $-\infty$, we find from (3.48) and (3.51):

(3.52) $$\operatorname*{Sup}_{\substack{p^* \in L^\infty(\Omega) \times L^2(\Omega) \\ \Delta p_1^* + p_0^* = 0}} \left[-\langle p^*, \Lambda\phi \rangle - \frac{1}{2} \|p_0\|^2_{L^\infty(\Omega)} - \frac{1}{2} \|p_1^*\|^2_{L^2(\Omega)} \right].$$

Theorem III.4.1 applies, the stability condition III(4.8) being easily verified: problem (3.52) possesses at least one solution $\bar{p}^*$ and $\inf \mathscr{P} = \sup \mathscr{P}^*$. In fact the solution of (3.52) is unique; if $\bar{p}^*$ and $\bar{q}^*$ are two different solutions, we obtain from the strict convexity of the function $p_1^* \mapsto \frac{1}{2}\|p_1^*\|^2_{L^2(\Omega)}$ $\bar{p}_1^* = \bar{q}_1^*$ and so $\bar{p}^* = \bar{q}^*$ since $\bar{p}_0^* = -\Delta \bar{p}_1^* = -\Delta \bar{q}_1^* = \bar{q}_0^*$.

We apply now the extremality relations (Proposition III.4.1, Remarks III.4.1 and III.4.2):

(3.53) $$F(\bar{u}) + F^*(\Lambda^* \bar{p}^*) = \langle \Lambda^* \bar{p}^*, \bar{u} \rangle,$$

holds automatically, but we also have

(3.54) $$\frac{1}{2} \|\bar{u}\|^2_{L^1(\Omega)} + \frac{1}{2} \|\bar{p}_0^*\|^2_{L^\infty(\Omega)} = -\langle \bar{p}_0^*, \bar{u} \rangle,$$

$$(3.55)\qquad \tfrac{1}{2}\,\|\Delta\bar{u}\|^2_{L^2(\Omega)} + \tfrac{1}{2}\,\|\bar{p}^*_1\|^2_{L^2(\Omega)} = -\langle \bar{p}^*_1, \Delta\bar{u}\rangle.$$

Relation (3.55) obviously implies

$$(3.56)\qquad \Delta\bar{u} = -\bar{p}^*_1.$$

On the other hand

$$-\langle \bar{p}^*_0, \bar{u}\rangle = -\int_\Omega \bar{p}^*_0\bar{u}\,\mathrm{d}x \leqslant \left(\int_\Omega |\bar{u}|\,\mathrm{d}x\right)\cdot \|p^*_0\|_{L^\infty(\Omega)}$$
$$\leqslant \tfrac{1}{2}\,\|\bar{p}^*_0\|^2_{L^\infty(\Omega)} + \tfrac{1}{2}\,\|\bar{u}\|^2_{L^1(\Omega)},$$

and from (3.54) *all these inequalities are in fact equalities*, which implies successively that

$$(3.57)\qquad -\bar{p}^*_0(x)\bar{u}(x) = \|\bar{p}^*_0\|_{L^\infty(\Omega)}\,|\bar{u}(x)| \qquad \text{a.e.},$$

$$(3.58)\qquad \|\bar{p}^*_0\|_{L^\infty(\Omega)} = \|\bar{u}\|_{L^1(\Omega)}.$$

Thus

$$(3.59)\qquad -\bar{p}^*_0(x) = \|\bar{u}\|_{L^1(\Omega)}[\operatorname{sgn}\bar{u}(x)], \qquad \text{a.e. when} \quad \bar{u}(x)\neq 0.$$

Since $\Delta\bar{p}^*_1 = -\bar{p}^*_0$, (3.56) implies that $\Delta^2\bar{u} = \bar{p}^*_0 \in L^\infty(\Omega)$ and (3.59) implies that

$$(3.60)\qquad \Delta^2\bar{u}(x) = -\|\bar{u}\|_{L^1(\Omega)}\cdot[\operatorname{sgn}\bar{u}(x)], \qquad \text{a.e. when} \quad \bar{u}(x)\neq 0.$$

Proposition 3.4. (i) *Problems* (3.34) *and* (3.52) *are mutually dual and*

$$(3.61)\qquad \inf\mathscr{P} = \sup\mathscr{P}^*.$$

(ii) *Problem* (3.34) *possesses a unique solution* $\bar{u}$, *problem* (3.52) *a unique solution* $\bar{p}^*$ *and these solutions are linked by the extremality relations* (3.56), (3.57) *and* (3.58).

(iii) $u \in \phi + \mathscr{H}_0(\Omega)$ *is solution of* (3.34) *if and only if*

$$(3.62)\qquad \Delta^2 u \in L^\infty(\Omega)$$

and

$$(3.63)\qquad \Delta^2 u(x) = -\|u\|_{L^1(\Omega)}[\operatorname{sgn} u(x)], \qquad \textit{a.e. when} \quad \bar{u}(x)\neq 0.$$

Proof. Everything has already been established except (iii). It follows from (ii) that (3.62) and (3.36) are necessary conditions. Conversely, if $u \in \phi + \mathscr{H}_0(\Omega)$ satisfies these relations, let

$$p^*_1 = -\Delta u \in L^2(\Omega),$$

and

$$p_0^* = -\Delta p_1^* = \Delta^2 u \in L^\infty(\Omega).$$

It is easily verified that the extremality relations (3.53), (3.54) and (3.55) are satisfied by u and $p^* = (p_0^*, p_1^*)$, which, with Proposition III.4.2, implies that u is the solution of $\mathscr{P}$ (and p^* the solution of $\mathscr{P}^*$). ■

Remark 3.4. Equation (3.62) together with (3.63) constitute the Euler equation of problem (3.34). It has been obtained here using the extremality relations; it can also be obtained by a direct argument. On the other hand, the characterization of the solution of the problem by (3.62), (3.63) (and $u \in \phi + \mathscr{H}_0(\Omega)$) seems to be new, even in the special case where $n = 1$ and $\Omega =]0, \infty[$.

Remark 3.5. The solution of (3.34) has been determined explicitly in Berkovitz and Pollard [1] [2], when $n = 1$, $\Omega =]0, \infty[$, $u(0) = 0, u'(0) = 1$. The solution is the function u defined as follows:

On $[\alpha_0, \alpha_1]$ $(\alpha_0 = 0)$:

$$u(x) = -\frac{I}{24}x^4 + \frac{25}{12}I^4x^3 - \frac{5}{2}I^2x^2 + x.$$

On $[\alpha_n, \alpha_{n+1}]$, $n \geqslant 1$,

$$u(x) = -\rho^{-4}u(\rho(x - \alpha_n) + \alpha_{n-1})$$

$$\alpha_{n+1} - \alpha_n = \rho^{-1}(\alpha_n - \alpha_{n-1}).$$

On $[\lambda, +\infty[$, $u(x) = 0$, where $\lambda = \lim_{n\to\infty} \alpha^n$.

Moreover

$$\rho = \frac{3\xi^2}{(\xi - 2)^2}, \qquad I = \left(\frac{9}{125}\frac{\xi^3}{\xi^3 + 1}\right)^{1/5}, \qquad \alpha_1 = \frac{6\xi}{5I^2(\xi + 1)}$$

where ξ is the unique solution of the equation

$$9\xi^2 = (\xi + 1)(\xi - 2)^3$$

in the interval $]2, \infty[$. We have

$$I = \int_0^\infty |u(x)| \, dx.$$

By direct use of the characterization (3.62)–(3.63) of the solution, we can verify by elementary calculus that the above function is in fact the solution of (3.34).

3.3. Problems of filtering theory (2)

Another problem in filtering theory proposed by L.D. Berkovitz and H. Pollard [3] [4] is the following:

$$\text{(3.64)} \qquad \operatorname{Inf} \frac{1}{2}\left[\left(\int_0^{\infty} |u(x)| \, dx\right)^2 + \int_0^{\infty} |Au(x)|^2 \, dx\right]$$

the infimum being taken among those functions u which are regular over $\bar{\Omega}, \Omega =]0,+\infty[$, such that $u \in L^1(\Omega)$, $u(0) = 0$ and $Au \in L^2(\Omega)$ where A is the integro-differential operator defined by

$$\text{(3.65)} \quad Au(x) = -\frac{du}{dx}(x) + \exp(-x) + \int_0^x \exp(-(x-t)) \frac{du}{dt}(t) \, dt \, ;$$

this is equivalent to

$$\text{(3.66)} \qquad Au = \Lambda_1 u + \rho,$$

with

$$\text{(3.67)} \qquad \Lambda_1 u = -u' + u' \star \rho,$$

where ρ denotes the function

$$\text{(3.68)} \qquad \rho(x) = \left| \begin{array}{ll} \exp(-x), & x > 0 \\ 0, & x < 0 \end{array} \right.$$

and $v \star w$ denotes the convolution product of v and w

$$\text{(3.69)} \qquad v \star w(x) = \int_{-\infty}^{+\infty} v(t) w(x-t) \, dt \, .^{(1)}$$

To formulate problem (3.64) a little more precisely, we shall need the following lemma

Lemma 3.4. *Let u be an absolutely continuous function in $\bar{\Omega}$, $u \in L^1(\Omega)$, $u' \in L^1_{\mathrm{loc}}(\bar{\Omega})$. Then $\Lambda_1 u \in L^1_{\mathrm{loc}}(\bar{\Omega})$ and*

(1) It is to be understood that the functions defined over $]0, \infty[$, are extended by 0 for $x < 0$ and in such a case

$$v \star w(x) = \int_0^x v(t) w(x-t) \, dt, \quad x > 0 \quad \text{and} \quad v \star w(x) = 0, \quad x < 0.$$

(3.70) $$u(x) = u(0) - \int_0^x \Lambda_1 u(t)\,\mathrm{d}t - \int_0^x (x - t)\Lambda_1 u(t)\,\mathrm{d}t.$$

Proof. If g is a function defined over Ω, we term $\tilde{g}$ the function which is equal to g over Ω and to 0 for $x \leqslant 0$. We have the standard result (*cf.* Schwartz [1]):

(3.71) $$\tilde{u}' = \tilde{u}' + u(0)\delta,$$

where δ is the Dirac δ-function at 0.

We now consider the expression

(3.72) $$\tilde{\Lambda}_1 u = -\tilde{u}' + \tilde{u}' \star \rho = (\rho - \delta) \star \tilde{u}'$$

which from (3.67) and (3.71) is equal to

$$\tilde{\Lambda}_1 u = \widetilde{\Lambda_1 u} - u(0)\delta + u(0)\rho.$$

The inverse of the convolution of $(\rho - \delta)$ is $-(\phi_0 + \delta)$, where ϕ_0 is the Heaviside function

$$\phi_0(t) = \begin{cases} 1 & t > 0. \\ 0 & t < 0. \end{cases}$$

Thus we deduce from (3.72) that

$$\tilde{u}' = -(\phi_0 + \delta) \star (\widetilde{\Lambda_1 u} - u(0)\delta + u(0)\rho)$$

which gives us after simplification

$$\tilde{u}' = -\widetilde{\Lambda_1 u} - \phi_0 \star \widetilde{\Lambda_1 u} + u(0)\delta;$$

hence, in Ω,

$$u' = -\Lambda_1 u - \phi_0 \star \Lambda_1 u,$$

$$u'(x) = -\Lambda_1 u(x) - \int_0^x \Lambda_1 u(t)\,\mathrm{d}t, \qquad x > 0\,{}^{(1)}.$$

Another integration with respect to x implies

$$u(x) = u(0) - \int_0^x \Lambda_1 u(t)\,\mathrm{d}t - \int_0^x (x - t)\Lambda_1 u(t)\,\mathrm{d}t,$$

which is in fact (3.70).

(1) We can show by a similar argument that if $u \in L^1(\Omega)$ and if the restriction of $\tilde{\Lambda}_1 u$ to Ω is in $L^1(\Omega)$, then u is absolutely continuous and we have the conclusions of the lemma with fewer hypotheses.

We note that, from (3.73), $\forall T > 0$:

$$\|u'\|_{L^1([0,T])} \leqslant c(T)\,\|\Lambda_1 u\|_{L^1([0,T])}, \tag{3.74}$$

where $c(T)$ = a constant depending on T. ∎

We now define the space V,

$$V = \{\, u \in L^1(\Omega),\ u' \in L^1_{\mathrm{loc}}(\bar{\Omega}),\ \Lambda_1 u \in L^2(\Omega) \,\}. \tag{3.75}$$

By Lemma 3.4, V is a Banach space for the norm

$$\|u\|_{L^1(\Omega)} + \|\Lambda_1 u\|_{L^2(\Omega)}.$$

For $u \in V$, we can speak of $u(0)$, and the precise formulation of the problem is then to take the *infimum in* (3.64) *among those functions* $u \in V$ *such that* $u(0) = 0$.

Duality for problem (3.64)

The space V has already been defined; we set $Y = L^1(\Omega) \times L^2(\Omega)$, $V^* =$ the dual of V, $Y^* = L^\infty(\Omega) \times L^2(\Omega)$, and

$$\left|\begin{aligned} &\Lambda u = \{\, \Lambda_0 u, \Lambda_1 u \,\}, \\ &\Lambda_1 u = -u' + u' \star \rho, \qquad \Lambda_0 u = u. \end{aligned}\right. \tag{3.76}$$

For $u \in V$, let

$$F(u) = \left|\begin{aligned} &0 \quad \text{if } u(0) = 0, \\ &+\infty \quad \text{otherwise,} \end{aligned}\right. \tag{3.77}$$

and for $p = (p_0, p_1) \in Y$,

$$G(p) = G_0(p_0) + G_1(p_1) \tag{3.78}$$

$$G_0(p_0) = \tfrac{1}{2}\|p_0\|^2_{L^1(\Omega)}\,, \qquad G_1(p_1) = \tfrac{1}{2}\|p_1 + \rho\|^2_{L^2(\Omega)}. \tag{3.79}$$

The problem

$$\operatorname*{Inf}_{u \in V}\,[F(u) + G(\Lambda u)] \tag{3.80}$$

is indeed identical to problem (3.64).

We shall determine the conjugate functionals explicitly. From Remark III.4.2 and Proposition I.4.2,

$$G^*(p^*) = G_0^*(p_0^*) + G_1^*(p_1^*) \tag{3.81}$$

$$G_0^*(p_0^*) = \tfrac{1}{2}\,\|p_0^*\|^2_{L^\infty(\Omega)}, \tag{3.82}$$

$$G_1^*(p_1^*) = \tfrac{1}{2}\,\|p_1^*\|^2_{L^2(\Omega)} - \langle p_1^*, \rho \rangle. \tag{3.83}$$

Lemma 3.5. *Writing* $\check{\rho}(x) = \rho(-x)$, *we have*

$$F^*(\Lambda^* p^*) = 0 \tag{3.84}$$

if

$$\frac{\mathrm{d}}{\mathrm{d}x}(p_1^* - p_1^* \star \check{\rho}) + p_0^* = 0 \tag{3.85}$$

and $+\infty$ *otherwise.*

Proof. For $p^* \in Y^*$,

$$F^*(\Lambda^* p^*) = \sup_{\substack{u \in V \\ u(0)=0}} \langle p^*, \Lambda u \rangle$$

$$= \sup_{\substack{u \in V \\ u(0)=0}} \int_0^\infty [p_0^* u + p_1^*(-u' + u' \star \rho)]\,\mathrm{d}x.$$

But

$$\int_0^\infty p_1^*(u' \star \rho)\,\mathrm{d}x = \int_0^\infty (p_1^* \star \check{\rho}) u'\,\mathrm{d}x,$$

and we thus have

$$F^*(\Lambda^* p^*) = \sup_{\substack{u \in V \\ u(0)=0}} \int_0^\infty [p_0^* u + (p_1 \star \check{\rho} - p_1^*) u']\,\mathrm{d}x$$

$$\geqslant \sup_{u \in \mathscr{D}(\Omega)} \int_0^\infty [p_0^* u + (p_1^* \star \check{\rho} - p_1^*) u']\,\mathrm{d}x.$$

The latter supremum is equal to $+\infty$ if (3.85) is not satisfied and so $F^*(\Lambda^* p^*) = +\infty$ in this case. If (3.85) is satisfied, a legitimate integration by parts gives

$$\int_0^\infty [p_0^* u + (p_1^* \star \check{\rho} - p_1^*) u']\,\mathrm{d}x =$$

$$+ \int_0^\infty \left[p_0^* + \frac{\mathrm{d}}{\mathrm{d}x}(p_1^* - p_1^* \star \check{\rho})\right] u\,\mathrm{d}x = 0$$

and in this case we indeed have $F^*(\Lambda^* p^*) = 0$. ■

We are now in a position to state explicitly the dual problem

$$\text{(3.86)} \qquad \sup_{p^* \in Y^*} [-F^*(\Lambda^* p^*) - G^*(-p^*)].$$

It can be written as

$$\text{(3.87)} \qquad \sup_{\substack{p^* \in L^\infty(\Omega) \times L^2(\Omega) \\ p_0^* + (p_1^* - p_1^* \star \rho)' = 0}} \left[-\langle p_1^*, \rho \rangle - \frac{1}{2} \|p_0^*\|^2_{L^\infty(\Omega)} - \frac{1}{2} \|p_1^*\|^2_{L^2(\Omega)} \right].$$

Proposition 3.5. (i) *Problems* (3.64) *and* (3.87) *are mutually dual and*

$$\text{(3.88)} \qquad \operatorname{Inf} \mathscr{P} = \operatorname{Sup} \mathscr{P}^*.$$

(ii) *Problem* (3.64) *possesses a unique solution* $\bar{u}$, *problem* (3.87) *possesses a unique solution* $\bar{p}^*$ *and these two solutions are linked by the following extremality relations*:

$$\text{(3.89)} \qquad \bar{p}_0^*(x) = -\|\bar{u}\|_{L^1(\Omega)} \operatorname{sgn} \bar{u}(x) \quad \text{a.e. when} \quad \bar{u}(x) \neq 0,$$

$$\text{(3.90)} \qquad \|\bar{p}_0^*\|_{L^\infty(\Omega)} = \|\bar{u}\|_{L^1(\Omega)},$$

$$\text{(3.91)} \qquad \bar{p}_1^* = \bar{u}' - \bar{u}' \star \rho - \rho.$$

(iii) *A function u is the solution of* (3.64) *if and only if*

$$\text{(3.92)} \qquad u \in V, \qquad u(0) = 0,$$

$$\text{(3.93)} \qquad (q - q \star \check{\rho})' \in L^\infty(\Omega), \textit{ where } q = +u' - u' \star \rho - \rho$$

$$\text{(3.94)} \qquad (q - q \star \check{\rho})'(x) = +\|u\|_{L^1(\Omega)} \operatorname{sgn} u(x) \quad \textit{a.e. for} \quad u(x) \neq 0.$$

Proof of (i) *and* (ii). The relation (3.88) results from Theorem III.4.1 since condition III(4.8) is satisfied. The existence of a solution $\bar{p}^*$ of (3.87) also results from this theorem; this solution is unique since if $\bar{q}^*$ is another solution then $\bar{p}_1^* = \bar{q}_1^*$ as $p_1^* \mapsto -\frac{1}{2}\|p_1^*\|^2_{L^2(\Omega)}$ is strictly concave; and after

$$\bar{p}_0^* = (\bar{p}_1^* \star \rho - \bar{p}_1^*)' = (\bar{q}_1^* \star \rho - \bar{q}_1^*)' = \bar{q}_0^*$$

so that $\bar{p}^* = q^*$.

To prove the existence of a solution of (3.64) we use the direct method: if u_n is a minimizing sequence, then u_n is bounded in $L^1(\Omega)$, $\Lambda_1 u_n$ is bounded in $L^2(\Omega)$ and from (3.70) and (3.73) (and $u_n(0) = 0$), u_n is bounded in $L^2_{\text{loc}}(\bar{\Omega})$

and u_n is bounded in $L^\infty_{\mathrm{loc}}(\bar{\Omega})$. By compactness, there exists a sequence $n_j \to \infty$ and there exists $u \in L^\infty_{\mathrm{loc}}(\Omega)$ such that

$$
\begin{aligned}
&u_{n_j} \quad \to u \text{ in } L^\infty([0, T]) \text{ in the weak star-topology}, \qquad \forall T > 0,\\
&u'_{n_j} \quad \to u' \text{ weakly in } L^2([0, T]) \qquad \forall T > 0,\\
&\Lambda_1 u_{n_j} \to \Lambda_1 u \text{ weakly in } L^2(\Omega).
\end{aligned}
$$

We deduce from this that $u_{n_j} \to u$ uniformly on $[0, T]$, $\forall T > 0$, so that $u(0) = 0$; furthermore, application of Fatou's Lemma shows that $u \in L^1(\Omega)$ and

$$\int_0^\infty |u(x)|\,\mathrm{d}x \leqslant \liminf_{n_j \to \infty} \int_0^\infty |u_{n_j}(x)|\,\mathrm{d}x < +\infty.$$

Thus $u \in V$ and u is the solution of problem (3.64). The uniqueness of the solution results from the strict convexity of $\frac{1}{2}\|\ .\ \|^2_{L^2(\Omega)}$: if $\bar{u}$ and $\bar{v}$ are two different solutions then $\Lambda_1 \bar{u} = \Lambda_1 \bar{v}$ and from (3.70), $\bar{u} = \bar{v}$ since $\bar{u}(0) = \bar{v}(0) = 0$.

The extremality relations (*cf.* Proposition III.4.1 and Remarks III.4.1 and III.4.2) can be written

$$F(\bar{u}) + F^*(\Lambda^* \bar{p}^*) = \langle \Lambda^* \bar{p}^*, \bar{u} \rangle,$$

which is obvious,

$$G_0(\Lambda_0 \bar{u}) + G_0^*(-\bar{p}_0^*) = -\langle \bar{p}_0^*, \Lambda_0 \bar{u} \rangle, \tag{3.95}$$

$$G_1(\Lambda_1 \bar{u}) + G_1^*(-\bar{p}_1^*) = -\langle \bar{p}_1^*, \Lambda_1 \bar{u} \rangle. \tag{3.96}$$

Relation (3.95) can be written

$$\tfrac{1}{2}\|\bar{u}\|^2_{L^1(\Omega)} + \tfrac{1}{2}\|\bar{p}_0^*\|^2_{L^\infty(\Omega)} = -\langle \bar{p}_0, \bar{u} \rangle,$$

and is equivalent to (3.89) (3.90) (*cf.* no. 3.2). Relation (3.96) can be written:

$$
\begin{aligned}
&\tfrac{1}{2}\|-\bar{u}' + \bar{u}' \star \rho + \rho\|^2_{L^2(\Omega)} + \tfrac{1}{2}\|\bar{p}_1^*\|^2_{L^2(\Omega)} - \langle \bar{p}_1^*, \rho \rangle =\\
&= -\langle \bar{p}_1^*, \bar{u}' - \bar{u}' \star \rho \rangle
\end{aligned}
$$

and is equivalent to (3.91).

(iii) Let us assume that a function u satisfies (3.92)–(3.94). We then set

$$
\begin{aligned}
p_1^* &= q = u' - u' \star \rho - \rho \in L^2(\Omega)\ ^{(1)},\\
p_0^* &= (q \star \check{\rho} - q) \in L^\infty(\Omega).
\end{aligned}
$$

We verify that $u \in V$ and $p^* \in Y^*$ satisfy the extremality relations and u is thus the solution of (3.64) and $\bar{p}^*$ the solution of (3.87). ■

(1) u and $\rho \in L^2(\Omega)$; $u' \star \rho \in L^2(\Omega)$ by the convolution inequality since $\rho \in L^\infty(\Omega)$.

Remark 3.6. The solution of (3.64) has been determined explicitly in Berkovitz and Pollard [3] [4]. It is the function

$$u(t) = -\frac{I}{24}(t-\lambda)^2 t(r-t) \quad \text{if} \quad t<\lambda \tag{3.97}$$

$$= 0 \quad \text{if} \quad t \geqslant \lambda,$$

where $I = \|u\|_{L^1(\Omega)}$, whereas r, λ, I are linked to the auxiliary coefficient ν by the algebraic relations

$$\begin{aligned} I\lambda^2 r &= 24\nu \\ I(\lambda^2 + 2r\lambda + 12) &= 6(1-\nu^2) \\ I(r+2\lambda) &= 2(1-\lambda)^2 \\ \frac{24}{\lambda^4} + \frac{\lambda}{30} &= \frac{r}{12}. \end{aligned}$$

By direct use of the characterization (3.92)–(3.94) of the solution, we can verify by extremely elementary calculus that (3.97) is indeed the solution of (3.64).

3.4. The problem of elasto-plastic torsion

Let us consider in $H_0^1(\Omega)$, the set

$$\mathscr{K} = \{ v \in H_0^1(\Omega) \mid |\text{grad } v(x)| \leqslant 1 \text{ p.p.} \}$$

which can easily be proved to be closed and convex. The problem of elasto-plastic torsion is the problem:

$$\operatorname*{Inf}_{v \in \mathscr{K}} \left\{ \frac{1}{2} \int_\Omega [(\text{grad } u)^2 - 2fu] \, \mathrm{d}x \right\} \tag{3.98}$$

which is equivalent to finding the solution $u \in K$ of the variational inequality

$$((u, v-u)) - (f, v-u) \geqslant 0, \qquad \forall v \in K. \tag{3.99}$$

According to a result of Brezis and Stampacchia [1], if f is given in $L^\alpha(\Omega)$, $1 < \alpha < \infty$, then (3.99) possesses a unique solution $\bar{u}$ satisfying $\bar{u} \in W^{2,\alpha}(\Omega)$. To simplify matters, we shall assume that

$$f \in L^\infty(\Omega), \tag{3.100}$$

so that

$$\bar{u} \in W^{2,\alpha}(\Omega), \qquad \forall \alpha, \ 1 \leqslant \alpha < \infty. \tag{3.101}$$

Because of the Sobolev embedding theorems[1] $\bar{u}$ satisfies

$$\bar{u} \in \mathscr{C}^1(\bar{\Omega}). \tag{3.102}$$

We set $V = H_0^1(\Omega)$, $V^* = H^{-1}(\Omega)$, $Y = Y^* = L^2(\Omega)^n$, $\Lambda = \text{grad}$ and

$$F(u) = \tfrac{1}{2}\|u\|^2 - \langle f, u \rangle$$

$$G(p) = \begin{cases} 0 & \text{if} \quad |p(x)| \leqslant 1 \quad \text{a.e.} \\ +\infty & \text{otherwise} \end{cases}$$

and problem (3.98) can be written:

$$\operatorname*{Inf}_{u \in V} [F(u) + G(\Lambda u)]. \tag{3.103}$$

We easily check that

$$F^*(u^*) = \tfrac{1}{2}\|u^* + f\|_*^2,$$

where $\| \, . \, \|$ is the dual norm of

$$\|u\| = \left[\int_\Omega (\text{grad } u)^2 \, \mathrm{d}x\right]^{1/2}$$

in $H^{-1}(\Omega)$. We also have

$$G^*(p^*) = \int_\Omega |p^*(x)| \, \mathrm{d}x.$$

The dual problem of (3.103) can therefore be written as

$$\operatorname*{Sup}_{p^* \in L^2(\Omega)^n} \left[-\frac{1}{2}\|\text{div } p^* - f\|_*^2 - \int_\Omega |p^*(x)| \, \mathrm{d}x \right]. \tag{3.104}$$

Problem (3.103) does not satisfy the stability criterion (4.21) given in Chapter III. On the other hand, the dual problem (3.104) satisfies the analogous condition to III(4.21) and owing to the complete symmetry of problems $\mathscr{P}$ and $\mathscr{P}^*$, we can easily transcribe the results contained in Theorems III.4.1 and III.4.2: $\mathscr{P}^*$ is stable, which implies that $\mathscr{P}$ possesses a solution (which we already knew), and that

$$\text{Inf } \mathscr{P} = \text{Sup } \mathscr{P}^*. \tag{3.105}$$

[1] $W^{m,\alpha}(\Omega) \subset \mathscr{C}^0(\bar{\Omega})$ with a continuous embedding provided that $n < m\alpha$ and that Ω is sufficiently regular (Ω is $\mathscr{C}^r$, $r \geqslant m + 2$). *Cf.* Sobolev [1], Lions [2].

If $\mathscr{P}^*$ possesses a solution $\bar{p}^*$ then the extremality relations imply that

$$(3.106) \qquad -\Delta\bar{u} + \operatorname{div}\bar{p}^* = f$$

and

$$(3.107) \qquad \int_\Omega |\bar{p}^*(x)|\,\mathrm{d}x = -\int_\Omega \bar{p}^*(x).\,\Lambda\bar{u}(x)\,\mathrm{d}x,$$

which is equivalent to

$$(3.108) \qquad |\bar{p}^*(x)| = -\bar{p}^*(x)\cdot\Lambda\bar{u}(x) \qquad \text{a.e.} \quad x\in\Omega,$$

which yields almost everywhere in Ω,

$$(3.109) \qquad \left|\begin{array}{l} |\bar{p}^*(x)| = 0 \quad \text{if } |\operatorname{grad}\ \bar{u}(x)| < 1 \\ \bar{p}^*(x) = -\lambda(x)\dfrac{\operatorname{grad}\bar{u}(x)}{|\operatorname{grad}\bar{u}(x)|} \quad \text{if } |\operatorname{grad}\bar{u}(x)| = 1 \end{array}\right.$$

where

$$(3.110) \qquad \lambda(x) = |\bar{p}^*(x)|.$$

Then $\bar{u}$ satisfies

$$(3.111) \qquad -\Delta\bar{u} - \operatorname{div}\left(\lambda\frac{\operatorname{grad}\bar{u}}{|\operatorname{grad}\bar{u}|}\right) = f.$$

Conversely, if there exists $\lambda\in L^2(\Omega)$ such that $\lambda(x)=0$ for $|\operatorname{grad}\bar{u}(x)| < 1$, and which satisfies (3.111), we define $\bar{p}^*$ from (3.109) and the extremality relations (3.106) and (3.107) are satisfied, which means that $\bar{p}^*$ is the solution of $\mathscr{P}^*$ (and $\bar{u}$ the solution of $\mathscr{P}$). ■

Remark 3.7. As problem (3.104) is not coercive in $L^2(\Omega)^n$,[1] it does not necessarily possess a solution.

Remark 3.8. If problem $\mathscr{P}^*$ possesses a solution, then we necessarily have (3.106). For this reason it is sufficient to maximize (3.104) among the $p^*\in L^2(\Gamma)^n$ which satisfy (3.106). Setting

$$\xi = f + \Delta u$$

and changing the signs, we obtain

$$(3.112) \qquad \operatorname*{Inf}_{\substack{p^*\in L^2(\Omega)^n \\ \operatorname{div}p^* = \xi}} \left\{\int_\Omega |p^*(x)|\,\mathrm{d}x\right\}.$$

[1] It is only coercive in $L^1(\Omega)^n$, which is not a reflexive space.

In the two-dimensional case, $n=2$, there exists $\phi \in H_0^1(\Omega)$ such that $\Delta\phi = \xi$, and setting $q = p^* - \mathrm{grad}\,\phi$, we have $\mathrm{div}\, q = 0$. If furthermore Ω is simply connected, there exists $\psi \in H^1(\Omega)$ such that,

$$q_1 = \partial\psi/\partial x_2, \qquad q_2 = -\,\partial\psi/\partial x_1, \tag{3.113}$$

and problem (3.112) is therefore equivalent to the following problem for ψ,

$$\operatorname*{Inf}_{\psi \in H^1(\Omega)} \int_\Omega \left[\left(\frac{\partial\psi}{\partial x_2} + \frac{\partial\phi}{\partial x_1}\right)^2 + \left(\frac{\partial\psi}{\partial x_1} - \frac{\partial\phi}{\partial x_2}\right)^2\right]^{1/2} \mathrm{d}x. \tag{3.114}$$

Like (3.104), this problem is not coercive (or at least it is coercive in $W^{1,1}(\Omega)$). It can be compared with the problems which will be studied in Chapter V, no. 4.1. ∎

We shall now see that it is possible to define a generalized solution of (3.104) according to the method outlined in Chapter III, no. 6.1.

We set $\tilde{V} = H_0^1(\Omega) \cap \mathscr{C}^1(\bar{\Omega})$ resp. ($\tilde{Y} = \mathscr{C}^0(\bar{\Omega})^n$) together with the topology induced by $H_0^1(\Omega)$ (resp. $L^2(\Omega)^n$). We set $Z =$ the space of bounded Radon measures on Ω with values in $\mathbf{R}^n$, $Z = \mathscr{M}_1(\Omega)^n$. The natural scalar product between $\tilde{Y}$ and Z extends the natural scalar product between $\tilde{Y}$ and Y^*, and the spaces $\tilde{Y}$ and Z (resp. $\tilde{Y}$ and Y^*) are mutually dual for the corresponding topologies $\sigma(\tilde{Y}, Z)$, $\sigma(Z, \tilde{Y})$ (resp. $\sigma(\tilde{Y}, Y^*)$, $\sigma(Y^*, \tilde{Y})$).

The extension of G^* to Z is easy:

$$\tilde{G}^*(p^*) = \int_\Omega |p^*| = \text{variation of } |p^*| \text{ on } \Omega, \quad \forall p^* \in Z,$$

and we have

$$F^*(\Lambda^* p^*) = \begin{cases} \frac{1}{2}\|\mathrm{div}\, p^* - f\|_*^2 & \text{if } \mathrm{div}\, p^* \in H^{-1}(\Omega) \\ +\infty & \text{otherwise.} \end{cases}$$

After that we consider the problem $\tilde{\mathscr{P}}$,

$$\operatorname*{Inf}_{v \in \tilde{V}} [F(u) + G(\Lambda u)]$$

which admits the same infimum as (3.103) and the same solution $\bar{u}$ because of (3.102). We also consider the problem $\tilde{\mathscr{P}}^*$, which is the bidual of $\mathscr{P}^*$:

$$\operatorname*{Sup}_{\substack{p^* \in Z \\ \mathrm{div}\, p^* \in H^{-1}(\Omega)}} \tfrac{1}{2}\left\{-\ \|\mathrm{div}\, p^* - f\|_*^2 - \int_\Omega |p^*|\right\}. \tag{3.115}$$

By analogous reasoning to that in Section III.6.1,

$$\operatorname{Sup} \mathscr{P}^* \leqslant \operatorname{Sup} \tilde{\mathscr{P}}^* \leqslant \operatorname{Inf} \tilde{\mathscr{P}},$$

and since $\operatorname{Sup}\mathscr{P}^* = \operatorname{Inf}\tilde{\mathscr{P}}$, there results the equality

$$\operatorname{Sup} \mathscr{P}^* = \operatorname{Sup} \tilde{\mathscr{P}}^*.$$

4. REGULARITY AND DUALITY

Orientation

The pairs of dual spaces considered in the general setting may be Hilbertian, and in this case we can take $V = V^*$ (resp. $Y = Y^*$) the pairing between V and V^* (resp. Y and Y^*) being the scalar product of V (resp. Y). This special case corresponds to Hilbertian duality.

Here we give an example of Hilbertian duality. This example will also illustrate the interconnection existing between the regularities of the solutions of the two dual problems. ■

The primal problem is:

$$\inf_{u \in H_0^1(\Omega)} \left[\tfrac{1}{2}\|u\|^2 - (f, u) + \psi(u) \right] \tag{4.1}$$

where $f \in L^2(\Omega)$ is given, $\psi \in \Gamma_0\,(H_0^1(\Omega))$ is given (convex, l.s.c. and proper), $\| \,.\, \|$ being the norm in $H_0^1(\Omega)$; the open set Ω is assumed to be of class $\mathscr{C}^2$.

We set $V = Y = H_0^1(\Omega)$, $V^* = Y^* = H_0^1(\Omega)$, the duality between V and V^* (resp. Y and Y^*) being the scalar product of V (resp. Y). The operator Λ is the identity,

$$F(u) = \tfrac{1}{2}\|u\|^2 - (f, u), \tag{4.2}$$

$$G(p) = \psi(u) \qquad (p = u). \tag{4.3}$$

For $u^* \in V^* = H_0^1(\Omega)$, we have

$$F^*(u^*) = \sup_{u \in H_0^1(\Omega)} \left[((u^*, u)) + (f, u) - \tfrac{1}{2}\|u\|^2 \right].$$

The supremum is attained at the point u, which is the solution of the Dirichlet problem

$$\left| \begin{aligned} & -\Delta u = f - \Delta u^*, \\ & u \in H_0^1(\Omega), \end{aligned} \right. \tag{4.4}$$

and the value of the supremum is

$$F^*(u^*) = +\tfrac{1}{2}\|u^* + \phi\|^2, \tag{4.5}$$

where ϕ is the solution of

$$\left| \begin{array}{l} -\Delta\phi = f, \\ \phi \in H_0^1(\Omega). \end{array} \right. \tag{4.6}$$

The conjugate function of G is

$$\psi^*(p^*) = \operatorname*{Sup}_{p \in H_0^1(\Omega)} [((p^*, p)) - \psi(p)]. \tag{4.7}$$

The dual problem of (4.1) has essentially the same form as (4.1):

$$\operatorname*{Sup}_{w \in H_0^1(\Omega)} [-\tfrac{1}{2}\|w + \phi\|^2 - \psi^*(w)] \tag{4.8}$$

(we use the variable w instead of p^*).

Theorem III.4.2 applies: hypothesis III(4.8) is easily verified and for coerciveness, we note that ψ possesses at least one continuous affine minorant function

$$\psi(u) \geqslant ((a, u)) + \lambda, \qquad a \in H_0^1(\Omega),\ \lambda \in \mathbf{R},$$

and hence

$$\begin{aligned} \tfrac{1}{2}\|u\|^2 - (f, u) + \psi(u) &\geqslant \tfrac{1}{2}\|u\|^2 - ((\phi, u)) - ((a, u)) - \lambda \\ &- \tfrac{1}{2}\|u - (\phi + a)\|^2 \geqslant \tfrac{1}{2}\|\phi + a\|^2 - \lambda, \end{aligned}$$

and III(4.8) follows.

Hence problem (4.1) possesses a unique solution $\bar{u}$, problem (4.8) a unique solution $\bar{w}$,

$$\min \mathscr{P} = \max \mathscr{P}^*,$$

and $\bar{u}$ and $\bar{w}$ are linked by the extremality relations

$$F(\bar{u}) + F^*(\bar{w}) = \langle \bar{w}, \bar{u} \rangle \tag{4.9}$$

$$\psi(\bar{u}) + \psi^*(\bar{w}) = -\langle \bar{w}, \bar{u} \rangle. \tag{4.10}$$

From the calculation of F^* (*cf.* (4.4)) we see that (4.9) amounts to saying

$$-\Delta\bar{u} = f - \Delta\bar{w}$$

or

$$\bar{u} = \phi + \bar{w}. \tag{4.11}$$

The exploitation of (4.10) depends on the explicit form of ψ.

A special case

Let $\mathscr{C} = \{u \in H_0^1(\Omega) \| u(x)| \leqslant 1 \text{ a.e.}\}$. This set is convex and closed in $H^6\Omega)$, and its indicator function $\chi_{\mathscr{C}}$ is convex, l.s.c. and proper. With $\psi = \chi_{\mathscr{C}}$, problem (4.1) can be written as

$$\text{(4.12)} \qquad \operatorname*{Inf}_{\substack{u \in H_0^1(\Omega) \\ |u(x)| \leqslant 1 \text{ a.e.}}} [\tfrac{1}{2}\|u\|^2 - (f, u)].$$

For the determination of ψ^* we have,

Lemma 4.1. *$\psi^*(w) = \int |\Delta w|$ if Δw is a bounded measure and $+\infty$ otherwise.*

Proof. Let us assume for the moment (*cf.* Lemma 4.2) that $\mathscr{D}(\Omega) \cap \mathscr{C}$ is dense in $\mathscr{C}$. Then we have

$$\text{(4.13)} \qquad \begin{aligned} \psi^*(w) &= \operatorname*{Sup}_{\substack{\theta \in \mathscr{D}(\Omega) \\ |\theta(x)| \leqslant 1}} ((w, \theta)) \\ &= \operatorname*{Sup}_{\substack{\theta \in \mathscr{D}(\Omega) \\ |\theta(x)| \leqslant 1}} \int_\Omega (-\Delta w) \cdot \theta \, dx \end{aligned}$$

and the required result follows.

Lemma 4.2. *$\mathscr{D}(\Omega) \cap \mathscr{C}$ is dense in $\mathscr{C}$.*

Proof. Let θ_ε ($\varepsilon > 0$) be a family of functions of $\mathscr{C}^2(\bar{\Omega})$, $0 \leqslant \theta_\varepsilon \leqslant 1$, and

$$\begin{aligned} \theta_\varepsilon(x) &= 0 \quad \text{if} \quad d(x, \Gamma) \leqslant \varepsilon \quad (\text{distance of } x \text{ to } \Gamma) \\ &= 1 \quad \text{if} \quad d(x, \Gamma) \geqslant 2\varepsilon. \end{aligned}$$

For every $u \in H_0^1(\Omega)$, $\theta_\varepsilon u \to u$ in $H_0^1(\Omega)$, when $\varepsilon \to 0$ (*cf.* Lions and Magenes [1]). In particular if $u \in \mathscr{C}$, $\theta_\varepsilon u \to u$, but we note that $\theta_\varepsilon u \in \mathscr{C}$, $\forall \varepsilon > 0$.

It thus remains for us to approximate a function $u \in \mathscr{C}$ with a compact support in Ω, by functions of $\mathscr{D}(\Omega) \cap \mathscr{C}$. The result follows immediately by regularization. ∎

Problem (4.8) becomes

$$\text{(4.14)} \qquad \operatorname*{Sup}_{\substack{w \in H_0^1(\Omega) \\ \Delta w \text{ bounded} \\ \text{measure}}} [-\tfrac{1}{2}\|w + \phi\|^2 - \int |\Delta w|].$$

This problem possesses a unique solution $\bar{w}$.

Remark 4.1. If Ω is regular, then from the standard regularity results $\phi \in H^2(\Omega)$ and from Brezis and Stampacchia [1], $\bar{u} \in H^2(\Omega)$. By virtue of (4.11):

$$\bar{w} \in H_0^1(\Omega) \cap H^2(\Omega) \tag{4.15}$$

which implies that the supremum in (4.14) is attained on $H_0^1(\Omega) \cap H^2(\Omega)$. But for $w \in H^2(\Omega)$, Δw is obviously a bounded measure and

$$\int |\Delta w| = \int_\Omega |\Delta w(x)| \, dx,$$

so that (4.14) becomes

$$\operatorname*{Sup}_{w \in H_0^1(\Omega) \cap H^2(\Omega)} \left[- \tfrac{1}{2} \|w + \phi\|^2 - \int_\Omega |\Delta w| \, dx \right] \tag{4.16}$$

and $\bar{w}$ is solution of (4.16).

This result is due in a different form, to H. Brezis [2].

5. GENERAL PROBLEMS IN THE CALCULUS OF VARIATIONS

Orientation

In this paragraph, we wish to study a form of duality for a general problem of the calculus of variations of the type

$$\operatorname*{Inf}_{u \in \mathscr{C}} \int_\Omega g(x, u(x), P(D)u(x)) \, dx, \tag{5.0}$$

where $\mathscr{C}$ is a convex set of functions, g a function which is convex with respect to its last two arguments and $P(D)$ a differential operator. ■

Let g be a Carathéodory function defined on $\Omega \times \mathbf{R}^l \times \mathbf{R}^m$ with values in $\mathbf{R}$, i.e.:

(5.1) for almost all $x \in \Omega$, $(u, \xi) \mapsto g(x, u, \xi)$ is continuous on $\mathbf{R}^l \times \mathbf{R}^m$,

(5.2) for all $(u, \xi) \in \mathbf{R}^l \times \mathbf{R}^m$, $x \to g(x, u, \xi)$, is measurable on Ω.

Additionally, we assume that for almost all $x \in \Omega$

(5.3) $\eta \to g(x, \eta)$ is convex from $\mathbf{R}^l \times \mathbf{R}^m$ into $\mathbf{R}$.

Setting

$$\begin{aligned} W &= L^{\alpha_1}(\Omega) \times \dots \times L^{\alpha_l}(\Omega) \\ Z &= L^{\beta_1}(\Omega) \times \dots \times L^{\beta_m}(\Omega), \end{aligned}$$

where $1 \leqslant \alpha_i < +\infty$, $1 \leqslant \beta_j < +\infty$ for all i and j, we assume that for all $u \in W$ and for all $\theta \in Z$, the measurable function

$$x \mapsto g(x, u(x), \theta(x)) \tag{5.4}$$

is in $L^1(\Omega)$.

It thus follows from the theorem of Krasnoselskii [1][(1)] that the mapping

$$(u, \theta) \to \{ x \mapsto g(x, u(x), \theta(x) \} \tag{5.5}$$

is a continuous function of $W \times Z$ in $L^1(\Omega)$.

We term $P(D)$ a differential operator continuous in the distribution space and we consider the space V,

$$V = \{ u \in W \mid P(D)u \in Z \},$$

which is a Banach space for the norm

$$\|u\|_V + \|P(D)u\|_Z.$$

Finally $\mathscr{C}$, the set of constraints, is a non-empty closed convex subset of V.

Problem (5.0) is now completely defined.

We shall now study its dual problem. Returning to the position in Chapter III, no. 4, the space V being already defined, we set

$$Y = W \times Z, \qquad V^* = \text{the dual of } V, \qquad Y^* = \text{the dual of } Y,$$
$$\Lambda u = (u, P(D)u), \qquad \forall u \in V \qquad (\text{naturally } \Lambda \in \mathscr{L}(V, Y)),$$
$$F = \chi_{\mathscr{C}} = \text{the indicator function of } \mathscr{C},$$

$$G(p) = \int_\Omega g(x, p(x))\,\mathrm{d}x, \qquad \forall p \in W \times Z.$$

The problem

$$\operatorname*{Inf}_{u \in V} \{ F(u) + G(\Lambda u) \} \tag{5.6}$$

is identical to (5.0)

F^* may be calculated easily,

$$F^* = \chi_{\mathscr{C}}^* = \text{the support function of } \mathscr{C}.$$

To calculate G^* we use Proposition 1.2 and we find that

$$G^*(p^*) = \int_\Omega g^*(x, p^*(x))\,\mathrm{d}x, \tag{5.7}$$

(1) Proposition 1.1.

where

$$g^*(x, \eta^*) = \operatorname*{Sup}_{\eta \in \mathbb{R}^l \times \mathbb{R}^m} [\eta^* \eta - g(x, \eta)], \quad \text{a.e. } x \in \Omega. \tag{5.8}$$

The dual problem of (5.0) can then be written

$$\operatorname*{Sup}_{p^* \in Y^*} \left\{ - \chi^*_{\mathscr{C}}(\Lambda^* p^*) - \int_\Omega g^*(x, p^*(x))\, \mathrm{d}x \right\}. \tag{5.9}$$

It is interesting to note that because of (5.5), Condition III(4.21) is satisfied and so with Theorems III.4.1 and III.4.2, and Proposition III.4.1, we have:

Proposition 5.1. *Problem* (5.9) *possesses at least one solution* $\bar{p}^*$ *and*

$$\inf (5.0) = \sup (5.9). \tag{5.10}$$

If problem (5.0) *possesses a solution* $\bar{u}$, *then we have the extremality relations*:

$$\chi_{\mathscr{C}}(\bar{u}) + \chi^*_{\mathscr{C}}(\Lambda^* \bar{p}^*) = \langle \Lambda^* \bar{p}^*, \bar{u} \rangle, \tag{5.11}$$

$$g(x, \Lambda \bar{u}(x)) + g^*(x, -\bar{p}^*(x)) = -\bar{p}^*(x). \Lambda \bar{u}(x) \qquad \text{a.e. } x \in \Omega. \tag{5.12}$$

Remark 5.1. It is noteworthy that we can demonstrate the existence of a solution for $\mathscr{P}^*$ with what little information we possess concerning g.

We may wonder what significance p^* has for $\mathscr{P}$ when $\mathscr{P}$ has no solution: *by means of* $\bar{p}^*$, *we can envisage the definition of a sort of generalized solution of* $\mathscr{P}$. This is exactly what we shall do in Chapter V for some very specialized cases of (5.0).

CHAPTER V

Applications of Duality to the Calculus of Variations (II) Minimal Hypersurface Problems

Orientation

In Section 1 we shall, as in the preceding chapter, make use of the techniques of duality of Chapter III for the non-parametric minimal hypersurfaces problem (with or without obstacles) and for various similar problems: the link between these problems is that they are only coercive in *non-reflexive* spaces [$L^1(\Omega)$ or a space built on $L^1(\Omega)$].

For each example, the position will be as follows: the primal problem may or may not possess a solution $\bar{u}$, the coerciveness being only available in a non-reflexive space. The dual problem, on the other hand, will possess a unique solution $\bar{p}^*$. The extremality relations will link $\bar{p}^*$ to the solution $\bar{u}$ of the primal problem, when this solution exists. When the primal problem possesses no solution, the dual problem will allow us to define a generalized solution to the problem: this is the aim of Sections 2 and 3, in Section 2 for the non-parametric minimal hypersurfaces problem and in Section 3 for a more general class of problems. Sections 2 and 3 will have recourse to numerous results in the theory of partial differential equations which will be recalled at the appropriate time. Moreover, a result of Brøndsted and Rockafellar concerning ε sub-differentials (Theorem I.6.2) plays a remarkable role here.

1. NON-PARAMETRIC MINIMAL HYPERSURFACES

1.1. The primal problem and the dual problem

We wish to minimize the integral

$$\int_\Omega \sqrt{1 + |\text{grad } u|^2}\, \mathrm{d}x, \tag{1.0}$$

among all the functions u which are equal to a given function ϕ on the boundary Γ of Ω. It is the problem of minimizing the area of a hypersurface among those which are graphs of a univocal function defined in Ω and are lying on the given contour (the points $(x,\phi(x)), x \in \Gamma$). This problem is equivalent to a Dirichlet problem for the equation of minimal hypersurfaces (the Euler equation of the problem):

$$(1.1)\qquad \left|\begin{array}{l} \displaystyle\sum_{i=1}^{n} D_i \frac{D_i u}{[1 + |\operatorname{grad} u|^2]^{1/2}} = 0 \quad \text{in} \quad \Omega. \\ u = \phi \quad \text{on } \Gamma. \end{array}\right.$$

To set the problem precisely, let us assume that Ω is a bounded set which is not necessarily regular. The space with which we shall be concerned is $W^{1,1}(\Omega)$ or occasionally(1) $\mathscr{W}^{1,1}(\Omega)$ = the closure in $W^{1,1}(\Omega)$ of indefinitely differentiable functions in $\bar{\Omega}$. The function ϕ is assumed to be given throughout Ω instead of just on Γ; this condition is only restrictive if Ω is not regular.(2) Thus we assume that ϕ is given satisfying:

$$(1.2)\qquad \phi \in \mathscr{W}^{1,1}(\Omega),$$

and we seek to minimize (1.0) among the functions $u \in \phi + W_0^{1,1}(\Omega)$ which exactly expresses, in a weak sense, the condition $u = \phi$ on Γ:

$$(1.3)\qquad \operatorname*{Inf}_{u \in \phi + W_0^{1,1}(\Omega)} \int_\Omega [1 + |\operatorname{grad} u(x)|^2]^{1/2}\,dx.$$

Just as we have done throughout Chapter IV, we revert to the situation in Chapter III(4.16) and we set

$$\begin{aligned}
V &= W^{1,1}(\Omega), \qquad Y = L^1(\Omega)^n, \\
V^* &= W^{1,1}(\Omega)^* = \text{the dual of } W^{1,1}(\Omega), \\
Y^* &= L^\infty(\Omega)^n, \qquad \Lambda = \operatorname{grad}, \\
F(v) &= \left|\begin{array}{l} 0 \quad \text{if} \quad v \in \phi + W_0^{1,1}(\Omega) \\ +\infty \text{ otherwise.} \end{array}\right. \\
G(p) &= \int_\Omega [1 + |p(x)|^2]^{1/2}\,dx, \qquad \forall p \in L^1(\Omega)^n.
\end{aligned}$$

We have thus reduced problem (1.3) to a problem III(4.16); we shall now describe the dual problem III(4.18).

(1) $\mathscr{W}^{1,1}(\Omega) = W^{1,1}(\Omega)$ if Ω is regular, *cf.* Lions [2].

(2) If Ω is sufficiently regular (for instance IV(1.2) with $r \geqslant 2$), then because of a result of Gagliardo [1], the trace operator γ_0 is defined and maps $W^{1,1}(\Omega)$ continuously and linearly into $L^1(\Gamma)$ and there is a continuous lifting operator $R \in \mathscr{L}(L^1(\Gamma), W^{1,1}(\Omega))$ ($\gamma_0 \circ R$ = the identity). Obviously, in this case, any given function $\theta \in L^1(\Gamma)$ can be "extended" inside Ω by a function $\phi \in W^{1,1}(\Omega)$ with $\gamma_0 \phi = \theta$.

Calculation of F^* and G^*

$$F^*(\Lambda^* p^*) = \sup_{v \in \phi + W_0^{1,1}(\Omega)} \langle p^*, \Lambda v \rangle$$
$$F^*(\Lambda^* p^*) = \langle p^*, \Lambda \phi \rangle + \sup_{v \in W_0^{1,1}(\Omega)} \langle p^*, \Lambda v \rangle.$$

Since $\mathscr{D}(\Omega)$ is dense in $W_0^{1,1}(\Omega)$,

$$\sup_{v \in W_0^{1,1}(\Omega)} \langle p^*, \Lambda v \rangle = \sup_{v \in \mathscr{D}(\Omega)} \langle p^*, \Lambda v \rangle = \sup_{v \in \mathscr{D}(\Omega)} [- \langle \operatorname{div} p^*, v \rangle]$$

and this supremum is 0 if $\operatorname{div} p^* = 0$, and $+\infty$ otherwise:

$$(1.4) \qquad F^*(\Lambda^* p^*) = \begin{cases} \langle p^*, \Lambda \phi \rangle = \displaystyle\int_\Omega p^* \operatorname{grad} \phi \, dx & \text{if } \operatorname{div} p^* = 0, \\ +\infty \text{ otherwise.} \end{cases}$$

Lemma 1.1.

$$G^*(p^*) = - \int_\Omega [1 - |p^*(x)|^2]^{1/2} \, dx, \quad \text{if } |p^*(x)| \leqslant 1 \quad \text{a.e.},$$

and $+\infty$ *otherwise.*

Proof.

$$(1.5) \qquad G^*(p^*) = \sup_{p \in L^1(\Omega)^n} \int_\Omega (p^*(x) \cdot p(x) - [1 + |p(x)|^2]^{1/2}) \, dx,$$

and to calculate G^*, by Proposition IV.1.2, it is sufficient to determine for almost all x

$$(1.6) \qquad \sup_{\xi \in \mathbb{R}^n} [p^*(x)\xi - (1 + |\xi|^2)^{1/2}].$$

If $\xi = \eta\rho$, $\rho = |\xi|$, the supremum in η for fixed ρ is attained for $\eta = p^*(x)/|p^*(x)|$ and its value is

$$(1.7) \qquad |p^*(x)| \rho - (1 + \rho^2)^{1/2};$$

it is then necessary to maximize (1.7) for $\rho \geqslant 0$. A simple calculation shows that the supremum is equal to $+\infty$ if $|p^*(x)| > 1$, to 0 (and this is achieved at $\rho = +\infty$) if $|p^*(x)| = 1$; finally if $|p^*(x)| < 1$, the supremum is equal to $-\sqrt{1 - |p^*(x)|^2}$ and this is attained at

$$(1.8) \qquad \rho = \frac{-p^{*(x)}}{[1 - |p^*(x)|^2]^{1/2}}.$$

The lemma then follows. ∎

Remark 1.1. It is important for what follows to observe that the relation

$$G(p) + G^*(p^*) = \langle p^*, p \rangle, \tag{1.9}$$

implies that $|p^*(x)| \leqslant 1$ almost everywhere and that for almost all $x \in \Omega$:

$$\begin{aligned} -[1 - |p^*(x)|^2]^{1/2} &= \operatorname*{Sup}_{\xi \in \mathbb{R}^n} [p^*(x) \cdot \xi - (1 + |\xi|^2)^{1/2}] \\ &= p^*(x) \cdot p(x) - [1 + |p(x)|^2]^{1/2}. \end{aligned}$$

From the proof of Lemma 1.1, this means that

$$|p^*(x)| < 1 \quad \text{a.e.,} \tag{1.10}$$

and

$$p(x) = \frac{p^*(x)}{[1 - |p^*(x)|^2]^{1/2}} \quad \text{a.e.,} \tag{1.11}$$

or

$$p^*(x) = \frac{p(x)}{[1 + |p(x)|^2]^{1/2}} \quad \text{a.e.} \quad \blacksquare \tag{1.12}$$

The dual problem

Because of (1.4) and Lemma 1.1, this can be written as

$$\operatorname*{Sup}_{\substack{p^* \in L^\infty(\Omega)^n \\ |p^*(x)| \leqslant 1 \text{ a.e.} \\ \operatorname{div} p^* = 0}} \left[- \int_\Omega p^*(x) \cdot \Lambda\phi(x)\, \mathrm{d}x + \int_\Omega [1 - |p^*(x)|^2]^{1/2}\, \mathrm{d}x \right]. \tag{1.13}$$

Hypothesis III(4.21) is satisfied since the function G is continuous in $L^1(\Omega)^n$ and Theorem III.4.1 applies: problem (1.13) possesses a solution $\bar{p}^*$, and this solution is unique since

$$p^* \mapsto \int_\Omega [1 - |p^*(x)|^2]^{1/2}\, \mathrm{d}x$$

is strictly concave on the set $\{p^* \in L^\infty(\Omega)^n, |p^*(x)| \leqslant 1 \text{ a.e.}\}$; furthermore

$$\inf \mathscr{P} = \max \mathscr{P}^*. \tag{1.14}$$

We note that the primal problem (1.3) may or may not possess a solution (*cf.* Remark 1.2). If such a solution $\bar{u}$ exists,[1] Proposition III.4.1 applies and entails the two extremality relations

$$F(\bar{u}) + F^*(\Lambda^*\bar{p}^*) = \langle \Lambda^*\bar{p}^*, \bar{u} \rangle, \tag{1.15}$$

$$G(\Lambda\bar{u}) + G^*(-\bar{p}^*) = -\langle \bar{p}^*, \Lambda\bar{u} \rangle, \tag{1.16}$$

where the first is trivial and the second is interpreted in the light of Remark 1.1. This leads us to

Proposition 1.1. *Problem* (1.3) (*the non-parametric minimal hypersurface problem*) *and problem* (1.13) *are mutually dual,*

$$\inf \mathscr{P} = \sup \mathscr{P}^*. \tag{1.17}$$

Problem (1.3) *may or may not possess a solution, whereas problem* (1.13) *possesses a unique solution* $\bar{p}^*$. *If problem* (1.3) *possesses a solution* $\bar{u}$ *then*

$$|\bar{p}^*(x)| < 1 \quad \text{a.e.} \quad x \in \Omega, \tag{1.18}$$

and

$$\operatorname{grad} \bar{u}(x) = -\frac{p^*(x)}{[1 - |p^*(x)|^2]^{1/2}} \quad \text{a.e.} \quad x \in \Omega. \tag{1.19}$$

Remark 1.2. When problem (1.3) does not possess a solution, we may ask whether the right-hand side of (1.19) is sufficiently regular to represent the gradient of a function of $W^{1,1}(\Omega)$ and if so what is the meaning of this function for the minimal hypersurface problem. This is the fundamental problem to be studied hereafter and, at the same time, we shall be in a position to demonstrate the regularity properties of $\bar{p}^*$.

1.2. Fundamental property of a minimizing sequence

In the rest of this chapter, we shall make considerable use of the following important remark. Let us consider in a general way a minimization problem $\mathscr{P}$,

$$\operatorname*{Inf}_{u \in V} [F(u) + G(\Lambda u)]. \tag{1.20}$$

$F \in \Gamma_0(V)$, $G \in \Gamma_0(V)$ and let $\mathscr{P}^*$ be its dual:

$$\operatorname*{Sup}_{p^* \in Y^*} [-F^*(\Lambda^*p^*) - G^*(-p^*)]. \tag{1.21}$$

[1] It is necessarily unique since $\int_\Omega [1 + \Lambda u(x)^2]^{1/2} dx$ being strictly convex, if u_1 and u_2 are two solutions, grad $u_1 =$ grad u_2 which together with $u_1 - u_2 \in W_0^{1,1}(\Omega)$ ($u_1 = u_2$ on Γ) implies that $u_1 = u_2$.

We shall assume that

$$\inf \mathscr{P} = \sup \mathscr{P}^* \tag{1.22}$$

and

$$\mathscr{P}^* \text{ possesses a solution } \bar{p}^*. \tag{1.23}$$

This is exactly the case for problems (1.3) and (1.13).

Let v_m be a minimizing sequence of $\mathscr{P}$: $v_m \in V$, $\forall m$, and

$$F(v_m) + G(\Lambda v_m) \to \inf \mathscr{P}, \tag{1.24}$$

i.e.

$$F(v_m) + G(\Lambda v_m) = \inf \mathscr{P} + \rho_m, \tag{1.25}$$

$$\rho_m \geqslant 0, \qquad \rho_m \to 0, \qquad m \to \infty. \tag{1.26}$$

Proposition 1.2. *Under hypotheses* (1.22) *and* (1.23), *if* v_m *is a minimizing sequence of* $\mathscr{P}$ *satisfying* (1.25) *then,* $\forall m$,

$$\Lambda^* \bar{p}^* \in \partial_{\rho_m} F(v_m) \tag{1.27}$$

$$- \bar{p}^* \in \partial_{\rho_m} G(\Lambda v_m). \tag{1.28}$$

Proof. Using (1.22) and (1.23), we can write

$$F(v_m) + G(\Lambda v_m) = \inf \mathscr{P} + \rho_m = - F^*(\Lambda^* \bar{p}^*) - G^*(- \bar{p}^*) + \rho_m,$$

whence

$$[F(v_m) + F^*(\Lambda^* \bar{p}^*) - \langle \Lambda^* p^*, v_m \rangle] + [G(\Lambda v_m) + G^*(- \bar{p}^*) + \langle \bar{p}^*, \Lambda v_m \rangle] = \rho_m ;$$

but since, by definition of a polar function (*cf.* I(4.3)), each of the expressions within square brackets is $\geqslant 0$, we see that

$$0 \leqslant F(v_m) + F^*(\Lambda^* \bar{p}^*) - \langle \Lambda^* \bar{p}^*, v_m \rangle \leqslant \rho_m,$$
$$0 \leqslant G(\Lambda v_m) + G^*(- \bar{p}^*) + \langle \bar{p}^*, \Lambda v_m \rangle \leqslant \rho_m,$$

which is identical to (1.27) and (1.28) by definition (*cf.* I(6.1)).

1.3. Regularity of the solution of (1.13)

The following results are valid if the hypothesis

$$\phi \in W^{1,1}(\Omega) \cap L^\infty(\Omega),$$

holds, but to simplify the presentation we shall assume that

$$\phi \in W^{1,1}(\Omega) \cap \mathscr{C}(\bar{\Omega}), \tag{1.29}$$

($\mathscr{C}(\bar{\Omega})$ = real continuous functions on $\bar{\Omega}$).

Paragraph 2 will deal (for much more general situations) with the case where $\phi \in W^{1,1}(\Omega) \cap L^\infty(\Omega)$.

Lemma 1.2. *There exist minimizing sequences of* (1.3), *constituted with continuous functions* v_m, *with*

$$\|v_m\|_{L^\infty(\Omega)} \leqslant \text{constant}, \qquad \|v_m\|_{W^{1,1}(\Omega)} \leqslant \text{constant}. \tag{1.30}$$

Proof. If v_m is a minimizing sequence of (1.3),

$$\int_\Omega (1 + |\text{grad } v_m|^2)^{1/2} \, dx \leqslant \text{constant},$$

so that

$$\|\text{grad } v_m\|_{L^1(\Omega)^n} \leqslant \text{constant}$$

and since $v_m - \phi \in W_0^{1,1}(\Omega)$, Poincare's inequality

$$\|v_m - \phi\|_{L^1(\Omega)} \leqslant c(\Omega) \, \|\text{grad } (v_m - \phi)\|_{L^1(\Omega)^n}$$

implies that $\|v_m\|_{L^1(\Omega)^n} \leqslant$ constant and thus any minimizing sequence of (1.3) is bounded in $W^{1,1}(\Omega)$.

Since $\mathscr{D}(\Omega)$ is dense in $W_0^{1,1}(\Omega)$, there exist minimizing sequences of the type

$$v_m = \phi + \theta_m, \qquad \theta_m \in \mathscr{D}(\Omega),$$

and in this case $v_m \in \mathscr{C}(\bar{\Omega})$. For a sequence v_m of this type, let w_m be the function defined by

$$w_m(x) = \begin{vmatrix} v_m(x) & \text{if } |v_m(x)| \leqslant M, \\ M & \text{if } \ v_m(x) \geqslant M, \\ -M & \text{if } \ v_m(x) \leqslant -M, \end{vmatrix}$$

where $M = \|\phi\|_{\mathscr{C}(\bar{\Omega})}$. The functions w_m are likewise continuous and $w_m \in W^{1,1}(\Omega)$, $\forall m$, the truncations being continuous in the $W^{1,\alpha}(\Omega)$ (*cf.* Stampacchia [1]); furthermore (*cf.* Stampacchia loc. cit.) for almost all $x \in \Omega$

$$\frac{\partial w_m}{\partial x_i}(x) = \begin{vmatrix} \dfrac{\partial v_m}{\partial x_i}(x) & \text{if } |v_m(x)| \leqslant M \\ 0 \quad \text{if } |v_m(x)| > M, \end{vmatrix}$$

and therefore

$$\int_{\Omega} [1 + |\Lambda w_m(x)|^2]^{1/2} \, dx \leqslant \int_{\Omega} [1 + |\Lambda v_m(x)|^2]^{1/2} \, dx,$$

so that w_m is likewise a minimizing sequence of (1.3) [naturally, we have $w_m - \phi \in W_0^{1,1}(\Omega)$].

Lemma 1.3. *Let $\mathcal{O}$ be a ball, $\mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega$. We can find a minimizing sequence of* (1.3) *which satisfies the properties stated in Lemma* 1.2 *and furthermore*

$$\sup_{x \in \mathcal{O}} |\operatorname{grad} w_m(x)| \leqslant c, \tag{1.31}$$

$$\| v_m|_{\mathcal{O}} \|_{H^2(\mathcal{O})} \leqslant c, \tag{1.32}$$

(c denoting constants independent of m).

Proof. There exists a ball $\mathcal{O}'$, $\bar{\mathcal{O}} \subset \mathcal{O}' \subset \bar{\mathcal{O}}' \subset \Omega$. Let us consider a sequence v_m as given by Lemma 1.2. For all m, the problem

$$\inf_{\substack{\psi \in W^{1,1}(\mathcal{O}') \\ \psi = v_m \text{ on } \partial\mathcal{O}'}} \int_{\mathcal{O}'} (1 + |\operatorname{grad} \psi|^2)^{1/2} \, dx, \tag{1.33}$$

admits a unique solution ψ_m, since $v_m|_{\partial\Omega}$ is continuous and $\mathcal{O}'$ is strictly convex (*cf.* M. Miranda [1]). The function ψ_m is the solution of the Dirichlet problem associated with the equation of minimal hypersurfaces:

$$\begin{aligned} &\sum_{i=1}^{n} D_i \frac{D_i \psi}{(1 + |\operatorname{grad} \psi|^2)^{1/2}} = 0, \qquad \text{in } \mathcal{O}' \\ &\psi = v_m \quad \text{on} \quad \partial\mathcal{O}'. \end{aligned} \tag{1.34}$$

Obviously, we have

$$\int_{\mathcal{O}'} (1 + |\operatorname{grad} \psi_m|^2)^{1/2} \, dx \leqslant \int_{\mathcal{O}'} (1 + |\operatorname{grad} v_m|^2)^{1/2} \, dx,$$

and hence the function w_m

$$w_m(x) = \begin{vmatrix} v_m(x), & x \notin \mathcal{O}', \\ \psi_m(x), & x \in \mathcal{O}', \end{vmatrix}$$

is in $\phi + W_0^{1,1}(\Omega)$, and constitutes *a fortiori* a minimizing sequence of (1.3).

From the maximum principle applied to (1.34),

$$\sup_{\mathcal{O}'} |\psi_m(x)| \leqslant \sup_{\partial\mathcal{O}'} |v_m(x)| \leqslant \|\phi\|_{\mathscr{C}(\bar{\Omega})}. \tag{1.35}$$

By virtue of a fundamental result of Bombieri, De Giorgi and Miranda [1], which gives an *a priori* estimate of the solutions of the equation of minimal hypersurfaces, (1.34) and (1.35) imply that ($\mathcal{O} \subset \bar{\mathcal{O}} \subset \mathcal{O}'$):

$$\text{(1.36)} \qquad \operatorname*{Sup}_{x\in\mathcal{O}} |\operatorname{grad} \psi_m(x)| \leqslant c = c(\mathcal{O}, \mathcal{O}', \|\phi\|_{L^\infty(\Omega)})$$

$$\text{(1.37)} \qquad \left\| \frac{\partial^2 \psi_m}{\partial x_i\, \partial x_j} \right\|_{L^2(\Omega)} \leqslant c = c(\mathcal{O}, \mathcal{O}', \|\phi\|_{L^\infty(\Omega)})^{(1)}.$$

The minimizing sequence w_m satisfies all the required properties. ■

Lemma 1.4. *The solution $\bar{p}^*$ of* (1.13) *satisfies*

$$\text{(1.38)} \qquad \operatorname*{Sup}_{x\in\mathcal{O}} |\bar{p}^*(x)| \leqslant 1 - \eta(\mathcal{O}), \qquad \eta(\mathcal{O}) > 0,$$

for any set $\mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega$.

Proof. To establish (1.38) we assume that $\mathcal{O}$ is a weakly compact ball in Ω; it is then easy to pass on to the case where $\mathcal{O}$ is a weakly compact subset of Ω by taking a finite covering of such balls.

For a fixed ball $\mathcal{O}$ we consider the minimizing sequence v_m defined in Lemma 1.3. By virtue of (1.30), (1.31) and (1.32), there exists $v \in L^\infty(\Omega)$, $v|_\mathcal{O} \in H^2(\mathcal{O}) \cap W^{1,\infty}(\mathcal{O})$, and there exists a subsequence $m' \mapsto \infty$ such that

$$\text{(1.39)} \qquad v_{m'} \to u \text{ in the weak star topology of } L^\infty(\Omega),$$

$$\text{(1.40)} \qquad \partial v_{m'}/\partial x_i \to \partial u/\partial x_i \text{ in the weak star topology of } L^\infty(\mathcal{O})\ (1 \leqslant i \leqslant n),$$

$$\text{(1.41)} \qquad v_{m'} \to u \text{ weakly in } H^2(\mathcal{O}).$$

Since the injections of $W^{1,1}(\Omega)$ into $L^1(\Omega)$ and of $H^2(\Omega)$ into $H^1(\mathcal{O})$ are compact (*cf.* J. L. Lions [2]),

$$\text{(1.42)} \qquad v_{m'} \to u \text{ strongly in } L^1(\Omega),$$

$$\text{(1.43)} \qquad v_{m'} \to u \text{ strongly in } H^1(\mathcal{O}).$$

Thus we can choose the sequence m' so that

$$\text{(1.44)} \qquad v_{m'}(x) \to u(x) \quad \text{a.e.} \quad x \in \Omega,$$

$$\text{(1.45)} \qquad \frac{\partial v_{m'}}{\partial x_i}(x) \to \frac{\partial u}{\partial x_i}(x) \quad \text{a.e.} \quad x \in \mathcal{O}, \qquad (1 \leqslant i \leqslant n).$$

Obviously we do not know if $u \in W^{1,1}(\Omega)$ nor *a fortiori* if $u \in \phi + W^{1,1}(\Omega)$ and it is this which prevents us from concluding that u is a solution of (1.3); in general u is not a solution of (1.3).

(1) Only inequality (1.36) is given explicitly in the work of Bombieri, De Giorgi and Miranda. But (1.37) follows at once; *cf.* R. Temam [3], p. 140, for a very similar situation.

Using, for the moment, Proposition 1,2, we see that

(1.46) $$-\bar{p}^* \in \partial_{\rho_m} G(\Lambda v_m),$$

where ρ_m is a sequence which converges to 0 ($\rho_m \geqslant 0$). Theorem I.6.2 applies: there exists $p_m \in L^1(\Omega)^n$, $p_m^* \in L^\infty(\Omega)^n$, with

(1.47) $$\|p_m - \Lambda v_m\|_{L^1(\Omega)^n} \leqslant \sqrt{\rho_m},$$

(1.48) $$\|p_m^* - \bar{p}^*\|_{L^\infty(\Omega)^n} \leqslant \sqrt{\rho_m},$$

and

(1.49) $$-p_m^* \in \partial G(p_m),$$

which, from Remark 1.1, means that

(1.50) $$p_m^*(x) = -\frac{p_m(x)}{[1 + |p_m(x)|^2]^{1/2}} \quad \text{a.e.}$$

From (1.47), $p_m - \Lambda v_m \to 0$ in $L^1(\Omega)^n$, for $m \to \infty$, and so we can choose the sequence m' so that

(1.51) $$p_{m'}(x) - \Lambda v_{m'}(x) \to 0, \quad \text{a.e.} \quad x \in \Omega.$$

With (1.45), (1.48), (1.50) and (1.51) we obtain on passing to the limit

(1.52) $$\bar{p}^*(x) = -\frac{\operatorname{grad} u(x)}{[1 + |\operatorname{grad} u(x)|^2]^{1/2}} \quad \text{a.e.} \quad x \in \mathcal{O},$$

and as $u \in W^{1,\infty}(\mathcal{O})$, property (1.38) follows.

Remark 1.3. To be precise, the constant η of (1.38) depends on $\mathcal{O}$, Ω and $\|\phi\|_{L^\infty(\Omega)}$ (*cf.* (1.36) and (1.37)).

1.4. Generalized solution of (1.3)

Lemma 1.5. *There exists an analytic function u which is bounded ($u \in L^\infty(\Omega)$), which is the solution in Ω of the equation of the minimal hypersurfaces and which satisfies*

(1.53) $$\operatorname{grad} u(x) = -\frac{\bar{p}^*(x)}{[1 - |\bar{p}^*(x)|^2]^{1/2}}, \qquad \forall x \in \Omega.$$

Proof. We consider a bounded minimizing sequence of (1.3), e.g. v_m (*cf.* Lemma 1.2). Then, as in Lemma 1.4, we have a subsequence v_m which satisfies

(1.54) $$v_{m'} \to u \text{ in } L^\infty(\Omega) \text{ in the weak star sense,}$$

$$v_{m'} \to u \text{ strongly in } L^1(\Omega). \tag{1.55}$$

As in Lemma 1.4, we have

$$-\bar{p}^* \in \partial_{\rho_m} G(\Lambda v_m), \qquad \rho_m \geqslant 0, \qquad \rho_m \to 0, \tag{1.56}$$

and there exists $p_m \in L^1(\Omega)^n$, $p_m^* \in L^\infty(\Omega)^n$, such that

$$\|p_m - \Lambda v_m\|_{L^1(\Omega)^n} \leqslant \sqrt{\rho_m}, \tag{1.57}$$

$$\|p_m^* - \bar{p}^*\|_{L^\infty(\Omega)^n} \leqslant \sqrt{\rho_m}, \tag{1.58}$$

$$p_m(x) = -\frac{p_m^*(x)}{[1 - |p_m^*(x)|^2]^{1/2}} \quad \text{a.e.} \quad x \in \Omega. \tag{1.59}$$

From (1.58), for $m \to \infty$,

$$\frac{p_m^*(x)}{[1 - |p^*(x)|^2]^{1/2}} \to \frac{\bar{p}^*(x)}{[1 - |\bar{p}^*(x)|^2]^{1/2}} \quad \text{a.e.} \tag{1.60}$$

If $\mathcal{O}$ is a weakly compact open subset of Ω, then by (1.38), if m is sufficiently large

$$|p_m^*(x)| \leqslant 1 - \tfrac{1}{2}\eta(\mathcal{O}),$$

so that the Lebesgue theorem allows us to conclude that

$$\frac{p_m^*}{(1 - |p^*|^2)^{1/2}} \to \frac{p^*}{(1 - |\bar{p}^*|^2)^{1/2}}$$

in $L^1(\mathcal{O})^n$. With (1.59) and (1.57),

$$\Lambda v_m \to -\frac{\bar{p}^*}{(1 - |\bar{p}^*|^2)^{1/2}} \tag{1.61}$$

in $L^1(\mathcal{O})^n$ and as $\Lambda v_m \to \Lambda u$ in the distribution sense in $\mathcal{O}$

$$\operatorname{grad} u(x) = -\frac{\bar{p}^*(x)}{[1 - |\bar{p}^*(x)|^2]^{1/2}} \quad \text{a.e.} \quad x \in \mathcal{O},$$

and since $\mathcal{O}$ is any relatively compact subset, we deduce that (1.53) holds.

By virtue of (1.53), the property $\operatorname{div}\bar{p}^* = 0$ (*cf.* (1.13)) means that u is the solution of the equation of the minimal hypersurfaces (*cf.* (1.1)).

It follows from (1.53) that $u \in W^{1,\infty}(\mathcal{O})$ for every relatively compact open subset $\mathcal{O}$ of $\Omega(\bar{\mathcal{O}} \subset \Omega)$ and since $\bar{u}$ is solution of the minimal hypersurface equation, $\bar{u}$ is analytic in Ω (*cf.* De Giorgi [1]).

Lemma 1.6. *For every minimizing sequence v_m of* (1.3),

(1.62) $$v_m \to u \quad in \quad L^1(\Omega)/R,$$

(1.63) $$\frac{\partial v_m}{\partial x_i} \to \frac{\partial u}{\partial x_i} \quad in \quad L^1(\mathcal{O}), \quad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega, \qquad 1 \leqslant i \leqslant n.$$

Proof. Exactly the same reasoning as for Lemma 1.5 shows that as the limit is independent of the particular subsequence, the convergences (1.55) (modified in (1.62)) and (1.61) hold for the complete sequence.

2. GENERALIZED SOLUTION OF THE DIRICHLET PROBLEM FOR THE EQUATION OF THE MINIMAL HYPERSURFACES

Orientation

In Section 2.1 we shall give the main theorem which depends on the lemmas of Section 1 and establishes the existence of a generalized solution of problem (1.1) or (1.3). In the subsequent subsections we shall develop various complementary results and remarks.

2.1. Statement of the main result

Theorem 2.1. *Let ϕ be given, satisfying*

(2.1) $$\phi \in \mathscr{W}^{1,1}(\Omega) \cap L^\infty(\Omega).$$

Problem (1.3) *admits problem* (1.13) *as its dual and*

(2.2) $$\inf \mathscr{P} = \sup \mathscr{P}^*.$$

Problem $\mathscr{P}^*$ possesses a unique solution $\bar{p}^*$, which is analytic and satisfies

(2.3) $$\operatorname*{Sup}_{x \in \mathcal{O}} |\bar{p}^*(x)| < 1, \qquad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega.$$

There also exists an analytic function u which is uniquely defined to within an additive constant:

(2.4) $$\operatorname{grad} u(x) = -\frac{\bar{p}^*(x)}{[1 - |\bar{p}^*(x)|^2]^{1/2}}, \qquad x \in \Omega,$$

which is the solution of (1.3) *if such a solution exists and which in every case satisfies:*

(2.5) $$u \in L^\infty(\Omega),$$

(2.6) u *is solution in Ω of the minimal hypersurface equation,* (2.7) $\left| \begin{array}{l} \textit{every minimizing sequence } \{v_m\} \textit{ of } (1.3) \textit{ converges to } u \textit{ in the} \\ \textit{following sense: } v_m \to u \textit{ in } L^1(\Omega)/\mathbf{R}, \\ \partial v_m/\partial x_i \to \partial u/\partial x_i \quad \textit{in} \quad L^1(\mathcal{O}), \quad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega, \quad 1 \leqslant i \leqslant n. \end{array} \right.$

Proof. If, instead of (2.1), $\phi \in \mathscr{W}^{1,1}(\Omega) \cap \mathscr{C}(\bar{\Omega})$, the theorem results as a direct consequence of Lemmas 1.4, 1.5 and 1.6. We pass on to the more general case by approximation of ϕ. If ϕ satisfies (2.1), there exists a sequence of functions $\phi_m \in \mathscr{W}^{1,1}(\Omega) \cap \mathscr{C}(\bar{\Omega})$ such that

(2.8) $$\phi_m \to \phi \quad \text{in} \quad \mathscr{W}^{1,1}(\Omega),$$

(2.9) $$\|\phi_m\|_{\mathscr{C}(\bar{\Omega})} \leqslant \|\phi\|_{L^\infty(\Omega)}.$$

Indeed, by definition of $\mathscr{W}^{1,1}(\Omega)$, there exists a sequence of indefinitely differentiable functions ψ_m, such that $\psi_m \to \phi$ in $\mathscr{W}^{1,1}(\Omega)$; the functions ϕ_m defined by

$$\phi_m(x) = \left| \begin{array}{rl} \psi_m(x) & \text{if } |\psi_m(x)| \leqslant M, \\ M & \text{if } \psi_m(x) > M, \\ -M & \text{if } \psi_m(x) < -M, \end{array} \right.$$

$(M = \|\phi\|)_{L^\infty(\Omega)}$ are in $\mathscr{W}^{1,1}(\Omega) \cap \mathscr{C}(\bar{\Omega})$ and satisfy (2.8) and (2.9).

Then, let $\bar{p}_m^*$ be the solution of problem (1.13) corresponding to ϕ_m and $\bar{p}^*$ the solution of problem (1.13) corresponding to ϕ. The crucial point is to verify (1.38). But the relation (1.38) for $\bar{p}_m^*$, remark 1.3 and (2.9) imply

(2.10) $$\operatorname*{Sup}_{\mathcal{O}} |\bar{p}_m^*(x)| \leqslant 1 - \eta(\mathcal{O}, \Omega, \|\phi_m\|_{\mathscr{C}(\bar{\Omega})})$$

(2.11) $$\operatorname*{Sup}_{\mathcal{O}} |\bar{p}_m^*(x)| \leqslant 1 - \eta(\mathcal{O}, \Omega, \|\phi\|_{L^\infty(\Omega)}).$$

We have thus shown that

(2.12) $$\operatorname*{Sup}_{\mathcal{O}} |\bar{p}^*(x)| \leqslant 1 - \eta(\mathcal{O}, \Omega, \|\phi\|_{L^\infty(\Omega)}) < 1, \qquad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega,$$

if we show that $\bar{p}_m^* \to \bar{p}^*$, for example in the weak star sense in $L^\infty(\Omega)$ (which implies that for all $\mathcal{O} \subset \Omega$, the restriction of $\bar{p}_m^*$ to $\mathcal{O}$ converges to the restriction of $\bar{p}^*$ to $\mathcal{O}$ in the weak star sense in $L^\infty(\mathcal{O})$).

Since $|\bar{p}_m^*(x)| \leqslant 1$ almost everywhere, there exists a subsequence $\bar{p}_m^*$ and q^* $L^\infty(\Omega)$, such that

(2.13) $$\bar{p}_{m'}^* \to q^* \quad \text{for the weak star topology of } L^\infty(\Omega).$$

Thus $\bar{p}_m^*$ being the solution of (1.13) for ϕ_m, we have

$$-\int_\Omega \bar{p}_m^*(x)\Lambda\phi_m(x)\,\mathrm{d}x + \int_\Omega [1-|\bar{p}_m^*(x)|^2]^{1/2}$$

$$\geqslant -\int_\Omega \bar{p}^*(x)\Lambda\phi_m(x)\,\mathrm{d}x + \int_\Omega [1-|p^*(x)|^2]^{1/2}\,\mathrm{d}x\,;$$

this gives us for $m' \to \infty$, and on taking the upper limit of each term and using the u.s.c. of the functionals:

$$-\int_\Omega \bar{p}^*(x)\Lambda\phi(x)\,\mathrm{d}x + \int_\Omega [1-|q^*(x)|^2]^{1/2}\,\mathrm{d}x$$

$$\geqslant -\int_\Omega \bar{p}^*(x)\Lambda\phi(x)\,\mathrm{d}x + \int_\Omega [1-|\bar{p}^*(x)|^2]^{1/2}\,\mathrm{d}x.$$

As $\bar{p}^*$ is the only solution of (1.13), we deduce that $q^* = \bar{p}^*$, and the whole sequence $\bar{p}_m^*$ converges to $\bar{p}^*$ in $L^\infty(\Omega)$ in the weak star sense.

As stated above, (2.11) gives (2.12) in the limit. It is easily seen that from (2.12) we can take up point by point the proofs of Lemmas 1.5 and 1.6 and thereby obtain all the stated results in Theorem 2.1. ■

Remark 2.1. The techniques used in Section 3 for more general problems than (1.3) show in fact that

$$u \in W^{1,1}(\Omega). \tag{2.14}$$

Remark 2.2. We cannot hope in (2.7) for strong convergence of $\partial v_m/\partial x_i$ to $\partial u/\partial x_i$ in the whole set Ω (e.g. in $L^1(\Omega)$): for then we would have $u = v_m = \phi$ on Γ which is not necessarily true (*cf.* Remark 2.3).

Remark 2.3. If problem (1.3) possesses a solution, then, obviously, u is equal to this solution (to within a constant): we establish this result by applying (2.7) to the minimizing sequence

$$v_m = u, \qquad \forall m \geqslant 1.$$

If problem (1.3) does not possess a solution (which is what happens in many standard examples), *the function u is not a solution of problem* (1.3) and this is evident from the fact that it does not satisfy the required limiting condition $u = \phi$ on Γ. This condition is only satisfied on one possibly empty subset of Γ. By virtue of property (2.7), *it is nonetheless reasonable to term this function u the generalized solution of problem* (1.3).

Remark 2.4. (Open problems).

(i) The essential problem set by Theorem 2.1 is to know whether $u = \phi$ on F in at least one part of Γ.

In Section 2.2, we shall give some partial answers to this question (*cf.* Theorem 2.2, Propositions 2.2 and 2.3).

(ii) When the open set Ω is regular, we can define because of Gagliardo [1] the trace of u on Γ, $\gamma_0 u \in L^1(\Gamma)$; this trace only depends on the trace of ϕ on Γ, $\gamma_0\phi$. It would be interesting to study the correspondence between $\gamma_0\phi$ and $\gamma_0 u$. In this case, too, we shall give some incomplete results in Section 2.2. ■

2.2 Boundary values of the generalized solution

Proposition 2.1. *Let us assume that the open set Ω is regular, of class $\mathscr{C}^2$, and that $\phi \in \mathscr{C}^3(\bar{\Omega})$. Let us also assume that Γ has a non-negative mean curvature everywhere. Then the function u given by Theorem* 2.1 *satisfies*[1] $u = \phi$ *on Γ and is a solution of* (1.3).

Proof. From Serrin [1] (*cf.* also D. Gilbarg [1], G. Stampacchia [3]), problem (1.3) possesses a solution and from Remark 2.3, this solution differs from u by a constant which we may choose to be 0.

After this easy result, we note the following result concerning the behaviour of u on Γ (*cf.* Lichnewsky [1] [2]).

Theorem 2.2. *Assume that Ω is a lipschitzian set and that Γ_1 is an open subset of Γ. Assume that Γ_1 is a $\mathscr{C}^3$ manifold with a non-negative mean curvature and that besides* (2.1)

$$\phi \in \mathscr{C}^2(\Gamma_1). \tag{2.15}$$

Then there exists a unique function u which satisfies all the conclusions of Theorem 2.1 *and furthermore*

$$u = \phi \quad on \quad \Gamma_1. \tag{2.16}$$

(2.17) *every minimizing sequence $\{v_m\}$ of* (1.3) *converges to u in the following sense*

$$v_m \to u \quad in \quad L^1(\Omega)^{(2)}$$

$$\frac{\partial v_m}{\partial x_i} \to \frac{\partial u}{\partial x_i} \quad in \quad L^1(\mathcal{O}), \qquad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega.$$

The proof of this theorem is based on the utilization of local barriers.

[1] More precisely one of the functions $u + c$, $c \in \mathbf{R}$.

[2] And not only in $L^1(\Omega)/\mathbf{R}$.

Let x_0 belong to Γ; an upper local barrier at x_0 for the Dirichlet Problem (1.1) is a function $\psi = \psi_\lambda^{x_0}$, $\lambda > 0$ satisfying

$$(2.18)\quad \left|\begin{array}{l} \text{(i) } \psi \in \mathscr{C}^2(\bar{\Omega} \cap \mathcal{O}_1), \\ \text{(ii) } \operatorname{div}\left(\dfrac{\operatorname{grad}\psi}{[1+|\operatorname{grad}\psi|^2]^{1/2}}\right) \leqslant 0 \quad \text{in} \quad \Omega \cap \mathcal{O}_1, \\ \text{(iii) } \psi(y_0) = \phi(y_0), \qquad \text{(iv) } \psi(y) \geqslant \phi(y) \quad \text{on} \quad \Gamma \cap \mathcal{O}_1, \\ \text{(v) } \psi \geqslant \lambda \quad \text{in} \quad \bar{\Omega} \cap (\mathcal{O}_1 - \mathcal{O}_2), \end{array}\right.$$

where $\mathcal{O}_1$ and $\mathcal{O}_2$ are two open neighbourhoods of x_0, $\bar{\mathcal{O}}_2 \subset \mathcal{O}_1$.

A lower local barrier is defined in a similar way, and we have

Lemma 2.1. *Under the assumptions of Theorem* 2.2, *at each point* $x_0 \in \Gamma_1$, *there exists a lower and an upper local barrier for the Problem* (1.1) *with* λ *arbitrary.*

This result is essentially classical and can be obtained by a slight improvement of the techniques of J. Serrin [1]; see, for instance, A. Lichnewsky [2].

Proof of Theorem 2.2. For each $x_0 \in \Gamma_1$, we consider the barrier $\psi_\lambda^{x_0}$, with fixed λ,

$$(2.19)\qquad \lambda > \|\phi\|_{L^\infty(\Omega)}.$$

Let u_n be a bounded minimizing sequence of (1.3) with:

$$\left|\begin{array}{l} u_n = \phi \quad \text{on} \quad \partial\Omega, \qquad \|u_n\|_{L^\infty(\Omega)} \leqslant \lambda \\ \displaystyle\int_\Omega \sqrt{1+|\operatorname{grad} u_n|^2}\,dx \to \inf \mathscr{P}. \end{array}\right.$$

Because of Lemma 1.2 such a minimizing sequence exists.

Let $v_n = \inf(\psi, u_n)$ in $\mathcal{O}_1$ and $v_n = u_n$ in $\Omega - \bar{\mathcal{O}}_2$. The proof will consist in showing first that v_n is also minimizing sequence of (1.3) and then to take advantage of this result.

(a) We choose lipschitzian open set $\mathcal{O}$, $\bar{\mathcal{O}}_2 \subset \mathcal{O}$, $\bar{\mathcal{O}} \subset \mathcal{O}_1$, and we note that $v_n \in W^{1,1}(\mathcal{O})$, $v_n \in W^{1,1}(\Omega - \bar{\mathcal{O}}_2)$ and then $v_n \in W^{1,1}(\Omega)$. It is also clear that $v_n = \phi$ on $\partial\Omega$ (see (2.18) (iv)).

Now we denote as θ,

$$(2.20)\qquad \theta = \operatorname{div}\left(\frac{\operatorname{grad}\psi}{[1+|\operatorname{grad}\psi|^2]^{1/2}}\right) \leqslant 0 \quad \text{in} \quad \Omega \cap \mathcal{O}.$$

It is clear with (2.20) that ψ is the unique solution of the variational problem

$$(2.21)\qquad \operatorname*{Inf}_{\substack{\sigma=\psi \\ \text{on } \partial(\Omega\cap\mathcal{O})}} \left\{\int_{\Omega\cap\mathcal{O}} [1+|\operatorname{grad}\sigma|^2]^{1/2}\,dx + \int_{\Omega\cap\mathcal{O}} \mathcal{O}\sigma\,dx\right\}.$$

The function $w_n = \operatorname{Sup}(\psi, u_n)$ satisfies $w_n = \psi$ on $\partial(\Omega \cap \mathcal{O})$; indeed $w_n = \psi$ on $\Gamma \cap \partial(\Omega \cap \mathcal{O})$ since $\psi \geqslant \phi = u_n$ on this part of the boundary and $w_n = \psi$ on

$\Omega \cap \partial(\Omega \cap \mathcal{O})$ since $\psi \geqslant \lambda \geqslant u_n$ in $\mathcal{O}_1 - \overline{\mathcal{O}}_2$. Since ψ is the unique solution of (2.21) and since $\theta(w_n - \psi) \leqslant 0$, we have:

$$\int_{\Omega \cap \mathcal{O}} [1 + |\operatorname{grad} \psi|^2]^{1/2} \, dx \leqslant \int_{\Omega \cap \mathcal{O}} [1 + |\operatorname{grad} w_n|^2]^{1/2} \, dx.$$

We observe that $\operatorname{grad} w_n = \operatorname{grad} u_n$ a.e. on the set

$$\mathscr{A} = \{x \in \mathcal{O}, \psi(x) \leqslant u_n(x)\}$$

and $\operatorname{grad} w_n = \operatorname{grad} \psi$ a.e. on the set $\mathcal{O} - \mathscr{A}$. Hence

$$\int_{\mathscr{A}} [1 + |\operatorname{grad} \psi|^2]^{1/2} \, dx \leqslant \int_{\Omega} [1 + |\operatorname{grad} u_n|^2]^{1/2} \, dx. \tag{2.22}$$

For v_n, we have $\operatorname{grad} v_n = \operatorname{grad} u_n$ a.e. in $\Omega - \mathscr{A}$ and $\operatorname{grad} v_n = \operatorname{grad} \psi$ a.e. in $\mathscr{A}$. From (2.22) we then find

$$\int_{\Omega} [1 + |\operatorname{grad} v_n|^2]^{1/2} \, dx \leqslant \int_{\Omega} [1 + |\operatorname{grad} u_n|^2]^{1/2} \, dx,$$

and since $v_n = \phi$ on $\partial\Omega$, v_n is also a minimizing sequence of our problem (1.3).

(b) Theorem 2.1 is now applicable to both sequences, v_n and u_n; in particular

$$u_n \to u \quad \text{in} \quad L^1(\Omega)$$
$$v_n \to u + c \quad \text{in} \quad L^1(\Omega).$$

Since $v_n = u_n$ on a set of positive measure (in $\Omega - \overline{\mathcal{O}}_2$), we must have $c = 0$.

We recall that $v_n \leqslant \psi$ (the upper barrier) in some neighbourhood $\mathcal{O}$ of $x_0 \in \Gamma_1$. Then

$$u \leqslant \psi \quad \text{in} \quad \mathcal{O}. \tag{2.23}$$

If u is continuous in $\bar{\Omega}$ then (2.23) and (2.18) (iii) imply that

$$u(x_0) \leqslant \phi(x_0), \qquad \forall x_0 \in \Gamma_1. \tag{2.24}$$

If u is not continuous and belongs only to $W^{1,1}(\Omega)$, we use the fact that the trace of u over Γ is the limit of u along the normals of Γ, for almost all these normals ($d\Gamma$ measure). We then deduce from (2.23) that for almost all $x_0 \in \Gamma_1$,

$$u(x_0) \leqslant \phi(x_0), \tag{2.25}$$

which is a weaker form of (2.24), the best result we can expect if we only know that $u \in W^{1,1}(\Omega)$.

Using similarly lower local barriers we find that

$$u(x_0) \geqslant \phi(x_0), \quad \text{a.e.} \quad x_0 \in \Gamma, \tag{2.26}$$

and Theorem 2.2 follows from this and (2.25). ■

Another type of result related to the behaviour of u on Γ is given by next proposition.

Proposition 2.2. *The hypotheses are those of Theorem* 2.1 *and we also assume that one of the functions u defined by Theorem* 2.1 *satisfies*

$$\left| \begin{array}{l} \exists\, x_0 \in \Gamma \text{ such that} \\ \overline{\lim\limits_{\substack{x \to x_0 \\ x \in \Omega}}}\, |\mathrm{grad}\, u(x)| < +\infty. \end{array} \right. \tag{2.27}$$

Then there exists a unique function u which satisfies all the conclusions of Theorem 2.1 *and, furthermore,*

$$\left| \begin{array}{l} u = \phi \text{ on a subset of } \Gamma \text{ with non-zero measure and more precisely on} \\ \{ x \in \Gamma \,|\, \overline{\lim\limits_{\substack{y \to x \\ y \in \Omega}}}\, |\mathrm{grad}\, u(y)| < +\infty \}; \end{array} \right. \tag{2.28}$$

$$\left| \begin{array}{l} \text{every minimizing sequence } \{v_m\} \text{ of (1.3) converges to } u \text{ in the} \\ \text{following sense} \\ v_m \to u \quad \text{in} \quad L^1(\Omega), \\ \partial v_m/\partial x_i \to \partial u/\partial x_i \quad \text{in} \quad L^1(\mathcal{O}), \qquad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega. \end{array} \right. \tag{2.29}$$

Proof. From (2.15) it follows that $\mathrm{grad}\, u(x)$ is bounded over an open set $\Omega_0 = B_\rho(x_0) \cap \Omega$, where $B_\rho(x_0)$ is an open ball with centre x_0. From (2.4) this implies that:

$$\operatorname*{Sup}_{x \in \Omega_0} |\bar{p}^*(x)| < 1. \tag{2.30}$$

We can then repeat the reasoning of Lemmas 1.5 and 1.6. For every minimizing sequence $\{v_m\}$ of (1.3) we have, as in these lemmas, the existence of a subsequence m' such that

$$\begin{array}{ll} v_{m'} \to u + c \quad \text{strongly in} \quad L^1(\Omega_0) & \\ \partial v_{m'}/\partial x_i \to \partial u/\partial x_i \quad \text{strongly in} \quad L^1(\Omega_0), & 1 \leqslant i \leqslant n, \end{array} \tag{2.31}$$

where u is one of the functions obtained from Theorem 2.1 and $c \in \mathbf{R}$ is some suitable constant.

Since $v_m = \phi$ on Γ, we have $v_m = \phi$ on $\Gamma_0 = \Gamma \cap \partial\Omega_0$ and so in the limit

$$u - c - \phi = 0 \quad \text{on} \quad \Gamma_0. \tag{2.32}$$

The constant c is thus independent of the chosen subsequence $v_{m'}$ and it is the whole sequence v_m which gives rise to the convergence limits in (2.31). We can thus take $c = 0$, and because of (2.4) and (2.32) the function u is uniquely defined.

The reasoning which has just been followed for the point x_0 can be repeated for every point $x \in \Gamma$ such that

$$\overline{\lim_{\substack{y \to x \\ y \in \Omega}}} |\text{grad } u(y)| < +\infty. \blacksquare$$

Remark 2.5. (In conjunction with Remark 2.4.) It would be interesting to determine some conditions on x_0, Ω, and ϕ which give (2.28) *a priori.*

Example. 2.1. If Ω is a ring of $\mathbf{R}^2$, $1 < |x| < 2$, and if ϕ is given on Γ, $\phi = 0$ for $|x| = 2$, $\phi = c$ for $|x| = 1$, problem (1.3) shows an axial symmetry and can be reduced easily to a one-dimensional problem:

$$\text{(2.33)} \qquad \operatorname*{Inf}_{\substack{u(1)=c \\ u(2)=0}} \int_1^2 [1 + (\mathrm{d}u/\mathrm{d}\rho)^2]^{1/2} \rho \, \mathrm{d}\rho.$$

If c is small, $c \leqslant c_*$, the solution of (2.33) is an arc of a catenary. If $c > c_*$ there is no solution of (2.33). The meridians of the minimal hypersurface are arcs of a catenary with vertical tangent at $|x| = 1$, and a segment $|x| = 1$, $c_* \leqslant u \leqslant c$. It can be verified from Theorem 2.2 or Proposition 2.2 that this arc of a catenary is the generalized solution of the problem. ■

In conjunction with the problem posed in Remark 2.4(ii), we have

Proposition 2.3. *Let us assume that Ω is regular of class $\mathscr{C}^2$, and for every $\phi \in W^{1,1}(\Omega) \cap L^\infty(\Omega)$, we term $\bar{p}^*(\phi)$ the solution of* (1.13).

Let us also assume that, for a given function ϕ, $\bar{p}^*(\phi)\nu = +1$[(1)] *on a subset Γ_+ of Γ, and $\bar{p}^*(\phi)\nu = -1$ on a subset Γ_- of Γ.*

Then if $\phi' \in \mathscr{W}^{1,1}(\Omega) \cap L^\infty(\Omega), \phi' \leqslant \phi$ on $\Gamma_+, \phi' \geqslant \phi$ on $\Gamma_-, \phi' = \phi$ on $\Gamma_0 = \Gamma - (\Gamma_+ \cup \Gamma_-)$, we have

$$\text{(2.34)} \qquad \bar{p}^*(\phi) = \bar{p}^*(\phi').$$

Proof. To simplify the notation we write $\bar{p}^*(\phi) = p, p^*(\phi') = p'$. From Lions and Magenes [1] (*cf.* note (1)) we have, as Ω is regular:

(1) From a trace theorem of Lions and Magenes [1], since $\bar{p}^* \in L^2(\Omega)^n$ and since $\operatorname{div} \bar{p}^* \in L^2(\Omega)$ ($\operatorname{div} \bar{p}^* = 0$), we can define the trace of $\bar{p}^* \cdot \nu$ on Γ, as an element of $H^{-1/2}(\Gamma)$, the dual space of $H^{1/2}(\Gamma)$. Since moreover $\bar{p}^* \in L^\infty(\Omega)^n$, it is easy to show, using the methods of Lions and Magenes [1], that the trace of $\bar{p}^* \cdot \nu$ belongs to $L^\infty(\Gamma)$ and

$$|p^* \cdot \nu|_{L^\infty(\Gamma)} \leqslant |\bar{p}^*|_{L^\infty(\Omega)^n}.$$

$$-\int_\Omega p \operatorname{grad} \phi \, \mathrm{d}x = -\int_\Gamma p\nu\phi \, \mathrm{d}\Gamma + \int_\Omega (\operatorname{div} p)\phi \, \mathrm{d}x$$

$$= -\int_\Gamma p\nu\phi \, \mathrm{d}\Gamma - \int_\Omega p' \operatorname{grad} \phi' \, \mathrm{d}x = -\int_\Gamma p'\nu\phi' \, \mathrm{d}\Gamma.$$

By definition of p' we then have

$$-\int_\Gamma p\nu\phi' \, \mathrm{d}\Gamma + \int_\Omega (1 - |p|^2)^{1/2} \, \mathrm{d}x \leqslant$$

$$\leqslant -\int_\Gamma p'\nu\phi' \, \mathrm{d}\Gamma + \int_\Omega (1 - |p'|^2)^{1/2} \, \mathrm{d}x.$$

Decomposing Γ into $\Gamma_0 \cup \Gamma_+ \cup \Gamma_-$, we write:

$$\begin{aligned} &-\int_{\Gamma_0} p\nu\phi \, \mathrm{d}\Gamma - \int_{\Gamma_+} (1 - p'\nu)\phi' \, \mathrm{d}\Gamma + \int_{\Gamma_-} (1 + p'\nu)\phi' \, \mathrm{d}\Gamma \\ &+ \int_\Omega (1 - |p|^2)^{1/2} \, \mathrm{d}x \leqslant -\int_{\Gamma_0} p'\nu\phi \, \mathrm{d}\Gamma + \int_\Omega (1 - |p'|^2)^{1/2} \, \mathrm{d}x. \end{aligned} \tag{2.35}$$

But

$$-\int_{\Gamma_+} (1 - p'\nu)\phi \, \mathrm{d}\Gamma \leqslant -\int_{\Gamma_+} (1 - p'\nu)\phi' \, \mathrm{d}\Gamma,$$

$$\int_{\Gamma_-} (1 + p'\nu)\phi \, \mathrm{d}\Gamma \leqslant \int_{\Gamma_-} (1 + p'\nu)\phi' \, \mathrm{d}\Gamma,$$

and (2.35) implies that:

$$-\int_{\Gamma_0} p\nu\phi \, \mathrm{d}\Gamma - \int_{\Gamma_+} (1 - p'\nu)\phi \, \mathrm{d}\Gamma + \int_{\Gamma_-} (1 + p'\nu)\phi \, \mathrm{d}\Gamma$$

$$+ \int_\Omega (1 - |p|^2)^{1/2} \, \mathrm{d}x \leqslant$$

$$\leqslant -\int_{\Gamma_0} p'\nu\phi \, \mathrm{d}\Gamma + \int_\Omega (1 - |p'|^2)^{1/2} \, \mathrm{d}x,$$

which yields the inequality

$$
(2.36) \qquad \begin{aligned} & -\int_\Gamma p\nu\phi \, \mathrm{d}\Gamma + \int_\Omega (1 - |p|^2)^{1/2} \, \mathrm{d}x \leqslant \\ & -\int_\Gamma p'\nu\phi \, \mathrm{d}\Gamma + \int_\Omega (1 - |p'|^2)^{1/2} \, \mathrm{d}x, \end{aligned}
$$

which implies that $p' = p$, by definition of p. ■

Remark 2.6. We will prove very simply in Section 2.3 that

$$
\begin{aligned} u \leqslant \phi \quad & \text{a.e. on} \quad \Gamma_+ \\ u \geqslant \phi \quad & \text{a.e. on} \quad \Gamma_- \\ u = \phi \quad & \text{a.e. on} \quad \Gamma_0. \end{aligned} \quad ■
$$

Remark 2.7. Proposition 2.3 may be interpreted in term of the generalized solution $u = u(\phi)$ of problem (1.3). Indeed we observe that $p.\nu = \pm 1$ implies $|\partial u| = +\infty$ and respectively $\partial u/\partial\nu < 0$, $\partial u/\partial\nu > 0$. It is conjectured that for smooth ϕ and $\partial\Omega$, $\partial u/\partial\tau$ is bounded up to the boundary and in this case $p.\nu = \pm 1$ will mean $\partial u/\partial\nu = \mp\infty$. Essentially if $u(\phi)$ satisfies $\partial u/\partial\nu = -\infty$ on Γ_-, $\partial/u\partial\nu = +\infty$ on Γ_+, $\partial u/\partial\nu$ finite on Γ_0, and if we "augment" ϕ on Γ_+, we "diminish" ϕ on Γ_-, and leave ϕ unchanged on Γ_0, then the generalized solution $u = u(\phi)$ is unaltered. ■

2.3. Connection with a problem of de Giorgi

Giusti, de Giorgi and Miranda have introduced a problem which is intimately connected with (1.3) and which can be written, when Ω is regular and of class $\mathscr{C}^2$, as:

$$
(2.37) \qquad \operatorname*{Inf}_{u \in W^{1,1}(\Omega)} \left\{ \int_\Omega (1 + |\operatorname{grad} u|^2)^{1/2} \, \mathrm{d}x + \int_\Gamma |u - \phi| \, \mathrm{d}\Gamma \right\}.
$$

We shall compare (1.3) and (2.37) using duality; we refer to de Giorgi [2] for a direct comparison of the problems, and a direct study of problem (2.37).

To specify the dual problem to (2.37), we set

$$
\begin{aligned}
& V = W^{1,1}(\Omega), \qquad Y = L^1(\Gamma) \times L^1(\Omega)^n, \\
& \Lambda u = \{ \gamma_0 u, \operatorname{grad} u \}, \qquad \forall u \in V, \\
& V^* = \text{the dual of } W^{1,1}(\Omega), \qquad Y^* = L^\infty(\Gamma) \times L^\infty(\Omega)^n, \\
& F(u) = 0, \qquad \forall u \in V, \\
& G(p) = G_0(p_0) + G_1(p_1), \quad p = (p_0, p_1),\ p_0 \in L^1(\Gamma),\ p_1 \in L^1(\Omega)^n,
\end{aligned}
$$

$$G_0(p_0) = \int_\Gamma |p_0 - \phi| \, d\Gamma,$$

$$G_1(p_1) = \int_\Omega (1 + |p_1|^2)^{1/2} \, dx.$$

Clearly problem III(4.16) is then identical to (2.37). We have

Lemma 2.2. *$F^*(\Lambda^* p^*) = +\infty$ unless*

(2.38) $$\operatorname{div} p_1^* = 0, \quad \textit{and} \quad p_0^* + p_1^*, \nu \,|_\Gamma = 0,$$

in which case, $F^*(\Lambda p^*) = 0$.

Proof. It is easily seen that $F^*(u^*) = 0$ if $u^* = 0$, $+\infty$ otherwise. It is thus sufficient to interpret the condition $\Lambda^* p^* = 0$ which amounts to:

$$\langle p^*, \Lambda u \rangle = \int_\Gamma p_0^* \gamma_0 u \, d\Gamma + \int_\Omega p_1^* \operatorname{grad} u \, dx = 0, \quad \forall u \in W^{1,1}(\Omega).$$

We easily obtain (2.38), using Lions and Magenes [1] which provides a meaning for $p^* \nu|_\Gamma$, ν = the outward normal to Γ. ■

Lemma 2.3.

(2.39) $$G^*(p^*) = G_0^*(p_0^*) + G_1^*(p_1^*, \ldots, p_n^*),$$

(2.40) $$G_0^*(p_0^*) = \begin{cases} \displaystyle\int_\Gamma p_0^* \phi \, d\Gamma & \textit{if } |p_0^*(x)| \leqslant 1 \quad \textit{a.e.} \quad x \in \Gamma, \\ +\infty & \textit{otherwise.} \end{cases}$$

(2.41) $$G_1^*(p_1^*) = \begin{cases} \displaystyle -\int_\Omega (1 - |p_1^*|^2)^{1/2} \, dx & \textit{if } |p_1^*(x)|^2 \leqslant 1 \quad \textit{a.e.}, \\ +\infty & \textit{otherwise} \end{cases}$$

Proof. For (2.39) we use III(4.27) and III(4.28); (2.41) is identical to Lemma 1.1. Finally for (2.40),

$$\begin{aligned} G_0^*(p_0^*) &= \operatorname*{Sup}_{p_0 \in L^1(\Gamma)} \int_\Gamma [p_0^* p_0 - |p_0 - \phi|] \, d\Gamma \\ &= \int_\Gamma p_0^* \phi + \operatorname*{Sup}_{q_0 \in L^1(\Gamma)} \int_\Gamma [p_0^* q_0 - |q_0|] \, d\Gamma, \end{aligned}$$

and we obtain the stated result. ■

The dual problem of (2.37),

$$\operatorname*{Sup}_{q^* \in Y^*} [- F^*(\Lambda^* q^*) - G^*(- q^*)], \tag{2.42}$$

is thus

$$\operatorname{Sup} \left[\int_\Gamma q_0^* \phi \, d\Gamma - \int_\Omega (1 - |q_1^*|^2)^{1/2} \, dx \right], \tag{2.43}$$

the supremum being taken on the set of $p^* \in L^\infty(\Gamma) \times L^\infty(\Omega)^n$ which satisfy

$$\operatorname{div} q_1^* = 0, \tag{2.44}$$

$$q_0^* + q_1^* \nu |_\Gamma = 0, \tag{2.45}$$

$$|q_1^*(x)| \leqslant 1, \quad \text{a.e.} \quad x \in \Omega, \tag{2.46}$$

$$|q_0^*(x)| \leqslant 1, \quad \text{a.e.} \quad x \in \Gamma. \tag{2.47}$$

Using Proposition III.4.1, we ascertain that (2.43) possesses a solution and by strict concavity this solution, denoted by q^*, is unique; furthermore

$$\inf [(2.25)] = \operatorname{Sup} [(2.31)]. \tag{2.48}$$

If problem (2.37) possesses a solution $\bar{v}$, we can apply Proposition III.4.2 and Remark III.4.2 to obtain the extremality relations:

$$G_0^*(- \bar{q}_0^*) + G_0(\gamma_0 \bar{v}) = - \langle \bar{q}_0^*, \gamma_0 \bar{v} \rangle, \tag{2.49}$$

$$G_1^*(- \bar{q}_1^*) + G_1(\operatorname{grad} \bar{v}) = - \int_\Omega \bar{q}_1^* \operatorname{grad} \bar{v} \, dx. \tag{2.50}$$

These entail respectively:

$$\bar{q}_0^*(x)(\gamma_0 \bar{v}(x) - \phi(x)) = |\gamma_0 \bar{v}(x) - \phi(x)|, \quad \text{a.e.} \quad x \in \Gamma, \tag{2.51}$$

$$\operatorname{grad} \bar{v}(x) = - \frac{\bar{q}_1^*(x)}{[1 - |\bar{q}_1^*(x)|^2}, \quad \text{a.e.} \quad x \in \Omega. \tag{2.52}$$

These relations take on a special interest with the following lemma:

Lemma 2.4. *Problems* (1.13) *and* (2.43) *are identical. Their respective solutions $\bar{p}$ and $\bar{q}$ are linked by*

$$\bar{q}_0^* = - \bar{p}^* \nu |_\Gamma, \tag{2.53}$$

$$\bar{q}_1^* = \bar{p}^*. \tag{2.54}$$

Proof. If we set $p^* = q_1^*$, problem (2.43) can be expressed only in terms of p^* (from 2.45) and it is then identical to (1.13), whence the result; (2.47) follows from (2.46) and (2.45) according to a remark in footnote (1) p. 134.

Proposition 2.4. (i) *Problems* (1.3) *and* (2.37) *have the same infimum and if* (1.3) *possesses a solution* $\bar{u}$ *then* $\bar{u}$ *is a solution of* (2.37).

(ii) *In all cases* (2.37) *possesses a solution* $\bar{v}$ *which is equal to within a constant to the generalized solution* u *of* (1.3).

Proof. It follows easily from (2.2), (2.48) and lemma 2.4 that the infimum are the same and (ii) is thus proved.

For (ii) we consider a minimizing sequence u_m of (2.37)

$$\mathscr{E}(u_m) = \inf(2.37) + \rho_m,$$

where $\mathscr{E}$ denotes the functional in (2.37), $\rho_m > 0$ and $\rho_m \to 0$ as $m \to \infty$.

We deduce from Proposition 1.2 and Lemma 2.4 that

$$-\bar{q}_1^* = \bar{p}^* \in \partial_{\rho_m} G_1(\operatorname{grad} v_m).$$

Using the same technique as in Lemma 1.4, we see that, for $m \to \infty$,

$$\left|\begin{array}{l} u_m \to u \quad \text{in} \quad L^1(\Omega) \\ \dfrac{\partial u_m}{\partial x_i} \to \dfrac{\partial u}{\partial x_i} \quad \text{in} \quad L^1(\mathcal{O}), \qquad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega, \quad 1 \leqslant i \leqslant n, \end{array}\right. \tag{2.55}$$

where u is one of the generalized solutions of (1.3) given, within a constant, by Theorem 1.2.

We then prove that u is solution of (2.37) by establishing some kind of lower semicontinuity property of $\mathscr{E}$.

Let Ω_0 denote an open ball containing $\bar{\Omega}$, and let $\Omega' = \Omega_0 - \bar{\Omega}$. From Gagliardo [1], the function ϕ can be extended as a function in $W^{1,1}(\Omega_0)$. We then introduce the functions $\tilde{u}_m$ (resp. $\tilde{u}$) which are equal to u_m (resp. u) on Ω and ϕ and Ω'; let also

$$\mathscr{E}'(u_m) = \mathscr{E}(u_m) + \int_{\Omega'} [1 + |\operatorname{grad}\phi|^2]^{1/2}\, \mathrm{d}x$$
$$= \int [1 + |\operatorname{grad}\tilde{u}_m|^2]^{1/2}\, \mathrm{d}x + \int_\Gamma |u_m - \phi|\, \mathrm{d}\Gamma.$$

Now

$$\mathscr{E}'(u_m) = \operatorname{Sup}\left\{\int \left[\theta_0 + \sum_{i=1}^n \frac{\partial \theta_i}{\partial x_i}\,\tilde{u}_m\right] \mathrm{d}x\right\} \tag{2.56}$$

where the supremum is taken among the $\theta = \{\theta_0, \ldots, \theta_n\}$ of $\mathscr{D}(\Omega_0)^{n+1}$, such that

$$\sum_{i=0}^n \theta_\alpha^2 \leqslant 1.$$

Indeed using the Green formula we see that the right-hand side of (2.56) is equal to the supremum of

$$\int_{\Omega_0}\left(\theta_0+\sum_{i=1}^{m}\theta_i\frac{\partial\tilde{u}_m}{\partial x_i}\right)\mathrm{d}x-\int_{\Gamma}\left(\sum_{i=1}^{n}\theta_i\nu_i\right)(\gamma_0 u_m-\phi)\,\mathrm{d}\Gamma.$$

The supremums are essentially independent in Ω, Ω' and on Γ; whence (2.56).

Now, because of (2.55) we see that

$$\tilde{u}_m\to\tilde{u}\quad\text{in}\quad L^1(\Omega_0),\quad\text{as}\quad m\to\infty$$

and we can pass to the lower limit for the right-hand side of (2.56); we get

$$\begin{aligned}\mathscr{E}'(\tilde{u})&=\operatorname*{Sup}_{\theta}\left\{\int_{\Omega_0}\left[\theta_0+\sum_{i=1}^{n}\frac{\partial\theta_i}{\partial x_i}\tilde{u}\right]\mathrm{d}x\right\}\\&\leqslant\lim\mathscr{E}'(\tilde{u}_m).\end{aligned}$$

Whence $\mathscr{E}(u)\leqslant\inf(2.37)$ and u is a solution of (2.37). ■

The following is an easy consequence of Proposition 2.4 and the extremality relation (2.51).

Corollary 2.1. *One of the functions u given by Theorem 2.1 satisfies*

$$\begin{aligned}&\gamma_0 u\leqslant\phi\quad\text{on}\quad\Gamma_+=\{x\in\Gamma,\bar{p}^*\nu=-1\}\\&\gamma_0 u\geqslant\phi\quad\text{on}\quad\Gamma_-=\{x\in\Gamma,\bar{p}^*\nu=+1\}\\&\gamma_0 u=\phi\quad\text{on}\quad\Gamma_0=\{x\in\Gamma,-1<\bar{p}^*\nu<1\}.\end{aligned}$$

Remark 2.8. (i) A direct proof of the existence of a solution to (2.37) is due to Giusti, de Giorgi and Miranda.

(ii) We do not know if the solution of (2.37) is unique. Under the assumptions of Theorem 2.2, the solution $\bar{v}$ of (2.37) is unique and equal to u. The uniqueness is not known in the case where the boundary of Ω has everywhere a strictly negative mean curvature. If the solution is not unique, two different solutions u and $u+c$ will satisfy

$$\int_{\Gamma}|u-\phi|\,\mathrm{d}\Gamma=\int_{\Gamma}|u+c\ -\phi|\,\mathrm{d}\Gamma.\quad ■$$

Extension of this formulation

The introduction of problem (2.37) by Giusti, De Giorgi and Miranda was based on geometric considerations: the functional in (2.37) represents the area of the graph of u augmented by the area of the part of the vertical cylinder of section Γ, and limited by the graph of $u(x)$ and $\phi(x)$, $x\in\Gamma$.

Using the duality techniques and in particular the arguments of Lemma 2.4 and Proposition 2.4 we will show that besides (2.37) there are many similar variational problems which play the same role for (1.3).

Let $g_0 \in \Gamma_0(\mathbf{R})$ be an even continuous function with $g_0(0) = 0$, and consider the problem

$$\text{(2.57)} \qquad \operatorname*{Inf}_{u \in W^{1,1}(\Omega)} \left\{ \int_\Omega [1 + |\operatorname{grad} u|^2]^{1/2} \, \mathrm{d}x + \int_\Gamma g_0(u - \phi) \, \mathrm{d}\Gamma \right\}.$$

In order to compare (2.57) and (1.3) we dualize (2.57) using the same setting as for (2.37). The only difference will be in the definition of G_0; here we set

$$G_0(p_0) = \int_\Gamma g_0(u - \phi) \, \mathrm{d}\Gamma.$$

Its conjugate function G_0^* is

$$G_0^*(p_0^*) = \int_\Gamma g_0^*(p_0^*) \, \mathrm{d}\Gamma + \int_\Gamma p_0^* \phi \, \mathrm{d}\Gamma$$

where $g_0 \in \Gamma_0(\mathbf{R})$ is the conjugate of g_0.

Using Lemma 2.3 we can give the dual problem of (2.57). It is:

$$\text{(2.58)} \qquad \operatorname{Sup} \left\{ + \int_\Gamma q_0^* \phi \, \mathrm{d}\Gamma - \int_\Gamma g_0^*(-q_0^*) \, \mathrm{d}\Gamma + \int_\Omega [1 - |q_1^*|^2]^{1/2} \, \mathrm{d}x \right\}$$

the supremum being taken among those $q^* = (q_0^*, q_1^*) \in L^\infty(\Gamma) \times L^\infty(\Omega)^n$, which satisfy

$$\text{(2.59)} \qquad \operatorname{div} q_1^* = 0$$

$$\text{(2.60)} \qquad q_0^* + q_1^* . \nu = 0$$

$$\text{(2.61)} \qquad |q_1^*(x)| \leqslant 1, \quad \text{a.e.} \quad x \in \Omega.$$

This problem possesses a unique solution $\bar{q}^*$ ($\bar{q}_1^*$ is unique by strict convexity, and $\bar{q}_0^*$ because of (2.60)) and

$$\text{(2.62)} \qquad \operatorname{Inf}[(2.57)] = \operatorname{Sup}[(2.58)].$$

If $g_0^*(s) = 0$, for $s \in [-1, +1]$, then we will be in a position to repeat the proof of Lemma 2.4 and establish that the unique solution $\bar{q}^*$ of (2.58) is linked to $\bar{p}^*$ (solution of (2.31)) by

$$q_0^* = -\bar{p}^* \nu / \Gamma, \qquad \bar{q}_1^* = \bar{p}^*.$$

Thus problems (1.3) and (2.57) have the same infimum, if (1.3) possesses a solution $\bar{u}$, then $\bar{u}$ is solution of (2.57) and if (2.57) possesses a solution $\bar{v}$ this solution is equal to within a constant to the generalized solution u of (1.3).

In conclusion, if $g_0 \in \Gamma_0(\mathbf{R})$ is any even continuous function, with $g_0(0) = 0$, and such that $g_0^*(s) = 0$ for $|s| \leqslant 1$, the problem (2.57) plays exactly the same role as (2.37) (except for the existence of solution).

Remark 2.9. In particular we may choose $g_0(s) = \alpha|s|$, $\alpha > 1$, and we find

$$\text{(2.63)} \qquad \operatorname*{Inf}_{u \in W^{1,1}(\Omega)} \left\{ \int_\Omega [1 + |\operatorname{grad} u|^2]^{1/2} \, \mathrm{d}x + \alpha \int_\Gamma |u - \phi| \, \mathrm{d}\Gamma \right\}.$$

For each α, $\alpha \geqslant 1$, the solution of (2.63) is $u(+c)$. As $\alpha \to \infty$, (2.63) appears as a penalized form of (1.3). It is interesting to note that the penalized problems all have the same solution for $\alpha \geqslant 1$. ■

Remark 2.10. We can also apply the techniques of duality to other problems which are similar to the minimal hypersurface problem: hypersurfaces with obstacles (*cf.* D. H. Kinderlehrer [2] [3] [4], M. Miranda [5] [6] [7] [8], J. C. C. Nitsche [1] [5], G. Stampacchia [4]), hypersurfaces with given mean curvature (*cf.* M. Miranda [8], P. P. Mossolov [1], E. Giusti [1]), capillary problems (*cf.* P. Concus and R. Finn [1] [2], M. Emmer [1]).

For the numerical approximation of minimal hypersurfaces, see C. Jouron [1].

3. GENERALIZED SOLUTION OF CERTAIN PROBLEMS OF MINIMAL HYPERSURFACE TYPE

Orientation

Using the techniques of duality, we shall develop, as in Sections 1 and 2, a concept of generalized solution for a class of problems of minimal hypersurface type.

3.1. The primal problem and the dual problem

Let $g = g(x, u, \xi)$, $u \in \mathbf{R}$, $\xi \in \mathbf{R}^n$, be a real function, three times continuously differentiable on $\bar{\Omega} \times \mathbf{R}^{n+1}$, and let us assume that

$$\pi = (\pi_0 \ldots, \pi_n) \mapsto g(x, \pi) \text{ is convex, } \forall x \in \Omega, \tag{3.1}$$

(3.2) $\forall p \in L^\alpha(\Omega) \times L^1(\Omega)^n$, the function $x \to g(x, \rho(x))$ is summable over Ω (α fixed, $1 \leqslant \alpha < +\infty$).

In particular, if $u \in W^{1,1}(\Omega)$, $x \to g(x, u(x), \operatorname{grad} u(x))$ is a summable function over Ω.

In this section, we are concerned with the problem

$$\operatorname*{Inf}_{\substack{u \in \varphi + W_0^{1,1}(\Omega) \\ u \in L^\alpha(\Omega)}} \int_\Omega g(x, u(x), \operatorname{grad} u(x))\, \mathrm{d}x. \tag{3.3}$$

This problem is equivalent to the Dirichlet problem associated with the Euler equation of (3.3),

$$\sum_{i=1}^{n} \frac{\mathrm{d}}{\mathrm{d}x_i} \frac{\partial g}{\partial \xi_i}(x, u, \operatorname{grad} u) = \frac{\partial g}{\partial u}(x, u, \operatorname{grad} u), \quad \text{in } \Omega, \tag{3.4}$$

$$u = \varphi \quad \text{on } \Gamma. \tag{3.5}$$

In what follows, we shall specify the hypotheses concerning G; the class of functions g under consideration contains the function

$$g(x, \pi) = \left(1 + \sum_{i=1}^{n} \pi_i^2\right)^{1/2}, \tag{3.6}$$

for which problem (3.3) is none other than (1.3).

For the moment, we shall write (3.3) as a problem of type III(4.16) and determine the dual problem.

Let $V = W^{1,1}(\Omega) \cap L^\alpha(\Omega)$, $Y = L^\alpha(\Omega) \times L^1(\Omega)^n$; we denote by $p = (p_0, p_1)$, $p_0 \in L^\alpha(\Omega)$, $p_1 \in L^1(\Omega)^n$, the elements of Y. We set:

$$V^* = \text{the dual of } V, \qquad Y^* = L^{\alpha'}(\Omega) \times L^\infty(\Omega)^n,$$

where $1/\alpha' + 1/\alpha = 1$; we denote by $p^* = (p_0^*, p_1^*)$, the elements of Y^*, $p_0^* \in L^{\alpha'}(\Omega)$, $p_1^* \in L^\infty(\Omega)^n$.

We define the operator Λ,

$$\Lambda u = \{ u, \operatorname{grad} u \}, \qquad \forall u \in V, \tag{3.7}$$

and the functions F and G,

$$F(u) = \begin{vmatrix} 0 \quad \text{if} \quad u \in \varphi + W_0^{1,1}(\Omega) \quad (u \in V), \\ + \infty \quad \text{otherwise}, \end{vmatrix}$$

$$G(p) = \int_\Omega g(x, p(x)) \, \mathrm{d}x, \qquad \forall p \in L^\alpha(\Omega) \times L^1(\Omega)^n.$$

With the above notations, the problem

$$\operatorname*{Inf}_{u \in V} [F(u) + G(\Lambda u)] \tag{3.8}$$

is in fact identical to (3.3).

The function F is convex and l.s.c. on V; its conjugate function is

$$F^*(u^*) = \operatorname*{Sup}_{\substack{u \in \varphi + W_0^{1,1}(\Omega) \\ u \in L^\alpha(\Omega)}} \langle u^*, u \rangle,$$

and for $p^* \in Y^*$,

$$F^*(\Lambda^* p^*) = \langle p^*, \Lambda\varphi \rangle + \operatorname{Sup} \langle p^*, \Lambda v \rangle, \qquad v \in W_0^{1,1}(\Omega) \cap L^\alpha(\Omega).$$

The latter supremum is equal to

$$\operatorname*{Sup}_{v \in \mathscr{D}(\Omega)} \int_\Omega (p_0^* v + p_1 \operatorname{grad} v) \, \mathrm{d}x,$$

and this supremum is 0 or $+\infty$ according to whether $p_0^* - \operatorname{div} p_1^* = 0$ or not:

$$(3.9) \qquad F^*(\Lambda^* p^*) = \begin{vmatrix} \langle p^*, \Lambda\varphi \rangle & \text{if } p_0^* - \operatorname{div} p_1^* = 0, \\ +\infty & \text{otherwise.} \end{vmatrix}$$

From (3.1) and (3.2), the function G is defined and convex on $L^\alpha(\Omega) \times L^1(\Omega)^n$, with values in $\mathbf{R}$. It follows from (3.2), from the continuity of G and from Proposition IV.1.1 that:

(3.10) the function $p \mapsto G(p)$ is continuous over $L^1(\Omega)^{n+1}$.

The conjugate functional of G, G^*, is defined over $Y^* = L^{\alpha'}(\Omega) \times L^\infty(\Omega)^n$:

$$(3.11) \qquad G^*(p^*) = \operatorname*{Sup}_{p \in L^\alpha(\Omega) \times L^1(\Omega)^n} \left| \int_\Omega [p^*(x)p(x) - g(x, p(x))] \, \mathrm{d}x \right|.$$

By virtue of Proposition IV.1.2:

$$(3.12) \qquad G^*(p^*) = \int_\Omega g^*(x, p^*(x)) \, \mathrm{d}x,$$

where g^* possibly equal to $+\infty$ at some points, is the pointwise conjugate function of g:

$$(3.13) \qquad g^*(x, \pi^*) = \operatorname*{Sup}_{\pi \in \mathbf{R}^{n+1}} \{ \pi^* . \pi - g(x, \pi) \}.$$

It is easily verified that g^* is bounded below on $\Omega \times \mathbf{R}^{n+1}$

$$(3.14) \qquad g^*(x, \pi) \geqslant - \operatorname*{Sup}_\Omega g(x, 0) > -\infty.$$

From this and by virtue of the lower semi-continuity of g^*, the integral (3.12) is completely defined in $]-\infty, +\infty]$.

Remark 3.1. From the foregoing, the relation $q^* \in \partial G(q)$ is equivalent to:

$$(3.15) \qquad g^*(x, q^*(x)) + g(x, q(x)) = q^*(x)q(x) \quad \text{a.e.},$$

and this implies that

$$(3.16) \qquad g^*(x, q^*(x)) \in \mathbf{R}, \quad \text{a.e.} \quad x \in \Omega. \; \blacksquare$$

The dual problem of (3.3) can thus be written as

$$(3.17) \qquad \operatorname*{Sup}_{\substack{p^* \in L^{\alpha'}(\Omega) \times L^\infty(\Omega)^n \\ p_0^* = \operatorname{div} p_1^*}} \left[-\int_\Omega (p_0^* \varphi + p_1^* \operatorname{grad} \varphi) \, \mathrm{d}x - \int_\Omega g^*(x - p^*(x) \, \mathrm{d}x \right].$$

Proposition 3.1. *Problems* (3.3) *and* (3.17) *are mutually dual and*

$$\inf \mathscr{P} = \operatorname{Sup} \mathscr{P}^*. \tag{3.18}$$

Problem (3.17) *possesses at least one solution* $\bar{p}^*$. *If problem* (3.3) *possesses a solution* $\bar{u}$, *then*:

$$\begin{aligned} g^*(x, -\bar{p}_0^*(x), -\bar{p}_1^*(x)) + g(x, \bar{u}(x), \operatorname{grad} \bar{u}(x)) &= \\ = -\bar{p}_0^*(x)\bar{u}(x) - \bar{p}_1^*(x) \cdot \operatorname{grad} \bar{u}(x), &\qquad \text{a.e.} \end{aligned} \tag{3.19}$$

Proof. This is an immediate consequence of Theorem III.4.1, of Proposition III.4.1 and of (3.15).

3.2. New hypotheses for g

The extension of Section 2 to problem (3.3) can be made in the case where g satisfies the hypotheses given below whose role and significance will appear later.

We assume that, for all $M > 0$, there exist constants μ_1, μ_2 ..., depending on M, such that:

$$\forall x \in \Omega, \qquad \forall u \in \mathbf{R}, \qquad |u| \leqslant M, \qquad \forall \xi \in \mathbf{R}^n,$$

$$g(x, u, \xi) \geqslant \mu_0(M)\,|\xi| - \mu_1(M), \tag{3.20}$$

$$\frac{\partial g}{\partial \xi_i}(x, u, \xi) \leqslant \mu_2(M), \qquad 1 \leqslant i \leqslant n, \tag{3.21}$$

$$\sum_{i=1}^{n} \frac{\partial g}{\partial \xi_i}(x, u, \xi)\xi_i \geqslant \mu_3(M)(1 + |\xi|^2)^{1/2} - \mu_4(M), \qquad \mu_3(M) > 0, \tag{3.22}$$

$$\left|\frac{\partial g}{\partial u}(x, u, \xi)\right| \leqslant \mu_5(M), \tag{3.23}$$

$$\begin{gathered} \mu_6(M)\frac{|\eta'|^2}{(1 + \xi^2)^{1/2}} \leqslant \sum_{i,j=1}^{n} \frac{\partial^2 g}{\partial \xi_i \, \partial \xi_j}(x, u\xi)\,\eta_i\eta_j \leqslant \mu_7(M)\frac{|\eta'|^2}{(1 + \xi^2)^{1/2}} \\ \forall \eta = (\eta_1, \ldots, \eta_n) \in \mathbf{R}^n, \qquad \mu_6(M) > 0, \end{gathered} \tag{3.24}$$

where $|\eta'|$ is defined from ξ and η by

$$|\eta'|^2 = |\eta|^2 - \frac{(\eta \cdot \xi)^2}{1 + \xi^2}. \tag{3.25}$$

Similarly, for all $M > 0$, we assume that there exists $\mu = \mu(M)$ such that,

$$\forall x \in \Omega, \qquad \forall u \in \mathbf{R}, \qquad |u| \leqslant M, \qquad \forall \xi \in \mathbf{R}^n,$$

$$(3.26) \quad \left|\frac{\partial^3 g}{\partial \xi_i \, \partial \xi_\ell \partial u}\right| \leqslant \frac{\mu(M)}{|\xi|^2}, \qquad \left|\frac{\partial^3 g}{\partial \xi_i \, \partial \xi_\ell \, \partial x_j}\right| \leqslant \frac{\mu(M)}{|\xi|},$$

$$(3.27) \quad \left|\frac{\partial^3 g}{\partial \xi_\ell \partial \xi_j \, \partial u} \xi_j\right| \leqslant \frac{\mu(M)}{|\xi|^2}, \qquad \left|\frac{\partial^3 g}{\partial \xi_\ell \, \partial \xi_i \, \partial x_j} \xi_i \xi_j\right| \leqslant \mu(M),$$

$$(3.28) \quad \left|\frac{\partial^3 g}{\partial \xi_i \, \partial u^2} \xi_i - \frac{\partial^2 g}{\partial u^2}\right| \leqslant \frac{\mu(M)}{|\xi|^2}, \qquad \left|\frac{\partial^3 g}{\partial \xi_i \, \partial u \, \partial x_\ell} \xi_i - \frac{\partial^2 g}{\partial u \, \partial x_\ell}\right| \leqslant \mu(M),$$

$$(3.29) \quad \left|\frac{\partial^3 g}{\partial \xi_i \, \partial u \, \partial x_i}\right| \leqslant \frac{\mu(M)}{|\xi|}, \qquad \left|\frac{\partial^3 g}{\partial \xi_i \, \partial x_\ell \, \partial x_i}\right| \leqslant \mu(M).$$

We also assume that there exists $M_1 > 0$ such that:

$$\forall x \in \Omega, \qquad \forall u \in \mathbf{R}, \qquad |\mu| > M_1, \qquad \forall \xi \in \mathbf{R}^n,$$

$$(3.30) \qquad \sum_{i=1}^{n} \frac{\partial g}{\partial \xi_i}(x, u, \xi) \xi_i \geqslant 0,$$

$$(3.31) \qquad \frac{\partial g}{\partial u}(x, u, \xi). \operatorname{sgn} u \geqslant 0.$$

As in Chapter I, we shall term domg^* the set

$$(3.32) \qquad \operatorname{dom} g^* = \{ (x, \pi) \in \Omega \times \mathbf{R}^{n+1} \mid g^*(x, \pi) < +\infty \}.$$

We shall assume that

(3.33) $\qquad g^*$ is continuous over dom g^*,

and to simplify matters slightly,

(3.34) $\qquad (\pi_1^*, \ldots, \pi_n^*) \mapsto g^*(x, \pi^*)$ is strictly convex, $\forall x \in \Omega$,

$\qquad \forall \pi_0^* \in \mathbf{R}$ with $(x, \pi) \in \operatorname{dom} g^*$.

This last hypothesis is a sort of continuity hypothesis for ∂G and can be stated as follows:

(3.35)

Let q_m and q_m^* be two sequences such that

(i) $q_m \in L^\alpha(\Omega) \times L^1(\Omega)^n, \qquad q_m^* \in L^{\alpha'}(\Omega) \times L^\infty(\Omega)^n$,

(ii) $q_m^* \in \partial G(q_m)$,

(iii) $q_m^* \to q^*$ in $L^\alpha(\Omega) \times L^\infty(\Omega)^n$,

where $q^* \in \partial G(q)$ and satisfies

(iv) $\operatorname{Sup}_{\mathcal{O}} |q(x)| < +\infty, \qquad \mathcal{O} \subset \Omega$ measurable.

Then for $i = 1, \ldots, n$,

$(q_m)_i \to q_i$ in $L^1(\mathcal{O})$. ∎

Remark 3.2. It is easily verified that hypotheses (3.20)–(3.25) are satisfied for the following functions g certain of which lead to standard variational problems (*cf.* J. Serrin [1]):

$(1 + |\xi|^2)^{1/2}$ (minimal hypersurfaces),
$(1 + |\xi|^2)^{1/2} - f(x)u + \lambda\,|u|, \quad \lambda > 0, \quad f \in L^\infty(\Omega)$
(*cf.* Mossolov [1] when $\lambda = 0$),
$(1 + |\xi|^2)^{1/2} - f(x)u + u^2, \quad f \in L^\infty(\Omega) \quad (\alpha = 2)$,
$(1 + |x|^2 + |\xi|^2)^{1/2}$ considered by Bernstein [1] [2],
$(1 + u^2 + |\xi|^2)^{1/2}, \quad (1 + (1 + |\xi|^2)^1)^{s/2}, \quad s \geqslant 1/2$, etc. ■

3.3. Generalized solution of problem (3.3)

Statement of the principal result

Theorem 3.1. *With the hypotheses* (3.1), (3.2) *and* (3.20) *to* (3.25), *and* ϕ *given satisfying*

$$\varphi \in \mathscr{W}^{1,1}(\Omega) \cap L^\infty(\Omega), \tag{3.36}$$

problem (3.3) *admits problem* (3.17) *as its dual and*

$$\inf \mathscr{P} = \inf \mathscr{P}^*. \tag{3.37}$$

Problem (3.17) *possesses a unique solution* $\bar{p}^*$.

Furthermore, there exists a unique function u (*unique up to within an additive constant*) *such that*:

$$u \in W^{1,1}(\Omega) \cap L^\infty(\Omega), \tag{3.38}$$

$$\operatorname*{Sup}_{\mathcal{O}} |\text{grad } u(x)| < +\infty, \qquad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega, \tag{3.39}$$

$$g^*(x, -\bar{p}^*(x)) + g(x, \Lambda u(x)) = -\bar{p}^*(x)\cdot \Lambda u(x), \quad \text{a.e.} \quad x \in \Omega, \tag{3.40}$$

(3.41) *u is the solution of* (3.4) *in* Ω, *the Euler equation of* (3.3).

(3.42) *Every bounded minimizing sequence* $\{v_m\}$ *of* (3.3) *converges to* u *in the following sense*
$v_m \to u \quad$ *in* $L^1(\Omega)/\mathbf{R}$
$\partial v_m/\partial x_i \to \partial u/\partial x_i \quad$ *in* $L^1(\mathcal{O}), \quad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega, \quad 1 \leqslant i \leqslant n.$

The theorem will be proved in the following sections.

Remark 3.3. By virtue of (3.39), (3.41) and some standard results (*cf.* E. De Giorgi [1], E. Hopf [1], O. A. Ladyzhenskaya and N. N. Uralceva [2]), the function u is as regular in Ω as g (up to the analyticity in Ω).

Remark 3.4. We can apply Remarks 2.2, 2.3 and 2.4 concerning minimal hypersurfaces to this situation.

In particular, u may or may not be the solution of (3.3) according to whether $u = \phi$ on Γ or not. When u is not the solution of (3.3) we say, by virtue of (3.42), that *u is a generalized solution of this problem.*

Research into *a priori* conditions on Ω, g, ϕ, which guarantee that $u = \phi$ on Γ or on a subset of Γ remains an open problem.

Following J. Serrin [1], we can obtain analogous results to Proposition 2.1 for certain functions g and certain open spaces Ω.

We also have an *a posteriori* result analogous to Proposition 2.2, which we shall state below and prove in Section 3.8. ∎

Proposition 3.2. *With the hypotheses of Theorem* 3.1 *and also the assumption that one of the functions defined by Theorem* 3.1 *satisfies*

$$(3.43)\qquad \left|\begin{array}{l} \exists\, x_0 \in \Gamma \ \textit{such that} \\ \overline{\lim\limits_{\substack{x\to x_0\\ x\in\Omega}}}\, |\operatorname{grad} u(x)| < +\infty. \end{array}\right.$$

Then there exists a unique function u which satisfies all the conclusions of Theorem 3.1 *and, furthermore,*

$$(3.44)\qquad \left|\begin{array}{l} u = \phi \textit{ on a part of } \Gamma \textit{ with non-zero measure and more precisely on} \\ \{\, x \in \Gamma \mid \overline{\lim\limits_{\substack{y\to x\\ y\in\Omega}}}\, |\operatorname{grad} u(y)| < +\infty \,\}. \end{array}\right.$$

$$(3.45)\qquad \left|\begin{array}{l} \textit{Every bounded minimizing sequence } \{v_m\} \textit{ of } (3.3) \textit{ converges to } u \\ \textit{in the following sense:} \\ v_m \to u \ \textit{in} \ L^1(\Omega) \\ \partial v_m/\partial x_i \to \partial u/\partial x_i \ \textit{in} \ L^1(\mathcal{O}), \qquad \forall\, \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega. \end{array}\right.$$

3.4. Construction of a regular minimizing sequence of (3.3)

We assume, until Section 3.7, that

$$(3.46)\qquad \varphi \in H^1(\Omega) \cap L^\infty(\Omega),$$

and we shall first prove Theorem 3.1 with this supplementary hypothesis.

We consider, for fixed $\varepsilon > 0$, the function $u_\varepsilon \in H^1(\Omega)$ which is the solution of the problem[1]

$$(3.47) \qquad \operatorname*{Inf}_{\substack{v \in \phi + H_0^1(\Omega) \\ v \in L(\Omega)}} \left\{ \int_\Omega \left[\frac{\varepsilon}{2} |\text{grad } v(x)|^2 + g(x, v(x), \text{grad } v(x)) \right] \mathrm{d}x \right\}.$$

The existence and uniqueness of a solution $u_\varepsilon \in L^\infty(\Omega)$ follows from Ladyzhenskaya and Uralceva [1], J. L. Lions [3]. From the same authors, we have $u_\varepsilon \in \mathscr{C}^2(\Omega)$ at least, and u_ε is solution of the Euler equation of (3.47):

$$(3.48) \qquad \varepsilon \Delta u_\varepsilon + \sum_{i=1}^n \frac{\mathrm{d}}{\mathrm{d}x_i} \frac{\partial g}{\partial \xi_i}(x, u_\varepsilon, \text{grad } u_\varepsilon) = \frac{\partial g}{\partial u}(x, u_\varepsilon, \text{grad } u_\varepsilon).$$

To begin with, we prove

Lemma 3.1. *u_ε is a minimizing sequence of* (3.3).

Proof. If $v \in \phi + \mathscr{D}(\Omega)$, we have, $\mathscr{P}$ being the problem (3.3):

$$\inf \mathscr{P} \leqslant \int_\Omega g(x, u_\varepsilon(x), \text{grad } u_\varepsilon(x))\, \mathrm{d}x$$

$$\leqslant \int_\Omega \left[\frac{\varepsilon}{2} |\text{grad } u_\varepsilon(x)|^2 + g(x, u_\varepsilon(x), \text{grad } u_\varepsilon(x)) \right] \mathrm{d}x$$

(by definition of u_ε)

$$\leqslant \int_\Omega \left[\frac{\varepsilon}{2} |\text{grad } v(x)|^2 + g(x, v(x), \text{grad } v(x)) \right] \mathrm{d}x.$$

When $\varepsilon \to 0$, we obtain:

$$\inf \mathscr{P} \leqslant \varliminf_{\varepsilon \to 0} \int_\Omega g(x, u_\varepsilon, \text{grad } u_\varepsilon)\, \mathrm{d}x$$

$$\leqslant \varlimsup_{\varepsilon \to 0} \int_\Omega g(x, u_\varepsilon, \text{grad } u_\varepsilon)\, \mathrm{d}x$$

$$\leqslant \int_\Omega g(x, v, \text{grad } v)\, \mathrm{d}x.$$

[1] This is a regularized elliptic version of problem (3.3). For elliptic regularization in other contexts see Baouendi [1], Kohn and Niremberg [1], Lions [3].

Since $\mathscr{D}(\Omega)$ is dense in $W_0^{1,1}(\Omega) \cap L^\alpha(\Omega)$, by taking the infimum for the v under consideration we obtain,

$$\lim_{\varepsilon \to 0} \int_\Omega g(x, u_\varepsilon, \operatorname{grad} u_\varepsilon)\, dx = \inf \mathscr{P},$$

and the lemma follows.

Lemma 3.2. *The sequence u_ε is bounded in $L^\infty(\Omega)$ and in $W^{1,1}(\Omega)$.*

Proof. By definition of u_ε,

$$\int_\Omega \left[\frac{\varepsilon}{2}|\operatorname{grad} u_\varepsilon|^2 + g(x, u_\varepsilon, \operatorname{grad} u_\varepsilon)\right] dx \leqslant \tag{3.49}$$

$$\leqslant \int_\Omega \left[\frac{\varepsilon}{2}|\operatorname{grad} \varphi|^2 + g(x, \varphi, \operatorname{grad} \varphi)\right] dx.$$

Because of (3.20) we deduce that:

$$\sqrt{\varepsilon}\, \|\operatorname{grad} u_\varepsilon\|_{L^2(\Omega)^n} \leqslant c, \tag{3.50}$$

$$\|\operatorname{grad} u_\varepsilon\|_{L^1(\Omega)^n} \leqslant c, \tag{3.51}$$

which with Poincare's inequality ($u_\varepsilon - \phi \in H_0^1(\Omega)$) implies that

$$\sqrt{\varepsilon}\, \|u_\varepsilon\|_{H^1(\Omega)} \leqslant c, \tag{3.52}$$

$$\|u_\varepsilon\|_{W^{1,1}(\Omega)} \leqslant c. \tag{3.53}$$

On the other hand, let

$$M_0 = \max(\|\varphi\|_{L^\infty(\Omega)}, M_1) \tag{3.54}$$

where M_1 is the constant introduced in (3.30) and (3.31). Applying the maximum principle to (3.48) we verify that

$$\|u_\varepsilon\|_{L^\infty(\Omega)} \leqslant M_0. \tag{3.55}$$

To show in fact that:

$$u_\varepsilon(x) \leqslant M_0 \quad \text{a.e.,} \tag{3.56}$$

we multiply (3.48) by $(u_\varepsilon - M_0)_+$, we integrate and, using Green's formula, we find that[1]

$$\int_{\Omega\cap\{u_\varepsilon > M_0\}} \sum_{i=1}^{n} \left[\varepsilon \frac{\partial u_\varepsilon}{\partial x_i} + \frac{\partial g}{\partial \xi_i}(x, u_\varepsilon, \operatorname{grad} u_\varepsilon)\right] \frac{\partial u_\varepsilon}{\partial x_i}\,dx =$$

$$= -\int_{\Omega\cap\{u_\varepsilon > M_0\}} \frac{\partial g}{\partial u}(x, u_\varepsilon, \operatorname{grad} u_\varepsilon)(u_\varepsilon - M_0)\,dx \leqslant 0 \text{ (by (3.31))}.$$

This implies, with (3.30), that $\partial u_\varepsilon/\partial x_i = 0$ almost everywhere on the set

$$\{\, x \mid u_\varepsilon(x) > M_0 \,\}, \quad 1 \leqslant i \leqslant n,$$

and (3.56) is proved.

Lemma 3.3. *For every open set $\mathcal{O}$ relatively compact in Ω*

$$(3.57) \qquad \|u_\varepsilon\|_{W^{1,\infty}(\mathcal{O})} \leqslant c(\mathcal{O}, \Omega, M_0),$$

$$(3.58) \qquad \|u_\varepsilon\|_{H^2(\mathcal{O})} \leqslant c(\mathcal{O}, \Omega, M_0),$$

where these constants only depend on $\mathcal{O}$, Ω and M_0.

Proof. This is a direct consequence of a perturbation theorem established in R. Temam [3] (*cf.* Theor. 1.1) and which can be stated as follows:

Theorem of singular perturbation. *We assume that $u_\varepsilon \in \mathscr{C}^2(\Omega)$ verifies* (3.48) *and* (3.55), *whereas $g \in \mathscr{C}^2(\bar\Omega \times \mathbf{R}^{n+1})$ satisfies* (3.21) *to* (3.29).[2]

Then for every open set $\mathcal{O}$ relatively compact in Ω, we have (3.57) *and* (3.58).

Passage to the limit, $\varepsilon \to 0$

We can specify the behaviour of u_ε, for $\varepsilon \to 0$. With Lemmas 3.2 and 3.3 and by using the diagonalization process, we can extract from u_ε a sequence denoted by u_{ε_m} such that:

$$(3.59) \qquad u_{\varepsilon_m} \to u \text{ in } L^\infty(\Omega)\text{—weak*},$$

$$(3.60) \qquad \partial u_{\varepsilon_m}/\partial x_i \to \partial u/\partial x_i \text{ in } L^\infty(\mathcal{O})\text{—weak*}, \qquad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega, \qquad 1 \leqslant i \leqslant n,$$

$$(3.61) \qquad u_{\varepsilon_m} \to u \text{ weakly in } H^2(\mathcal{O}), \qquad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega.$$

[1] This is permissible, since $(u_\varepsilon - M)_+ \in H_0^1(\Omega)$, *cf.* Stampacchia [1].
[2] This is the reason for introducing these hypotheses.

This function $u \in L^\infty(\Omega)$ and satisfies (3.39).

Because of (3.53), since the embedding of $W^{1,1}(\Omega)$ into $L^1(\Omega)$ is compact, we can choose u_{ε_m} so that

$$(3.62) \qquad u_{\varepsilon_m} \to u \text{ strongly in } L^1(\Omega).$$

Similarly, with (3.58) and the compactness theorems in Sobolev spaces (Lions [2], Lions and Magenes [1]),

$$(3.63) \qquad u_{\varepsilon_m}|_{\mathcal{O}} \to u|_{\mathcal{O}} \quad \text{strongly in} \quad H^1(\mathcal{O}), \qquad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega.$$

Because of (3.63) and again using the diagonalization process, we can extract from u_{ε_m} a subsequence (again denoted by u_{ε_m}) such that

$$(3.64) \qquad \left| \begin{array}{l} u_{\varepsilon_m}(x) \to u(x) \quad \text{a.e.,} \\ \dfrac{\partial u_{\varepsilon_m}}{\partial x_i}(x) \to \dfrac{\partial u}{\partial x_i}(x), \qquad \text{a.e.,} \qquad i = 1, \ldots, n. \end{array} \right.$$

By virtue of (3.51), (3.63) and Fatou's lemma,

$$\frac{\partial u}{\partial x_i} \in L^1(\Omega), \qquad i = 1, \ldots, n,$$

and since $u \in L^\infty(\Omega)$, property (3.38) is satisfied.

We now pass to the limit in (3.48). By virtue of (3.55), (3.21) and (3.23), the functions

$$\frac{\partial g}{\partial \xi_i}(x, u_\varepsilon, \operatorname{grad} u_\varepsilon), \qquad \frac{\partial g}{\partial u}(x, u_\varepsilon, \operatorname{grad} u_\varepsilon),$$

are uniformly bounded on Ω, independently of ε. Since $\partial g/\partial \xi_i$, $\partial g/\partial u$ are continuous, (3.62) and (3.64) imply that

$$\frac{\partial g}{\partial \xi_i}(x, u_\varepsilon(x), \operatorname{grad} u_\varepsilon(x)) \to \frac{\partial g}{\partial \xi_i}(x, u(x), \operatorname{grad} u(x)), \quad \text{a.e.}$$

$$\frac{\partial g}{\partial \xi_i}(x, u_\varepsilon(x), \operatorname{grad} u_\varepsilon(x)) \to \frac{\partial g}{\partial \xi_i}(x, u(x), \operatorname{grad} u(x)), \quad \text{a.e.}$$

$$\frac{\partial g}{\partial u}(x, u_\varepsilon(x), \operatorname{grad} u_\varepsilon(x)) \to \frac{\partial g}{\partial u}(x, u(x), \operatorname{grad} u(x)), \quad \text{a.e.}$$

With the dominated convergence theorem, this implies that:

$$\frac{\partial g}{\partial \xi_i}(x, u_\varepsilon, \operatorname{grad} u_\varepsilon) \to \frac{\partial g}{\partial \xi_i}(x, u, \operatorname{grad} u),$$

$$\frac{\partial g}{\partial u}(x, u_\varepsilon, \operatorname{grad} u_\varepsilon) \to \frac{\partial g}{\partial u}(x, u, \operatorname{grad} u), \quad \text{in } L^1(\Omega).$$

Passing to the limit in (3.48), we now see that u satisfies the equation (3.4) and (3.41) is proved.

3.5. Regularity of the solution $\bar{p}^*$ of (3.17)

From Proposition 3.1, problem (3.17) possesses at least one solution.

By virtue of the supplementary hypothesis (3.34), if $\bar{p}^* = (\bar{p}_0^*, \bar{p}_1^*)$ and $\bar{q}^* = (\bar{q}_0^*, \bar{q}_1^*)$ are two solutions of (3.17), we have:

$$\bar{p}_1^* = \bar{q}_1^*.$$

But as $\bar{p}_0^* = \operatorname{div} \bar{p}_1^* = \operatorname{div} \bar{q}_1^* = \bar{q}_0^*$, we have $\bar{p}^* = \bar{q}^*$ which demonstrates the uniqueness of the solution of (3.17).[1] This solution will be denoted by $\bar{p}^*$.

At this point a study of the regularity of $\bar{p}^*$ will enable us to prove (3.40) which implies (for Section 3.4) a *characterization of u independent of the subsequence u_{ε_m} extracted from u_ε.*

From Lemma 3.1, u_ε is a minimizing sequence of (3.3); hence:

$$F(u_\varepsilon) + G(\Lambda u_\varepsilon) = G(\Lambda u_\varepsilon) = \inf \mathscr{P} + \rho_\varepsilon,$$
$$\rho_\varepsilon \geqslant 0, \quad \rho_\varepsilon \to 0 \quad \text{with } \varepsilon.$$

We can make use of Proposition 1.2:

$$- p_\varepsilon^* \in \partial \rho_\varepsilon \quad G(\Lambda u_\varepsilon). \tag{3.65}$$

Theorem I.6.2 then enables us to affirm the existence of p_ε and p_ε^* such that:

$$p_\varepsilon \in Y = L^\alpha(\Omega) \times L^1(\Omega)^n, \tag{3.66}$$

$$p_\varepsilon^* \in Y^* = L^{\alpha'}(\Omega) \times L^\infty(\Omega)^n, \tag{3.67}$$

$$\|p_\varepsilon - \Lambda u_\varepsilon\|_Y \leqslant \sqrt{\rho_\varepsilon}, \tag{3.68}$$

$$\|p_\varepsilon^* - \bar{p}^*\|_{Y^*} \leqslant \sqrt{\rho_\varepsilon}, \tag{3.69}$$

$$- p_\varepsilon^* \in \partial G(p_\varepsilon). \tag{3.70}$$

From (3.15) (Remark 3.1), (3.70) means that:

$$g^*(x, - p_\varepsilon^*(x)) + g(x, p_\varepsilon(x)) = - p_\varepsilon^*(x) \cdot p_\varepsilon(x), \quad \text{a.e.,} \tag{3.71}$$

and implies that:

$$g^*(x, - p_\varepsilon^*(x)) \in \operatorname{dom} g^*, \quad \text{a.e.} \quad x \in \Omega. \tag{3.72}$$

[1] We have not used (3.46) here.

By virtue of (3.68), we can, by extracting a new subsequence, assume that:

(3.73) $$p_{\varepsilon_m}(x) - \Lambda u_{\varepsilon_m}(x) \to 0 \quad \text{a.e.},$$

and, with (3.64),

(3.74) $$p_{\varepsilon_m}(x) \to \Lambda u(x) \quad \text{a.e.}$$

Obviously, by (3.69),

(3.75) $$p^*_{\varepsilon_m}(x) \to \bar{p}^*(x) \quad \text{a.e.}$$

Passing to the lower limit in (3.71), we obtain:

$$g^*(x, -\bar{p}^*(x)) \leqslant -\bar{p}^*(x).\,\Lambda u(x) - g(x, \Lambda u(x)) < +\infty, \quad \text{a.e.},$$

so that $(x, -\bar{p}^*(x)) \in \operatorname{dom} g^*$ for almost all $x \in \Omega$.

Hence, because of (3.33) we can pass to the limit in (3.71) and deduce the equality (3.40) from it.

3.6. Property of minimizing sequences of (3.3)

We shall now prove (3.42).

Let $\{v_m\}$ be a bounded minimizing sequence of (3.3):[1]

(3.76) $$\left|\begin{array}{ll} v_m \in \varphi + W_0^{1,1}(\Omega), & v_m \in L^\infty(\Omega), \quad \forall m \\ \|v_m\|_{L^\infty(\Omega)} \leqslant c, & \\ G(\Lambda v_m) \to \inf \mathscr{P}, & m \to \infty. \end{array}\right.$$

From this we can deduce with (3.20) that the sequence v_m is also bounded in $W^{1,1}(\Omega)$.

We set

$$G(\Lambda v_m) = \inf \mathscr{P} + \sigma_m, \quad \sigma_m \geqslant 0,$$

$\sigma_m \to 0$ for $m \to \infty$. From Proposition 1.2 we have again

$$-\bar{p}^* \in \partial_{\sigma_m} G(\Lambda v_m).$$

Using Theorem I.6.2, we obtain for all m, the existence of p_m and p_m^*, such that:

(3.77) $$p_m \in L^\alpha(\Omega) \times L^1(\Omega)^n, \qquad p_m^* \in L^{\alpha'}(\Omega) \times L^\infty(\Omega)^n,$$

(3.78) $$\|p_m - \Lambda v_m\|_Y \leqslant \sqrt{\sigma_m},$$

[1] In particular v_m can be the sequence $\{u_\varepsilon\}$ of Section 3.4.

(3.79) $$\|p_m^* - \bar{p}^*\|_{Y^*} \leqslant \sqrt{\sigma_m},$$

(3.80) $$p_m^* \in \partial G(p_m).$$

From (3.79), $p_m^* \to \bar{p}^*$ in Y^*, and from (3.19) and (3.40) $-\bar{p}^* \in \partial G(\Lambda u)$, with Λu satisfying (*cf.* (3.38) and (3.39)):

(3.81) $$\operatorname*{Sup}_{\mathcal{O}} |\Lambda u(x)| < +\infty, \qquad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega.$$

We can thus apply (3.35) and deduce that,

$$(p_m)_i \to (\Lambda u)_i \quad \text{in} \quad L^1_{\text{loc}}(\Omega), \qquad 1 \leqslant i \leqslant n.$$

But then from (3.78) we have

$$(\Lambda v_m)_i \to (\Lambda v)_i \quad \text{in} \quad L^1_{\text{loc}}(\Omega),$$

which implies that

$$\partial v_m/\partial x_i \to \partial v/\partial x_i \quad \text{in} \quad L^1_{\text{loc}}(\Omega), \qquad 1 \leqslant i \leqslant n.$$

Now, since v_m is bounded in $W^{1,1}(\Omega)$ and in $L^\infty(\Omega)$, we can extract a subsequence v_{m_j} such that:

(3.82) $$v_{m_j} \to \Psi \quad \text{in} \quad L^\infty(\Omega)\text{—weak* and strongly in } L^1(\Omega).$$

Necessarily, in each convex component of Ω we have $\Psi = u + \text{constant}$, and if we pass to the quotient by the constants, the limit in (3.82) is independent of the subsequence and the entire sequence v_m converges to Ψ; there exists a bounded sequence of numbers λ_m, such that:

$$v_m + \lambda_m \to u \text{ in } L^\infty(\Omega)\text{—weak*, and strongly in } L^1(\Omega),\ m \to \infty;$$

(3.42) is established.

Remark 3.5. Certain *a priori* upper bounds for u result from the foregoing:

(3.83) $$\|u\|_{L^\infty(\Omega)} \leqslant M_0 = \max\,(\|\varphi\|_{L^\infty(\Omega)}, M_1)$$

(by (3.54), (3.55) and (3.59)).

(3.84) $$\|\text{grad}\, u\|_{L^1(\Omega)^n} \leqslant c(M_0)\left(1 + \int_\Omega g(x, \varphi(x), \text{grad}\, \varphi(x)\, \mathrm{d}x\right)$$

(by (3.20), (3.49), (3.64), (3.83) and Fatou's Lemma).

(3.85) $$\|u\|_{W^{1,\infty}(\mathcal{O})} \leqslant c(\mathcal{O}, \Omega, M_0), \qquad \forall \mathcal{O} \subset \mathcal{O} \subset \Omega,$$

$$(3.86) \qquad \|u\|_{H^2(\mathcal{O})} \leqslant c(\mathcal{O}, \Omega, M_0), \qquad \forall \mathcal{O} \subset \mathcal{O} \subset \Omega,$$

(by (3.57), (3.58), (3.60), (3.61)). ∎

Orientation

Theorem 3.1 has been completely proved using the supplementary hypothesis (3.46). It is now necessary to suppress (3.46) and to replace this hypothesis with (3.36).

3.7. Conclusion of the proof of Theorem 3.1

We shall proceed by approximating ϕ, the idea being the same as for the proof of Theorem 2.1, but with several additional technical difficulties.

Just as in (2.8), (2.9), there exists a sequence of functions $\phi_m \in H'(\Omega) \cap L^\infty(\Omega)$, such that

$$(3.87) \qquad \phi_m \to \phi \quad \text{in} \quad W^{1,1}(\Omega),$$

$$(3.88) \qquad \|\phi_m\|_{L^\infty(\Omega)} \leqslant \|\phi\|_{L^\infty(\Omega)}.$$

We then call $\bar{p}_m^*$ the solution of problem (3.17) corresponding to ϕ_m and $\bar{p}^*$ the solution of problem (3.17) corresponding to ϕ. We call u_m the generalized solution of problem (3.3) corresponding to ϕ_m.

We ignore, for the moment, the existence of a generalized solution of problem (3.3) associated with ϕ and our purpose is to demonstrate its existence.

To do this, we shall give *a priori* estimates on $\bar{p}_m^*$ and u_m, and pass to the limit $m \to \infty$.

Lemma 3.4.

(3.89) *$\bar{p}^*$ remains in a bounded subset of $L^{\alpha'}(\Omega) \times L^\infty(\Omega)^n$, as $m -> \infty$.*

Proof. Since $\bar{p}_m^*$ is the solution of (3.17) corresponding to ϕ_m,

$$- \langle \bar{p}^*, \Lambda\phi_m \rangle - G^*(-\bar{p}^*) \leqslant \langle \bar{p}_m^*, \Lambda\phi_m \rangle - G^*(-\bar{p}_m^*),$$

whence:

$$(3.90) \qquad G^*(-\bar{p}_m^*) + \langle \bar{p}_m^*, \Lambda\phi_m \rangle \leqslant G^*(-\bar{p}^*) + \langle \bar{p}^*, \Lambda\phi_m \rangle \\ \leqslant c < +\infty,$$

as $G^*(-\bar{p}^*) < +\infty$ and $\langle \bar{p}^*, \Lambda\phi_m \rangle$ is bounded due to (3.87).

It is now necessary to show that (3.89) is a consequence of (3.87) and (3.90). This can be done if we show that:

$$(3.91) \qquad c + \|q^*\|_{Y^*} \leqslant G^*(q^*) - \langle q^*, \Lambda\phi_m \rangle, \qquad \forall q^* \in Y^*,$$

with c independent of m.

Let $\mathscr{B}$ be the ball of radius unity of $L^{\alpha}(\Omega) \times L^{1}(\Omega)^{n}$.

It follows from (3.10) that G is bounded on the bounded sets and thus:

$$|(q + \Lambda\phi_m)| \leqslant c < +\infty, \qquad \forall q \in \mathscr{B}, \quad \forall m.$$

This implies that:

$$G(q + \Lambda\phi_m) \leqslant c + \chi_{\mathscr{B}}(q), \qquad \forall q \in Y, \tag{3.92}$$

where $\chi_{\mathscr{B}}$ is the indicator function of $\mathscr{B}$. Inequality (3.92) passes to the conjugate functions; using I(4.5), I(4.9), I(4.10), and noting that:

$$\chi^*_{\mathscr{B}}(q^*) = \|q^*\|_{Y^*},$$

we find exactly (3.91).

Lemma 3.5. *When $m \to \infty$, $\bar{p}^*_m \to \bar{p}^*$ in $L^{\alpha'}(\Omega) \times L^{\infty}(\Omega)^n$, in the weak* sense.*

Proof. The reasoning is standard with (3.89) and the properties of G^* (convexity, l.s.c.).

Lemma 3.6. *The sequence u_m is bounded in $L^{\infty}(\Omega)$ and $W^{1,1}(\Omega)$. Furthermore for all $\mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega$:*

$$\|u_m\|_{W^{1,\infty}(\mathcal{O})} \leqslant c(\mathcal{O}, \Omega, M_0), \tag{3.93}$$

$$\|u_m\|_{H^2(\mathcal{O})} \leqslant c(\mathcal{O}, \Omega, M_0). \tag{3.94}$$

Proof. This depends essentially on Remark 3.5, where we replace ϕ and u by ϕ_m and u_m.

The fact that u_m is bounded in $L^{\infty}(\Omega)$ is a consequence of (3.83) and (3.88); u_m is bounded in $W^{1,1}(\Omega)$ by (3.84), (3.10) and (3.87).

The upper bounds (3.93) and (3.94) result directly from (3.85) and (3.86).

Lemma 3.7. *When $m \to \infty$, $u_m \to u$, where u satisfies* (3.38) *to* (4.32) (*i.e. the conclusions of Theorem* 3.1).

Proof. This is accomplished by extracting a suitable subsequence, using Lemma 3.6, then repeating the reasoning of 3.4, 3.5 and 3.6.

Theorem 3.1 is completely proved. ∎

3.8. Proof of Proposition 3.2

Let us assume that (3.43) holds true. Then $|\operatorname{grad} u(x)|$ is bounded on an open set $\Omega_0 = B_\rho(x_0) \cap \Omega$, $B_\rho(x_0)$ = some open ball centred at x_0 with radius $\rho > 0$. As well as (3.81) (which is (3.39)), we have

$$\operatorname*{Sup}_{\Omega_0} |\Lambda u(x)| < +\infty. \tag{3.95}$$

Because of (3.35) and (3.95), we can recommence the reasoning of Section 3.6, with $\mathcal{O} = \Omega_0$.

If $\{v_m\}$ is a bounded minimizing sequence, we thus obtain the existence of a subsequence v_{m_j}, such that:

$$(3.96) \qquad \left| \begin{array}{ll} v_{m_j} \to u + c \quad \text{strongly in} \quad L^1(\Omega_0), \\ \partial v_{m_j}/\partial x_i \to \partial u/\partial x_i \quad \text{strongly in} \quad L^1(\Omega_0), \qquad 1 < i < n, \end{array} \right.$$

where u is one of the functions given by Theorem 3.1 and $c \in \mathbf{R}$ an appropriate constant.

Since $v_m = \phi$ on Γ, we have $v_m = \phi$ on $\Gamma_0 = \partial\Omega_0 \cap \Gamma$ and hence at the limit

$$(3.97) \qquad u = \phi + c \quad \text{on} \quad \Gamma_0.$$

The constant c is thus independent of the subsequence v_{m_j} under consideration and it is the whole subsequence v_m which gives rise to the convergences (3.96). We can then take $c = 0$, as because of (3.97) and (3.40), the function u is defined in a unique way.

The reasoning which we have just provided in the neighbourhood of x_0 can be reproduced in the neighbourhood of any point $x \in \Gamma$ such that

$$\overline{\lim_{\substack{y \to x \\ y \in \Omega}}} |\text{grad } u(y)| < +\infty.$$

Remark 3.6. Results of the same type as those given in Theorem 2.2 are available for problem (3.3) too. Also, using the same techniques as in Section 2.3 we can associate with (3.3) problems similar to that of De Giorgi (2.37), or more generally (2.57).

4. OTHER PROBLEMS

Orientation

In this section we shall give two other examples which are related to the examples in Sections 1, 2 and 3: only the order of the operator (Section 4.2) or the boundary condition (Section 4.1) vary. The situation is then very different and much simpler, but the comparison seems interesting to us.

4.1. A Neuman type problem

We consider a function $g = g(x,u,\xi)$ defined over $\Omega \times \mathbf{R}^{n+1}$, three times continuously differentiable and satisfying the same hypotheses as in Section 3: namely (3.1), (3.2) and (3.20) to (3.35).

We are now concerned with problem (4.1)

$$(4.1) \qquad \operatorname*{Inf}_{u \in W^{1,1}(\Omega) \cap L^{\infty}(\Omega)} \left[\int_{\Omega} g(x, u(x), \operatorname{grad} u(x))\, \mathrm{d}x \right].$$

This is a Neuman type problem.

For this problem, the situation is much simpler than in Sections 2 and 3, for, as we shall see, *the problem possesses a solution.*

Proposition 4.1. *The open space Ω is assumed to be of class $\mathscr{C}^2$.*

Under hypotheses (3.1), (3.2) *and* (3.20) *to* (3.35), *problem* (4.1) *possesses a solution $u \in W^{1,1}(\Omega) \cap L^{\infty}(\Omega)$ and this solution is unique except perhaps to within an additive constant.*

Proof. We shall demonstrate the existence of a solution for regularized elliptic forms of problem of (4.1), give *a priori* estimates and obtain the existence of a solution of (4.1) by passages to the limit.

For fixed ε and $\lambda > 0$, we consider the problem (*cf.* (3.47))

$$(4.2) \qquad \operatorname*{Inf}_{v \in H^{1}(\Omega) \cap L^{\alpha}(\Omega)} \left\{ \int_{\Omega} \left[\frac{\varepsilon}{2} |\operatorname{grad} v|^2 + \lambda |v|^{\alpha} + g(x, v, \operatorname{grad} v) \right] \mathrm{d}x \right\}.$$

The existence and uniqueness of a solution $u_{\varepsilon\lambda}$ of (4.2) follow from Proposition II.1.2, because of the hypotheses for g.

We shall show that:

$$(4.3) \qquad \|u_{\varepsilon\lambda}\|_{H^1(\Omega)} \leqslant c/\sqrt{\varepsilon},$$

$$(4.4) \qquad \|u_{\varepsilon\lambda}\|_{L^{\infty}(\Omega)} \leqslant M_1$$

(*cf.* (3.30) and (3.31)), where c is independent of ε and λ. Effectively

$$\int_{\Omega} \left[\frac{\varepsilon}{2} |\operatorname{grad} u_{\varepsilon\lambda}|^2 + \lambda |u_{\varepsilon\lambda}|^{\alpha} + g(x, u_{\varepsilon\lambda}, \operatorname{grad} u_{\varepsilon\lambda}) \right] \mathrm{d}x \leqslant \int_{\Omega} g(x, 0, 0)\, \mathrm{d}x \leqslant c,$$

and then, with (3.20),

$$(4.5) \qquad \|\operatorname{grad} u_{\varepsilon\lambda}\|_{L^2(\Omega)^n} \leqslant c/\sqrt{\varepsilon}\,;$$

(4.3) is a direct consequence of (4.4) and (4.45).

Let us demonstrate (4.4); we shall merely make explicit the proof of:

$$(4.6) \qquad u_{\varepsilon\lambda}(x) \leqslant M_1 \quad \text{a.e. } x \in \Omega.$$

But $u_{\varepsilon\lambda}$ satisfies (*cf.* for example II.2.1):

$$\text{(4.7)} \qquad \int_\Omega \Big[\varepsilon(\operatorname{grad} u_{\varepsilon\lambda} \cdot \operatorname{grad} v) + \lambda\alpha\, |u_{\varepsilon\lambda}|^{\alpha-2} u_{\varepsilon\lambda} v + \sum_{i=1}^{n} \left(\frac{\partial g}{\partial \xi_i}(x, u_{\varepsilon\lambda}, \operatorname{grad} u_{\varepsilon\lambda}) \cdot \frac{\partial v}{\partial x_i} \right) + \frac{\partial g}{\partial u}(x, u_{\varepsilon\lambda}, \operatorname{grad} u_{\varepsilon\lambda}) \Big] \mathrm{d}x = 0,$$

for all $v \in H^1(\Omega) \cap L^\alpha(\Omega)$.

Taking $v = (u_{\varepsilon\lambda} - M_1)_+$, we obtain with (3.30)–(3.31):

$$\int_{\Omega \cap \{u_{\varepsilon\lambda} > M_1\}} [\varepsilon\, |\operatorname{grad} u_{\varepsilon\lambda}|^2 + \lambda\alpha\, |u_{\varepsilon\lambda}|^{\alpha-2} u_{\varepsilon\lambda}(u_{\varepsilon\lambda} - M_1)]\, \mathrm{d}x = 0.$$

This implies that $(u_{\varepsilon\lambda} - M_1)_+ = 0$ and (4.6) is proved.

A first passage to the limit $\lambda \to 0$ shows that there exists $u_\varepsilon \in H^1(\Omega) \cap L^\infty(\Omega)$, solution of

$$\text{(4.8)} \qquad \operatorname*{Inf}_{v \in H^1(\Omega) \cap L^\alpha(\Omega)} \left\{ \int_\Omega \left[\frac{\varepsilon}{2} |\operatorname{grad} v|^2 + g(x, v, \operatorname{grad} v) \right] \mathrm{d}x \right\}$$

which satisfies furthermore

$$\text{(4.9)} \qquad \|u_\varepsilon\|_{L^\infty(\Omega)} \leqslant M_1 \quad \text{(independent of } \varepsilon\text{)}.$$

The function u_ε is *possibly* only unique to within an additive constant.

This function u_ε satisfies the Euler equation of (4.8):

$$\text{(4.10)} \qquad \varepsilon\, \Delta u_\varepsilon + \sum_{i=1}^{n} \frac{\mathrm{d}g}{\mathrm{d}\xi_i}(x, u_\varepsilon, \operatorname{grad} u_\varepsilon) = \frac{\partial g}{\partial u}(x, u_\varepsilon, \operatorname{grad} u_\varepsilon),$$

and this implies that $u_\varepsilon \in \mathscr{C}^2(\Omega)$ (*cf.* for example Ladyzenskaya and Uralceva [2]).

By virtue of (4.9) we can again apply Theorem 1.1 of Temam [3] (*cf.* Lemma 3.3) to obtain

$$\text{(4.11)} \qquad \|u_\varepsilon\|_{W^{1,\infty}(\mathcal{O})} \leqslant c(\mathcal{O}, \Omega, M_1),$$

$$\text{(4.12)} \qquad \|u_\varepsilon\|_{H^2(\mathcal{O})} \leqslant c(\mathcal{O}, \Omega, H_1),$$

for any open set $\mathcal{O} \subset \overline{\mathcal{O}} \subset \Omega$.

On the other hand with (3.20), we easily obtain:

$$\|u_\varepsilon\|_{W^{1,1}(\Omega)} \leqslant c. \tag{4.13}$$

With (4.9), (4.11), (4.12) and (4.13), we can see that there exists a sequence $\varepsilon_m \to 0$ such that

$$u_{\varepsilon_m} \to u \text{ in } L^\infty(\Omega)\text{—weak}\bigstar\,, \tag{4.14}$$

$$\frac{\partial u_{\varepsilon_m}}{\partial x_i} \to \frac{\partial x_i}{\partial u} \text{ in } L^\infty(\mathcal{O})\text{—weak}\bigstar\,, \qquad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega, \quad 1 \leqslant i \leqslant n, \tag{4.15}$$

$$u_{\varepsilon_m}|_{\mathcal{O}} \to u|_{\mathcal{O}} \text{ weakly in } H^2(\mathcal{O}), \qquad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega, \tag{4.16}$$

$$u_{\varepsilon_m}|_{\mathcal{O}} \to u|_{\mathcal{O}} \text{ strongly in } H^1(\mathcal{O}), \qquad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega, \tag{4.17}$$

$$\left|\begin{array}{l} u_{\varepsilon_m}(x) \to u(x) \quad \text{a.e.} \\ \dfrac{\partial u_{\varepsilon_m}(x)}{\partial x_i} \to \dfrac{\partial u}{\partial x_i}(x) \qquad \text{a.e.} \end{array}\right. \tag{4.18}$$

(as in Section 3.4, we have made use of compactness results and of the diagonalization process).

By Fatou's lemma, (4.13) and (4.18), we have $u \in W^{1,1}(\Omega)$; hence

$$u \in W^{1,1}(\Omega) \cap L^\infty(\Omega), \qquad u|_{\mathcal{O}} \in H^2(\mathcal{O}) \cap W^{1,\infty}(\mathcal{O}) \qquad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega.$$

Finally, we verify that u is a solution of (4.1). Effectively, if $v \in H^1(\Omega) \cap L^\infty(\Omega)$, we verify as in (4.7) that

$$\int_\Omega \left[\varepsilon \operatorname{grad} u_\varepsilon \operatorname{grad} v + \sum_{i=1}^n \frac{\partial g}{\partial \xi_i}(x, u_\varepsilon, \operatorname{grad} u_\varepsilon)\frac{\partial v}{\partial x_i} + \frac{\partial g}{\partial u}(x, u_\varepsilon, \operatorname{grad} u_\varepsilon)v\right] dx = 0. \tag{4.19}$$

With (3.21), (3.23), the convergences (4.14) to (4.18) and Lebesgue's theroem we can pass to the limit in (4.19) in order to obtain

$$\int_\Omega \left[\sum_{i=1}^n \frac{\partial g}{\partial \xi_i}(x, u, \operatorname{grad} u)\cdot\frac{\partial v}{\partial x_i} + \frac{\partial g}{\partial u}(x, u, \operatorname{grad} u)\cdot v\right] dx = 0. \tag{4.20}$$

By density, since $H^1(\Omega) \cap L^\infty(\Omega)$ is dense in $W^{1,1}(\Omega) \cap L^\infty(\Omega)$, we deduce that (4.20) is true, $\forall v \in W^{1,1}(\Omega) \cap L^\infty(\Omega)$ and so (*cf.* Proposition II.2.1), u is a solution of (4.1).

The uniqueness of u to within an additive constant results from the fact that:

(4.21) $\xi \to g(x, u, \xi)$ is strictly convex,
$$\forall x \in \Omega, \quad \forall u \in \mathbf{R}, \quad \xi \in \text{a bounded set of } \mathbf{R};$$

(due to (3.24)).

Proposition 4.1 is proved. ■

Remark 4.1. The existence result given in Proposition 4.1 seems new; it was announced in R. Temam [4].

Duality for (4.1)

We set

$$\begin{aligned}
V &= W^{1,1}(\Omega) \cap L^{\alpha}(\Omega),\\
Y &= L^{\alpha}(\Omega) \times L^{1}(\Omega)^{n}, \qquad V^* = V', \qquad Y^* = L^{\alpha'}(\Omega) \times L^{\infty}(\Omega)^{n},\\
\Lambda u &= \{ u, \operatorname{grad} u \},\\
F(u) &= 0, \qquad \forall u \in V,\\
G(p) &= \int_{\Omega} g(x, p(x))\, \mathrm{d}x, \qquad \forall p \in Y.
\end{aligned}$$

The problem

$$\operatorname*{Inf}_{u \in V} \left[F(u) + G(\Lambda u) \right]$$

is indeed identical to (4.1).

We have

$$F^*(\Lambda^* p^*) = +\infty \text{ except if } \Lambda^* p^* = 0,$$

which means (*cf.* Lemma 2.1) that,

$$\operatorname{div} p_1^* = p_0^*, \qquad p_1^* \nu\,|_{\Gamma} = 0. \tag{4.22}$$

The relation (3.12) is still valid with regard to G^*.

The dual problem of (4.1) can be written

$$\sup_{\substack{p^* \in L^{\alpha'}(\Omega) \times L^{\infty}(\Omega)^n \\ \operatorname{div} p^* = p_0^* \\ p_1^* \nu|_{\Gamma} = 0}} \left[- \int_{\Omega} g^*(x, -p^*(x))\, \mathrm{d}x \right]. \tag{4.23}$$

Proposition 4.2. *Problems* (4.1) *and* (4.2) *are mutually dual and*

$$\inf \mathscr{P} = \sup \mathscr{P}^*. \tag{4.24}$$

Problem (4.1) *possesses a solution* $\bar{u}$ *which is unique except perhaps to within an additive constant. Problem* (4.23) *possesses a unique solution* $\bar{p}^*$ *and we have:*

$$\begin{aligned} g^*(x, -\bar{p}^*(x)) + g(x, \bar{u}(x), \operatorname{grad} \bar{u}(x)) &= \\ &= -\bar{p}_0^*(x) \cdot \bar{u}(x) - \bar{p}_1^*(x) \cdot \operatorname{grad} \bar{u}(x), \text{ a.e. } x \in \Omega. \end{aligned} \tag{4.25}$$

Proof. This can be proved with the help of standard reasonings and that of Section 3.

Remark 4.2. Compare (4.23) with (3.17).

4.2. A higher order problem

Let Ω be a bounded open set and let ϕ be given in $H^2(\Omega)$. We consider the problem

$$\operatorname*{Inf}_{u \in \varphi + H_0^2(\Omega)} \left[\int_\Omega (1 + |\Delta u|^2)^{1/2} \, \mathrm{d}x \right]. \tag{4.26}$$

We set:

$$V = H^2(\Omega), \qquad Y = L^1(\Omega), \qquad \Lambda = \Delta,$$
$$V^* = \text{the dual of } H^2(\Omega), \qquad Y^* = L^\infty(\Omega),$$

$$F(u) = \begin{cases} 0 & \text{if } u \in \varphi + H_0^2(\Omega), \\ +\infty & \text{otherwise,} \end{cases}$$

$$G(p) = \int_\Omega [1 + |p(x)|^2]^{1/2} \, \mathrm{d}x.$$

We have

$$F^*(\Lambda^* p^*) = \begin{cases} \langle p^*, \Delta\varphi \rangle & \text{if } \Delta p^* = 0, \\ +\infty & \text{otherwise.} \end{cases}$$

$$G^*(p^*) = \begin{cases} -\int_\Omega [1 - |p^*(x)|^2]^{1/2} \, \mathrm{d}x & \text{if } |p^*(x)| \leqslant 1 \quad \text{a.e.} \\ +\infty & \text{otherwise.} \end{cases}$$

Problem (4.26) is thus of the type III(4.16) and its dual can be written as

$$(4.27)\qquad \operatorname*{Sup}_{\substack{p^*\in L^2(\Omega)\\ \Delta p^*=0\\ |p^*(x)|\leqslant 1\ \text{a.e.}}} \left[\int_\Omega [-\,p^*(x)\,\Delta\varphi(x) + (1 - |p^*(x)^2)^{1/2}]\,dx\right].$$

By Proposition III.4.1, there exists a solution $\bar{p}^*$ of (4.27) (which is unique by strict concavity) and $\inf\mathscr{P} = \sup\mathscr{P}^*$. Since $\Delta\bar{p}^* = 0$, $\bar{p}^*$ is analytical in Ω, and as $\bar{p}^*(x)$ is not identical to ± 1, we deduce from the maximum principle that:

$$(4.28)\qquad \operatorname*{Sup}_{\mathcal{O}} |p^*(x)| \leqslant 1 - \eta, \qquad \eta = \eta(\mathcal{O}, \Omega, \varphi) \qquad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega.$$

Using the techniques of Sections 2 and 3, we arrive (much more easily this time!) at

Proposition 4.3. *Let ϕ given in $H^2(\Omega)$. Problems* (4.26) *and* (4.27) *are mutually dual and*

$$(4.29)\qquad \inf\mathscr{P} = \sup\mathscr{P}^*.$$

Problem (4.27) *possesses a unique solution $\bar{p}^*$, a harmonic function satisfying* (4.28).

Moreover, there exists a function u which is analytic in Ω, which satisfies $\Delta u \in L^1(\Omega)$ and

$$(4.30)\qquad \Delta u(x) = -\frac{p^*(x)}{[1 - |p^*(x)|^2]^{1/2}} \qquad \forall x \in \Omega,$$

$$(4.31)\qquad \Delta\frac{\Delta u}{(1 + |\Delta u|^2)^{1/2}} = 0 \ \ in\ \ \Omega.$$

(4.32) *Any minimizing sequence $\{v_m\}$ of* (4.26) *converges to u in the following sense*:

$$\Delta v_m \to \Delta u \quad in \quad L^1(\mathcal{O}), \qquad \forall \mathcal{O} \subset \bar{\mathcal{O}} \subset \Omega.$$

Finally u is the solution of (4.26) *if such a solution exists.* ■

CHAPTER VI

Duality by the Minimax Theorem

Orientation

We consider a minimization problem

$$\inf_{u \in V} \Phi(u). \tag{0.1}$$

Let us assume that we can write $\Phi(u)$ as a supremum in p of a function $L(u, p)$:

$$\Phi(u) = \sup_{p \in Z} L(u, p), \qquad \forall u \in V. \tag{0.2}$$

As the convex l.s.c. functions are the pointwise suprema of their affine continuous minorants, we can in general write $\Phi(u)$ in the form (0.2) and often in several different ways.

Problem (0.1) then takes the form

$$\inf_{u \in V} \sup_{p \in Z} L(u, p). \tag{0.3}$$

It is convenient to term the problem

$$\sup_{p \in Z} \inf_{u \in V} L(u, p), \tag{0.4}$$

the dual problem of (0.1).

We propose to study the connection between problems (0.1) (or (0.3)) and (0.4) somewhat as in Chapter III. This presentation of duality has implicitly appeared in Section 3 of Chapter III and we shall take up this point of view in a more systematic manner. This investigation is based upon the classical minimax theorems of Ky-Fan [1] and Sion [1]. The results which we will derive are essentially the same as those of Chapter III (their application to examples will also lead to the same conclusions). However, it may happen that one of the two approaches is more appropriate for a particular problem.

1. SADDLE POINT OF A FUNCTION. PROPERTIES

Our starting point is a function $L = L(u, p)$ defined on a product space and which we shall term a Lagrangian function. In Chapter III, it was convenient to consider functions with values in $\overline{\mathbf{R}}$, and this was justified by the fact that convex l.s.c. functions only "rarely" take the value $-\infty$. On the other hand, the Lagrangian functions which we are going to consider may generally take any value in $\overline{\mathbf{R}}$, and as the algebra of such functions is awkward, *we shall restrict the domain of definition of L in order to limit ourselves to the pairs* (u, p) *for which* $L(u, p)$ *is a real number.*

We therefore take a function L defined on a set $\mathscr{A} \times \mathscr{B}$ with real values. Ultimately the nature of the sets $\mathscr{A}$ and $\mathscr{B}$ will be stated precisely, but for the first results and definitions given below, $\mathscr{A}$ and $\mathscr{B}$ are merely two arbitrary sets.

Our first result is as follows:

Proposition 1.1. *If L is a function defined on* $\mathscr{A} \times \mathscr{B}$ *with real values,*

$$\sup_{p \in \mathscr{B}} \inf_{u \in \mathscr{A}} L(u, p) \leqslant \inf_{u \in \mathscr{A}} \sup_{p \in \mathscr{B}} L(u, p). \tag{1.1}$$

Proof. Indeed we have

$$\inf_{u \in \mathscr{A}} L(u, p) \leqslant L(v, p) \qquad \forall v \in \mathscr{A},\ \forall p \in \mathscr{B},$$

whence

$$\sup_{p \in \mathscr{B}} \inf_{u \in \mathscr{A}} L(u, p) \leqslant \sup_{p \in \mathscr{B}} L(v, p)$$

and

$$\sup_{p \in \mathscr{B}} \inf_{u \in \mathscr{A}} L(u, p) \leqslant \inf_{v \in \mathscr{A}} \sup_{p \in \mathscr{B}} L(v, p). \ \blacksquare$$

Remark 1.1. In connection with what we said in the Introduction, this Proposition 1.1 can be compared with Proposition III.1.1. ■

Let us now recall the definition of a saddle point:

Definition 1.1. *We say that a pair* $(\bar{u}, \bar{p}) \in \mathscr{A} \times \mathscr{B}$ *is the saddle point of L on* $\mathscr{A} \times \mathscr{B}$ *if*

$$L(\bar{u}, p) \leqslant L(\bar{u}, \bar{p}) \leqslant L(u, \bar{p}), \qquad \forall u \in \mathscr{A},\ \forall p \in \mathscr{B}. \tag{1.2}$$

The following proposition gives an existence criterion of a saddle point.

Proposition 1.2. *The function L defined on $\mathscr{A} \times \mathscr{B}$ with real values possesses a saddle point $(\bar{u}, \bar{p})$ on $\mathscr{A} \times \mathscr{B}$ if and only if*

$$\operatorname*{Max}_{p \in \mathscr{B}} \operatorname*{Inf}_{u \in \mathscr{A}} L(u, p) = \operatorname*{Min}_{u \in \mathscr{A}} \operatorname*{Sup}_{p \in \mathscr{B}} L(u, p), \text{[1]}, \tag{1.3}$$

and this number is then equal to $L(\bar{u}, \bar{p})$.

Proof. Let us assume that there exists a saddle point $(\bar{u}, \bar{p})$; then, from (1.2),

$$\operatorname*{Sup}_{p \in \mathscr{B}} L(\bar{u}, p) = L(\bar{u}, \bar{p}) = \operatorname*{Inf}_{u \in \mathscr{A}} L(u, \bar{p}). \tag{1.4}$$

But

$$\operatorname*{Inf}_{u \in \mathscr{A}} \operatorname*{Sup}_{p \in \mathscr{B}} L(u, p) \leqslant \operatorname*{Sup}_{p \in \mathscr{B}} L(\bar{u}, p), \tag{1.5}$$

$$\operatorname*{Inf}_{u \in \mathscr{A}} L(u, \bar{p}) \leqslant \operatorname*{Sup}_{p \in \mathscr{B}} \operatorname*{Inf}_{u \in \mathscr{A}} L(u, p), \tag{1.6}$$

so that

$$\operatorname*{Inf}_{u \in \mathscr{A}} \operatorname*{Sup}_{p \in \mathscr{B}} L(u, p) \leqslant \operatorname*{Sup}_{p \in \mathscr{B}} \operatorname*{Inf}_{u \in \mathscr{A}} L(u, p), \tag{1.7}$$

which together with (1.1) implies that (1.7) is in fact an equality. This further implies with (1.4) that (1.5) and (1.6) are also equalities and we then have

$$\begin{aligned} L(\bar{u}, \bar{p}) = \operatorname*{Sup}_{p \in \mathscr{B}} L(\bar{u}, p) &= \operatorname*{Min}_{u \in \mathscr{A}} \operatorname*{Sup}_{p \in \mathscr{B}} L(u, p) = \operatorname*{Inf}_{u \in \mathscr{A}} L(u, \bar{p}) = \\ &= \operatorname*{Max}_{p \in \mathscr{B}} \operatorname*{Inf}_{u \in \mathscr{A}} L(u, p). \end{aligned}$$

Conversely, let us assume that (1.3) holds, the minimum being attained in $\bar{u}$ and the maximum in $\bar{p}$. Clearly,

$$\operatorname*{Inf}_{u \in \mathscr{A}} L(u, \bar{p}) \leqslant L(u, \bar{p}) \leqslant \operatorname*{Sup}_{p \in \mathscr{B}} L(\bar{u}, p), \tag{1.8}$$

and from (1.3), the inequalities of (1.8) are in fact equalities, which shows that $(\bar{u}, \bar{p})$ is a saddle point of L.

Remark 1.2. If the function L possesses a saddle point, we then have, in particular

$$\operatorname*{Sup}_{p \in \mathscr{B}} \operatorname*{Inf}_{u \in \mathscr{A}} L(u, p) = \operatorname*{Inf}_{u \in \mathscr{A}} \operatorname*{Sup}_{p \in \mathscr{B}} L(u, p). \tag{1.9}$$

[1] We recall that the replacement of Inf (Sup) by Min (Max) means that the infimum (supremum) is attained.

We note the following criterion

Proposition 1.3. *If there exists* $u_0 \in \mathscr{A}$, $p_0 \in \mathscr{B}$ *and* $\alpha \in \mathbf{R}$ *such that*

$$L(u_0, p) \leqslant \alpha, \qquad \forall p \in \mathscr{B}, \tag{1.10}$$

$$L(u, p_0) \geqslant \alpha, \qquad \forall u \in \mathscr{A}, \tag{1.11}$$

then (u_0, p_0) *is a saddle point of* L *and*

$$\alpha = \operatorname*{Inf}_{u \in \mathscr{A}} \operatorname*{Sup}_{p \in \mathscr{B}} L(u, p) = \operatorname*{Sup}_{p \in \mathscr{B}} \operatorname*{Inf}_{u \in \mathscr{A}} L(u, p). \tag{1.12}$$

Proof. Obviously we have $\alpha = L(u_0, p_0)$ and

$$L(u_0, p) \leqslant L(u_0, p_0) \leqslant L(u, p_0), \qquad \forall u \in \mathscr{A}, \forall p \in \mathscr{B}.$$

Proposition 1.4. *The set of saddle points of* L *is of the form* $\mathscr{A}_0 \times \mathscr{B}_0$ *where*

$$\mathscr{A}_0 \subset \mathscr{A} \quad \textit{and} \quad \mathscr{B}_0 \subset \mathscr{B}.$$

Proof. We have to verify, for example, that if (u_0, p_0) and (u_1, p_1) are two saddle points of L on $\mathscr{A} \times \mathscr{B}$, then the same holds for (u_1, p_0).

From Proposition 1.2,

$$\alpha = L(u_0, p_0) = L(u_1, p_1),$$

and by definition of (u_0, p_0), (u_1, p_1),

$$L(u_1, p) \leqslant \alpha, \qquad \forall p \in \mathscr{B},$$
$$L(u, p_0) \geqslant \alpha, \qquad \forall p \in \mathscr{A}.$$

Proposition 1.3 implies that (u_1, p_0) is the saddle point.

Hypotheses concerning L

We take two reflexive Banach spaces, V and Z, and we assume henceforth that

$$\mathscr{A} \subset V \text{ is convex, closed and non-empty,} \tag{1.13}$$

$$\mathscr{B} \subset Z \text{ is convex, closed and non-empty.} \tag{1.14}$$

The function L of $\mathscr{A} \times \mathscr{B} \mapsto \mathbf{R}$ satisfies

$$\forall u \in \mathscr{A}, \; p \to L(u, p) \text{ is concave and u.s.c.,} \tag{1.15}$$

$$\forall p \in \mathscr{B}, \; a \to L(u, p) \text{ is convex and l.s.c.} \tag{1.16}$$

We can now make more precise the result of Proposition 1.4:

Proposition 1.5. *Under hypotheses* (1.13)–(1.16) *the set* $\mathscr{A}_0 \times \mathscr{B}_0$ *of the saddle points of L is convex.*

If $\{p \to L(u, p)\}$ *is strictly concave,* $\forall u \in \mathscr{A}$, *then* $\mathscr{B}_0$ *contains at most one point. If* $\{u \to L(u, p)\}$ *is strictly convex,* $\forall p \in \mathscr{B}$, *then* $\mathscr{A}$ *contains at most one point.*

Proof. Let us assume that $\mathscr{A}_0 \times \mathscr{B}_0$ is non-empty, and let

$$\alpha = \operatorname*{Inf}_{u \in \mathscr{A}} \operatorname*{Sup}_{p \in \mathscr{B}} L(u, p) = \operatorname*{Sup}_{p \in \mathscr{B}} \operatorname*{Inf}_{u \in \mathscr{A}} L(u, p).$$

If $u_1, u_2 \in \mathscr{A}_0$ and $\lambda \in \,]0, 1[$ we have,

$$L(u_1, p) \leqslant \alpha, \qquad L(u_2, p) \leqslant \alpha, \qquad \forall p \in \mathscr{B},$$

and thus, from (1.15),

$$L(\lambda u_1 + (1 - \lambda)u_2, p) \leqslant \alpha \qquad \forall p \in \mathscr{B}.$$

On the other hand $L(u, p_1) \geqslant \alpha$, $\forall u \in \mathscr{A}$, which shows, with Proposition 1.3, that $(\lambda u_1 + (1 - \lambda) u_2, p_1)$ is the saddle point of L and $\lambda u_1 + (1 - \lambda) u_2 \in \mathscr{A}_0$.

If L is strictly convex with respect to u, $\forall p \in \mathscr{B}$, then $\mathscr{A}_0$ is reduced to one point since if u_1, u_2 were two distinct points of $\mathscr{A}_0$, we would have

$$L(\lambda u_1 + (1 - \lambda)u_2, p_1) < \lambda L(u_1, p_1) + (1 - \lambda)L(u_2, p_1) = \alpha,$$

which is impossible as

$$L(\lambda u_1 + (1 - \lambda)u_2, p_1) = \alpha.$$

Characterization of a saddle point (differentiable functionals)

We shall now give a useful characterization of a saddle point, in the case where L is a differentiable function.

Proposition 1.6. *We assume, in addition to* (1.13)–(1.16), *that*

(1.17) $\qquad \forall u \in \mathscr{A},\ p \mapsto L(u, p)$ *is Gâteaux-differentiable,*

(1.18) $\qquad \forall p \in \mathscr{B},\ u \mapsto L(u, p)$ *is Gâteaux-differentiable.*

Then $(\bar{u}, \bar{p}) \in \mathscr{A} \times \mathscr{B}$ *is the saddle point of L if and only if*

(1.19) $$\left\langle \frac{\partial L}{\partial u}(\bar{u}, \bar{p}), u - \bar{u} \right\rangle \geqslant 0, \qquad \forall u \in \mathscr{A},$$

$$\left\langle \frac{\partial L}{\partial v}(\bar{u}, \bar{p}), p - \bar{p} \right\rangle \leqslant 0, \qquad \forall p \in \mathscr{B}. \tag{1.20}$$

This is a special case of the proposition given below.

Proposition 1.7. *We assume that* $L = m + \ell$ *with*

$$\begin{aligned} &\forall u \in \mathscr{A}, \ p \mapsto \ell(u, p) \ \text{ is concave and Gâteaux-differentiable,} \\ &\forall p \in \mathscr{B}, \ u \mapsto \ell(u, p) \ \text{ is convex and Gâteaux-differentiable,} \end{aligned} \tag{1.21}$$

$$\begin{aligned} &\forall u \in \mathscr{A}, \ p \mapsto m(u, p) \ \text{ is concave,} \\ &\forall p \in \mathscr{B}, \ u \to m(u, p) \ \text{ is convex.} \end{aligned} \tag{1.22}$$

Then $(\bar{u}, \bar{p}) \in \mathscr{A} \times \mathscr{B}$ *is a saddle point of* L *if and only if*

$$\left\langle \frac{\partial \ell}{\partial u}(\bar{u}, \bar{p}), u - \bar{u} \right\rangle + m(u, \bar{p}) - m(u, \bar{p}) \geqslant 0, \qquad \forall u \in \mathscr{A}, \tag{1.23}$$

$$-\left\langle \frac{\partial \ell}{\partial p}(\bar{u}, \bar{p}), p - \bar{p} \right\rangle + m(u, \bar{p}) - m(\bar{u}, \bar{p}) \geqslant 0, \qquad \forall p \in \mathscr{B}. \tag{1.24}$$

Proof. If $(\bar{u}, \bar{p})$ is a saddle point of L, then

$$L(\bar{u}, \bar{p}) \leqslant L((1 - \lambda)\bar{u} + \lambda\bar{u}, \bar{p}) \qquad \forall u \in \mathscr{A},$$

$$\begin{aligned} \ell(\bar{u}, \bar{p}) + m(\bar{u}, \bar{p}) &\leqslant \ell((1 - \lambda)\bar{u} + \lambda u, \bar{p}) + m((1 - \lambda)\bar{u} + \lambda u, \bar{p}) \\ &\leqslant \text{(by the convexity of } m) \\ &\leqslant \ell(\bar{u} + \lambda(u - \bar{u}), \bar{p}) + (1 - \lambda)m(\bar{u}, p) + \lambda m(u, \bar{p}), \end{aligned}$$

whence

$$\frac{\ell(\bar{u} + \lambda(u - \bar{u}), \bar{p}) - \ell(\bar{u}, \bar{p})}{\lambda} + m(u, \bar{p}) - m(\bar{u}, \bar{p}) \geqslant 0,$$

which implies (1.19) when $\lambda \searrow 0$.

The proof of (1.20) is analogous.

Conversely, if $(\bar{u}, \bar{p})$ satisfies (1.23) and (1.24), then

$$\forall u \in \mathscr{A}, \qquad \ell(u, \bar{p}) - \ell(\bar{u}, \bar{p}) - \left\langle \frac{\partial \ell}{\partial u}(\bar{u}, \bar{p}), u - \bar{u} \right\rangle \geqslant 0 \quad (\text{convexity of } \ell),$$

so that

$$\ell(u, \bar{p}) - \ell(\bar{u}, \bar{p}) + m(u, \bar{p}) - m(\bar{u}, \bar{p}) \geqslant 0,$$

or

$$L(\bar{u}, \bar{p}) \leqslant L(u, \bar{p}), \qquad \forall u \in \mathscr{A}.$$

In the same way, we could prove that

$$L(\bar{u}, p) \leqslant L(\bar{u}, \bar{p}), \qquad \forall p \in \mathscr{B},$$

and the result follows.

2. EXISTENCE RESULTS FOR SADDLE POINTS

2.1. Main results

Proposition 2.1. *We assume that the conditions* (1.13)–(1.16) *are satisfied and additionally that*

(2.1) $$\mathscr{A} \text{ and } \mathscr{B} \text{ are bounded.}$$

Then the function L possesses at least one saddle point $(\bar{u}, \bar{p})$ on $\mathscr{A} \times \mathscr{B}$ and

(2.2) $$L(\bar{u}, \bar{p}) = \underset{u \in \mathscr{A}}{\text{Min}}\ \underset{p \in \mathscr{B}}{\text{Max}}\ L(u, p) = \underset{p \in \mathscr{B}}{\text{Max}}\ \underset{u \in \mathscr{A}}{\text{Min}}\ L(u, p).$$

Proof. First of all, we consider the case where we also have

(2.3) $$\forall p \in \mathscr{B}, \qquad u \mapsto L(u, p) \text{ is strictly convex.}$$

Since V and Z are reflexive, $\mathscr{A}$ and $\mathscr{B}$ compact for the weak topologies of V and Z respectively. Moreover, the properties (1.15) and (1.16) of semi-continuity of L are also true for the weak topologies.

Under these conditions, for all $p \in \mathscr{B}$, the function $u \to L(u, p)$ being weakly l.s.c. is bounded from below on $\mathscr{A}$ and attains its minimum at a unique point by virtue of (2.3). This minimum is denoted by $f(p)$ and the point where the minimum is attained by $e(p) \in \mathscr{A}$:

(2.4) $$f(p) = \underset{u \in \mathscr{A}}{\text{Min}}\ L(u, p) = L(e(p), p).$$

The function $p \to f(p)$ is concave and weakly u.s.c. as a lower bound of such functions; it is therefore bounded above and attains its upper bound at a point $\bar{p}$:

(2.5) $$f(\bar{p}) = \underset{p \in \mathscr{B}}{\text{Max}}\ f(p) = \underset{p \in \mathscr{B}}{\text{Max}}\ \underset{u \in \mathscr{A}}{\text{Min}}\ L(u, p)$$

(2.6) $$f(\bar{p}) \leqslant L(u, p), \qquad \forall u \in \mathscr{A}.$$

Now, for $u \in \mathscr{A}$, $p \in \mathscr{B}$ and $\lambda \in]0,1[$, and by virtue of (1.15),

$$L(u, (1 - \lambda)\bar{p} + \lambda p) \geqslant (1 - \lambda)L(u, \bar{p}) + \lambda L(u, p).$$

In particular, by taking $u = e_\lambda = e((1 - \lambda)\bar{p} + \lambda p)$, we obtain

$$\begin{aligned} f(p) \geqslant f((1 - \lambda)\bar{p} + \lambda p) &= L(e_\lambda, (1 - \lambda)\bar{p} + \lambda p) \\ &\geqslant (1 - \lambda)L(e_\lambda, \bar{p}) + \lambda L(e_\lambda, p) \\ &\geqslant (1 - \lambda)f(\bar{p}) + \lambda L(e_\lambda, p), \end{aligned}$$

whence

$$f(\bar{p}) \geqslant L(e_\lambda, p), \qquad \forall p \in \mathscr{B}. \tag{2.7}$$

Since $\mathscr{A}$ is weakly sequentially compact, we can extract from e_λ a sequence $\lambda_n \to 0$ with e_{λ_n} converging weakly to some limit $\bar{u}$. We have $\bar{u} = e(\bar{p})$ and *this limit $\bar{u}$ is thus independent of p* and of the selected sequence λ_n:[1] indeed, by definition of e_λ,

$$L(e_\lambda, (1 - \lambda)\bar{p} + \lambda p) \leqslant L(u, (1 - \lambda)\bar{p} + \lambda p), \qquad \forall u \in \mathscr{A},$$

and from (1.16)

$$(1 - \lambda)L(e_\lambda, \bar{p}) + \lambda L(e_\lambda, p) \leqslant L(u, (1 - \lambda)\bar{p} + \lambda p) \qquad \forall u \in \mathscr{A}.$$

Since $L(e_\lambda, p)$ is bounded from below by $f(p)$, the passage to the limit in the latter inequality, $\lambda_n \to 0$, yields, because of (1.16):

$$L(\bar{u}, \bar{p}) \leqslant \varliminf_{\lambda_n \to 0} L(e_{\lambda_n}, \bar{p}) \leqslant \varlimsup_{\lambda_n \to 0} L(u, (1 - \lambda_n)p + \lambda_n p),$$

and hence $\bar{u} = e(\bar{p})$.

We can now pass to the limit in (2.8). Using (1.15) again, we find

$$f(\bar{p}) \geqslant L(\bar{u}, p), \qquad \forall p \in \mathscr{B}. \tag{2.8}$$

By virtue of (2.5), (2.8) and Proposition 1.3, $(\bar{u}, \bar{p})$ is the saddle point of L, and the proposition is proved by means of the supplementary hypothesis (2.3).

If hypothesis (2.3) is not satisfied by L, we introduce the perturbed Lagrangians L_ε,

$$L_\varepsilon(u, p) = L(u, p) + \varepsilon \|u\|_V, \qquad \varepsilon > 0,$$

defined on $\mathscr{A} \times \mathscr{B}$, which satisfy the same hypotheses as L and also (2.3).[2]

[1] By virtue of (2.3).

[2] Since V is a reflexive Banach space, we may assume that the norm of V is strictly convex; *cf.* E. Asplund [2].

By applying the above reasoning, we obtain the existence for L_ε of a saddle point $(\bar{u}_\varepsilon, \bar{p}_\varepsilon)$ on $\mathscr{A} \times \mathscr{B}$:

$$(2.9)\qquad \begin{aligned} &L(\bar{u}_\varepsilon, p) + \varepsilon \|\bar{u}_\varepsilon\|_V \leqslant L(\bar{u}_\varepsilon, \bar{p}_\varepsilon) + \varepsilon \|\bar{u}_\varepsilon\|_V \\ &\leqslant L(u, \bar{p}_\varepsilon) + \varepsilon \|u\|_V, \qquad \forall u \in \mathscr{A},\ \forall p \in \mathscr{B}. \end{aligned}$$

By the weak compactness of $\mathscr{A}$ and $\mathscr{B}$, there exists a sequence $\varepsilon_n \to 0$, with

$$\begin{aligned} \bar{u}_{\varepsilon_n} &\to \bar{u} \text{ weakly in } V, \\ \bar{p}_{\varepsilon_n} &\to \bar{p} \text{ weakly in } Z. \end{aligned}$$

Passing to the limit in (2.9), using (1.15), (1.16), we have

$$L(\bar{u}, p) \leqslant L(u, \bar{p}), \qquad \forall u \in \mathscr{A},\ \forall p \in \mathscr{B},$$

which proves that $(\bar{u}, \bar{p})$ is a saddle point of L and so proves the proposition.

Proposition 2.2. *We assume that conditions* (1.13) (1.16) *are satisfied and also that*

$$(2.10)\qquad \exists p_0 \in \mathscr{B} \text{ such that } \lim_{\substack{u \in \mathscr{A} \\ \|u\| \to \infty}} L(u, p_0) = +\infty,$$

$$(2.11)\qquad \exists u_0 \in \mathscr{A} \text{ such that } \lim_{\substack{p \in \mathscr{B} \\ \|p\| \to +\infty}} L(u_0, p) = -\infty.$$

Then L possesses at least one saddle point on L and

$$(2.12)\qquad L(\bar{u}, \bar{p}) = \operatorname*{Min}_{u \in \mathscr{A}} \operatorname*{Sup}_{p \in \mathscr{B}} L(u, p) = \operatorname*{Max}_{p \in \mathscr{B}} \operatorname*{Inf}_{p \in \mathscr{A}} L(u, p).$$

Proof. For $\mu > 0$ fixed, let

$$\begin{aligned} \mathscr{A}_\mu &= \{ u \in \mathscr{A} \mid \|u\| \leqslant \mu \}, \\ \mathscr{B}_\mu &= \{ p \in \mathscr{B} \mid \|p\| \leqslant \mu \}. \end{aligned}$$

The sets $\mathscr{A}_\mu$ and $\mathscr{B}_\mu$ are closed, convex and bounded and Proposition 2.1 shows that L possesses a saddle point $(\bar{u}_\mu, \bar{p}_\mu)$ on $\mathscr{A}_\mu \times \mathscr{B}_\mu$

$$(2.13)\qquad L(u_\mu, p) \leqslant L(\bar{u}_\mu, \bar{p}_\mu) \leqslant L(u, \bar{p}_\mu), \qquad \forall u \in \mathscr{A}_\mu,\ \forall p \in \mathscr{B}_\mu.$$

Assuming that μ is sufficiently large so that $u_0 \in \mathscr{A}_\mu$ and $p_0 \in \mathscr{B}_\mu$, we also have

$$(2.14)\qquad L(\bar{u}_\mu, p_0) \leqslant L(\bar{u}_\mu, \bar{p}_\mu) \leqslant L(u_0, \bar{p}_\mu).$$

The function $u \mapsto L(u, p_0)$ is convex, l.s.c. and coercive; whence by Proposition II.1.2 it is bounded from below

$$(2.15) \qquad -\infty < a \leqslant L(u, p_0) \qquad \forall u \in \mathscr{A}.$$

In particular

$$(2.16) \qquad -\infty < a \leqslant L(\bar{u}_\mu, p_0) \qquad \forall \mu.$$

Similarly $p \to L(u_0, p)$ is bounded from above

$$(2.17) \qquad L(u_0, p) \leqslant b < +\infty, \qquad \forall p \in \mathscr{B},$$

and

$$(2.18) \qquad L(u_0, \bar{p}_\mu) \leqslant b < +\infty, \qquad \forall \mu.$$

With (2.14) (2.16) and (2.18) we obtain

$$(2.19) \qquad L(\bar{u}_\mu, p_0) \leqslant b < +\infty, \qquad \forall \mu,$$

$$(2.20) \qquad L(u_0, p_\mu) \geqslant a > -\infty, \qquad \forall \mu.$$

By virtue of (2.10) and (2.19) (resp. of (2.11) and (2.20)) the sequence $\bar{u}_\mu$ (resp. $\bar{p}_\mu$) is bounded independently of μ. Since the numbers $L(\bar{u}_\mu, \bar{p}_\mu)$ are also bounded, there exists a sequence $\mu_j \to \infty$ such that

$$(2.21) \qquad L(\bar{u}_{\mu_j}, \bar{p}_{\mu_j}) \to \alpha,$$

$$(2.22) \qquad \bar{u}_{\mu_j} \to \bar{u} \quad \text{weakly in } \mathscr{A},$$

$$(2.23) \qquad \bar{p}_{\mu_j} \to \bar{p} \quad \text{weakly in } \mathscr{B}.$$

By virtue of (2.13)

$$L(\bar{u}, p) \leqslant \alpha \leqslant L(u, \bar{p}), \qquad \forall u \in \mathscr{A},\ \forall p \in \mathscr{B},$$

and (u, p) is a saddle point of L on $\mathscr{A} \times \mathscr{B}$.

Remark 2.1. Proposition 2.1 may be easily extended to the case where $\mathscr{A}$ and $\mathscr{B}$ are compact subsets of separated topological vector spaces.

The hypotheses of Propositions 2.1 and 2.2 can be combined to result in the existence of a saddle point when

(2.24) $\qquad$ *$\mathscr{A}$ is bounded and* (2.11) *is satisfied,*

or else

(2.25) $\qquad$ *$\mathscr{B}$ is bounded and* (2.10) *is satisfied.* ∎

2.2. A partial result

Proposition 2.3. *With the hypotheses* (1.13)–(1.16) *and*

(2.26) $\mathscr{A}$ *is bounded or else* (2.10) *is satisfied,*

then

(2.27) $$\operatorname*{Min}_{u\in\mathscr{A}} \operatorname*{Sup}_{p\in\mathscr{B}} L(u,p) = \operatorname*{Sup}_{p\in\mathscr{B}} \operatorname*{Inf}_{u\in\mathscr{A}} L(u,p).$$

Remark 2.2. From Proposition 1.2, Proposition 2.3 only provides *part* of the information necessary for the existence of a saddle point of L: The equality inf sup = sup inf and the fact that the infimum is attained for the inf sup, giving thus the existence of the first component $\bar{u}$ of the saddle point.

Remark 2.3. By virtue of the symmetrical role played by u and p, we have an analogous result:

If $\mathscr{B}$ *is bounded or if* (2.11) *is satisfied*

(2.28) $$\operatorname*{Inf}_{u\in\mathscr{A}} \operatorname*{Sup}_{p\in\mathscr{B}} L(u,p) = \operatorname*{Max}_{p\in\mathscr{B}} \operatorname*{Inf}_{u\in\mathscr{A}} L(u,p). \blacksquare$$

Proof of Proposition 2.3. For fixed $\varepsilon > 0$ we consider

$$L_\varepsilon(u,p) = L(u,p) - \varepsilon\,\|p\|, \qquad u\in\mathscr{A},\ p\in\mathscr{B}.$$

The hypotheses of Proposition 2.2 are satisfied (if necessary, see Remark 2.1); we thus deduce the existence of $(u_\varepsilon, p_\varepsilon)\in\mathscr{A}\times\mathscr{B}$ such that

(2.29) $$L(\bar{u}_\varepsilon, p) - \varepsilon\,\|p\| \leqslant L(\bar{u}_\varepsilon, \bar{p}_\varepsilon) - \varepsilon\,\|\bar{p}_\varepsilon\| \leqslant L(u, \bar{p}_\varepsilon) - \varepsilon\,\|\bar{p}_\varepsilon\|, \qquad \forall u\in\mathscr{A},\ \forall p\in\mathscr{B}.$$

From (2.29) it follows that

(2.30) $$L(\bar{u}_\varepsilon, \bar{p}_\varepsilon) \leqslant L(u, \bar{p}_\varepsilon), \qquad \forall u\in\mathscr{A}.$$

(2.31) $$L(\bar{u}_\varepsilon, \bar{p}_\varepsilon) \leqslant \operatorname*{Inf}_{u\in\mathscr{A}} L(u, \bar{p}_\varepsilon) \leqslant \operatorname*{Sup}_{p\in\mathscr{B}} \operatorname*{Inf}_{u\in\mathscr{A}} L(u,p) = \gamma.$$

By setting $p = p_0$ (in the case of hypothesis (2.10)), we also have

(2.32) $$L(\bar{u}_\varepsilon, p_0) \leqslant \varepsilon\,\|p_0\| + L(\bar{u}_\varepsilon, \bar{p}_\varepsilon) \leqslant \varepsilon\,\|p_0\| + \gamma.$$

Let us assume that $\gamma < +\infty$ (we shall ultimately consider the case $\gamma = +\infty$). In this case (2.32) means that $L(\bar{u}_\varepsilon, p_0)$ is bounded from above when $\varepsilon \to 0$ and, from (2.10),

(2.33) $\bar{u}_\varepsilon$ is bounded for $\varepsilon \to 0$.

This conclusion is immediate if, instead of (2.10), $\mathscr{A}$ is bounded.

There then exists a sequence $\varepsilon_j \to 0$ and $\bar{u} \in \mathscr{A}$, such that

$$\bar{u}_{\varepsilon_j} \to \bar{u} \text{ weakly in } \mathscr{A}. \tag{2.34}$$

If $p \in \mathscr{B}$ is fixed, then by (2.29) and (2.31),

$$L(\bar{u}, p) \leqslant \varliminf_{\varepsilon_j \to 0} L(u_{\varepsilon_j}, p) \leqslant \varliminf_{\varepsilon_j \to 0} L(\bar{u}_{\varepsilon_j}, \bar{p}_{\varepsilon_j}) \leqslant \gamma.$$

Whence

$$\operatorname*{Inf}_{u \in \mathscr{A}} \operatorname*{Sup}_{p \in \mathscr{B}} L(u, p) \leqslant \operatorname*{Sup}_{p \in \mathscr{B}} L(u, p) \leqslant \gamma = \operatorname*{Sup}_{p \in \mathscr{B}} \operatorname*{Inf}_{u \in \mathscr{A}} L(u, p). \tag{2.35}$$

Due to (1.1), all the inequalities of (2.25) are in fact equalities and the result is established.

There remains the case when $\gamma = +\infty$; from (1.1) we then have

$$+\infty = \operatorname*{Sup}_{p \in \mathscr{B}} \operatorname*{Inf}_{u \in \mathscr{A}} L(u, p) = \operatorname*{Inf}_{u \in \mathscr{A}} \operatorname*{Sup}_{p \in \mathscr{B}} L(u, p) = \operatorname*{Sup}_{p \in \mathscr{B}} L(\bar{u}, p), \qquad \forall \bar{u} \in \mathscr{A},$$

and the proposition is trivial in this case. ■

Finally, there is another existence result for saddle points which is more general than Propositions 2.1 and 2.2.

Proposition 2.4. *We assume that conditions* (1.13)–(1.16) *are satisfied and, moreover, that*

(2.36) *$\mathscr{A}$ is bounded, or else there exists $p_0 \in \mathscr{B}$ such that*

$$\lim_{\substack{\|u\| \to +\infty \\ u \in \mathscr{A}}} L(u, p_0) = +\infty,$$

(2.37) *$\mathscr{B}$ is bounded, or else*

$$\lim_{\substack{\|p\| \to +\infty \\ p \in \mathscr{B}}} \operatorname*{Inf}_{u \in \mathscr{A}} L(u, p) = -\infty.$$

Then L possesses a saddle point on $\mathscr{A} \times \mathscr{B}$.

Proof. From Proposition 2.3, (2.36) shows that

$$\operatorname*{Min}_{u \in \mathscr{A}} \operatorname*{Sup}_{p \in \mathscr{B}} L(u, p) = \operatorname*{Sup}_{p \in \mathscr{B}} \operatorname*{Inf}_{u \in \mathscr{A}} L(u, p). \tag{2.38}$$

Now the function $p \to \operatorname{Inf}_{u \in \mathscr{A}} L(u, p)$ is concave, u.s.c. and coercive on $\mathscr{B}$ (unless $\mathscr{B}$ is bounded) and so Proposition II.1.2 shows that the problem

$$\operatorname*{Sup}_{p \in \mathscr{B}} \operatorname*{Inf}_{u \in \mathscr{A}} L(u, p)$$

possesses a solution:

$$\operatorname*{Min}_{u \in \mathscr{A}} \operatorname*{Sup}_{p \in \mathscr{B}} L(u, p) = \operatorname*{Max}_{p \in \mathscr{B}} \operatorname*{Inf}_{u \in \mathscr{A}} L(u, p), \tag{2.39}$$

and, with Proposition 1.2, the result follows.

Remark 2.4. Clearly, we can interchange the roles of $\mathscr{A}$ and $\mathscr{B}$ in Proposition 2.4.

3. APPLICATIONS TO DUALITY. EXAMPLE

Orientation

In this section we shall show how the theorems of minimax allow of a different and *direct* approach towards duality. This point of view is developed in Section 3.1. All the examples of Chapters IV and V can be re-examined from this standpoint and we shall limit ourselves to a single example, to be developed in Section 3.2.

In the following section we shall render somewhat more precisely the analogy between this method of dualization and that of Fenchel and Rockafellar studied in Chapter III. Nevertheless, we wish to emphasize once more, and this will be even clearer after Sections 3 and 4, that these methods of duality are fundamentally the same as those in Chapter II and that they only differ in their *concrete and practical way* of dealing with a particular problem.

3.1. Use of the saddle point theorems in duality

We start with an optimization problem $\mathscr{P}$:

$$\operatorname*{Inf}_{u \in V} F(u) \tag{3.1}$$

which we can write, setting $\mathscr{A} = \operatorname{dom} F$, $\mathscr{A} \subset V$, in the form

$$\operatorname*{Inf}_{u \in \mathscr{A}} F(u). \tag{3.2}$$

As we indicated in the introduction to this chapter, we initially arrange to write F in the form of a supremum

$$F(u) = \operatorname*{Sup}_{p \in \mathscr{B}} L(u, p), \qquad \forall u \in \mathscr{A}, \tag{3.3}$$

so that (3.2) becomes

$$\operatorname*{Inf}_{u \in \mathscr{A}} \operatorname*{Sup}_{p \in \mathscr{B}} L(u, p). \tag{3.4}$$

Writing F in the form (3.3) is done essentially by using the theory of conjugate convex functions. For example if

$$F(u) = F_0(u) + F_1(u), \qquad \forall u \in V \tag{3.5}$$

and if F_1 is convex, l.s.c. and proper on V, then (*cf.* Chap. I):

$$F_1(u) = \operatorname*{Sup}_{u^* \in V^*} [\langle u, u^* \rangle - F_1^*(u)] \tag{3.6}$$

where $F_1^* \in \Gamma_0(V^*)$ is the conjugate functional of F_1 (V^* and V being in duality). We then have

$$F(u) = \operatorname*{Sup}_{u^* \in V^*} [\langle u, u^* \rangle + F_0(u) - F_1^*(u^*)], \qquad \forall u \in \mathscr{A}, \tag{3.7}$$

which is indeed of the form (3.3) with $\mathscr{B} = V^*$, $p = u^*$ and

$$L(u, p) = \langle u, p \rangle + F_0(u) - F_1^*(p). \tag{3.8}$$

Similarly if Λ is an operator, possibly non-linear, of V in another t.v.s. Y and if

$$F(u) = F_0(u) + F_1(\Lambda u), \tag{3.9}$$

where $F_1 \in \Gamma_0(Y)$, then we write

$$F_1(\Lambda u) = \operatorname*{Sup}_{p \in Z} [\langle \Lambda u, p \rangle - F_1^*(p)], \tag{3.10}$$

where Y and Z are two locally convex t.v.s. in duality, and $F_1^* \in \Gamma_0(Z)$ is the conjugate functional of F_1.

Problem (3.2) can then be written in the form (3.4) with $\mathscr{B} = Z$[1] and

$$L(u, p) = F_0(u) + \langle \Lambda u, p \rangle - F_1^*(p). \tag{3.11}$$

Having reduced the primal problem $\mathscr{P}$ to the form (3.4), we term problem $\mathscr{P}^*$ the dual problem of $\mathscr{P}$

$$\operatorname*{Sup}_{p \in \mathscr{B}} \operatorname*{Inf}_{u \in \mathscr{A}} L(u, p). \tag{3.12}$$

[1] Possibly a subspace of Z.

Let us now interpret the results of Sections 1 and 2.

Proposition 1.1 means that

$$-\infty \leqslant \operatorname{Sup} \mathscr{P}^* \leqslant \operatorname{Inf} \mathscr{P} \leqslant +\infty, \tag{3.13}$$

and it must be compared with Proposition III.1.1.

Proposition 2.3 gives a criterion which ensures that

$$\inf \mathscr{P} = \sup \mathscr{P}^*, \quad \textit{and } \mathscr{P} \textit{ has a solution.}$$

This is essentially the stability of problem $\mathscr{P}^*$ (*cf.* Proposition III.2.1). Remark 2.3 gives symmetrically a criterion which is sufficient to show that

$$\inf \mathscr{P} = \sup \mathscr{P}^*, \quad \textit{and } \mathscr{P}^* \textit{ has a solution.} \tag{3.14}$$

This is the stability of problem $\mathscr{P}$.

Finally, Propositions 2.1 and 2.2 give criteria which determine whether $\mathscr{P}$ and $\mathscr{P}^*$ are both stable; they are to be compared with Proposition II.2.5. When $\inf \sup L = \sup \inf L$, the existence of a saddle point for L is equivalent to the existence of solutions for $\mathscr{P}$ and $\mathscr{P}^*$ with together the stability of these problems. The saddle point is equal to the pair of solutions of $\mathscr{P}$ and $\mathscr{P}^*$ in this case.

There remain the extremality conditions. These are, when $(\bar{u},\bar{p})$ is a saddle point, the relations

$$L(\bar{u}, \bar{p}) = \operatorname*{Sup}_{p \in \mathscr{B}} L(\bar{u}, p), \tag{3.15}$$

$$L(\bar{u}, \bar{p}) = \operatorname*{Inf}_{u \in \mathscr{A}} L(u, \bar{p}). \tag{3.16}$$

Remark 3.1. It is clear that writing an optimization problem such as (3.1), (3.2) in the form (3.4) can be envisaged without difficulty in non-convex situations (*cf.* the passage from (3.5) to (3.7) or from (3.9) to (3.11)). Nevertheless, the duality theorems of Sections 1 and 2, which are valid for non-convex situations, are useless here. Duality for non-convex problems is one of the topics considered in Chapters IX and X.

Orientation

The foregoing was a general description of the application of saddle point methods to duality,[1] and of a purely heuristic parallel between this duality and that of Frenchel and Rockafellar.

[1] From the single viewpoint of the calculus of variations.

The object of Section 3.2 is to give an example of the use of duality by minimax. Section 4.1 and 4.2 will take up in a more exact manner the analysis of the connection between these two methods of duality.

3.2. An example

The preceding point of view applies to *all* the examples in the foregoing chapters. We shall limit ourselves to a single example: the second way of dualizing Mossolov's problem (Chapter IV, Section 3.1, (3.15) et seq.).

Problem (3.1) is here

$$\inf_{u \in H^1(\Omega)} \left[\int_\Omega \left(\frac{\alpha}{2} |\text{grad}\, u|^2 + \beta\, |\text{grad}\, u| - fu \right) \mathrm{d}x \right], \tag{3.17}$$

with f given in $L^2(\Omega)$.

By means of an analogous idea to that used in (3.10) we write (*cf.* the calculations in Section IV.3.1)

$$\int_\Omega \left(\frac{\alpha}{2} |\text{grad}\, u(x)|^2 + \beta\, |\text{grad}\, u(x)| \right) \mathrm{d}x$$

$$= \sup_{p \in L^2(\Omega)^n} \left\{ \int_\Omega \left[-\, p(x)\, \text{grad}\, u(x) - \frac{1}{2\alpha} (|p(x) - \beta)_+^2 \right] \mathrm{d}x \right\}.$$

Then, problem (3.17) in its form (3.4) can be written as

$$\inf_{u \in H_0^1(\Omega)} \sup_{p \in L^2(\Omega)^n} L(u, p), \tag{3.18}$$

with

$$L(u, p) = \int_\Omega \left[-\, p(x)\, \text{grad}\, u(x) - f(x) u(x) - \frac{1}{2\alpha} (|p(x)| - \beta)_+^2 \right] \mathrm{d}x. \tag{3.19}$$

The dual problem is (*cf.* (3.12))

$$\sup_{p \in L^2(\Omega)^n} \inf_{u \in H_0^1(\Omega)} L(u, p). \tag{3.20}$$

But

$$\operatorname*{Inf}_{u \in H_0^1(\Omega)} \int_\Omega (-\,p(x) \cdot \operatorname{grad} u(x) - f(x)u(x))\,\mathrm{d}x$$

$$= \operatorname*{Inf}_{u \in \mathscr{D}(\Omega)} \int_\Omega (-\,p(x) \operatorname{grad} u(x) - f(x)u(x))\,\mathrm{d}x$$

$$= \begin{vmatrix} 0 & \text{if } \operatorname{div} p = f, \\ -\infty & \text{otherwise.} \end{vmatrix}$$

Problem (3.18) is thus

$$\operatorname*{Sup}_{\substack{p \in L^2(\Omega)^n \\ \operatorname{div} p = f}} \left[-\frac{1}{2\alpha} \int_\Omega (|p(x)| - \beta)_+^2 \,\mathrm{d}x \right]. \tag{3.21}$$

The above proposition applies (in the interchanged circumstances note in Remark 2.4). We therefore have the existence of a solution $\bar{u}$ of the proble (and this solution is unique), of a solution $\bar{p}$ of problem $\mathscr{P}^*$ and, finall $\inf \mathscr{P} = \sup \mathscr{P}^*$. The extremality relations (3.15) and (3.16) give, as i Chapter IV,

$$\operatorname{grad} \bar{u}(x) = -\frac{\bar{p}(x)}{\alpha\,|\bar{p}(x)|}(|\bar{p}(x)| - \beta)_+, \qquad \text{a.e. } x \in \Omega. \tag{3.22}$$

We can proceed in an analogous way for the other examples.

4. COMPARISON OF THE METHODS OF DUALITY

We shall now specify, in a less heuristic way than in Section 3, the parallel which exists between duality by minimax and the duality of Fenchel and Rockafellar. We shall not proceed by an elaborate comparison of the two methods of which neither seems to be a consequence of the other. We shall restrict ourselves to the simple case where the functions are convex and l.s.c., and we shall see that in this case the two methods are identical. This is the purpose of Section 4.1. In Section 4.2 we shall recover by a direct demonstration the duality theorems of Arrow and Hurwicz and of Kuhn and Tucker proved in Chapter III, Section 5. We shall also ascertain that the results are far more easily established here.

4.1. Comparison of the methods of duality

Part of this comparison has been made in Chapter III.3. Our purpose here is essentially the inverse process of reducing the problems of type (3.4) to the situation of Chapter III.

In order to facilitate the comparison of the methods, *we here denote by* $L(u, p^*)$ *the Lagrangian function* defined on $\mathscr{A} \times \mathscr{B}$, where $\mathscr{A} \subset V$ and $\mathscr{B} \subset Y^*$, V and V^* on the one hand and Y and Y^* on the other hand being two locally convex t.v.s. in duality.

Taking our inspiration from III(3.2), we set, for $u \in V$ and $p \in Y$,

$$\Phi(u, p) = +\infty \quad \text{if } u \notin \mathscr{A}, \qquad \forall p \in Y, \tag{4.1}$$

$$\Phi(u, p) = \operatorname{Sup}\,[\langle p, p^*\rangle + L(u, p^*)], \qquad \forall u \in \mathscr{A},\ \forall p \in Y. \tag{4.2}$$

Lemma 4.1. *For all* $u \in V$, $p \mapsto \Phi(u, p)$ *is convex and l.s.c. on* Y.

This lemma is obvious.

Lemma 4.2. *We assume that*

$$\mathscr{A} \subset V \textit{ is closed and convex} \tag{4.3}$$

$$\forall p^* \in \mathrm{E}, \qquad u \to L(u, p^*) \textit{ is convex and l.s.c.} \tag{4.4}$$

Then

$$(u, p) \mapsto \Phi(u, p) \textit{ is convex and l.s.c. on } V \times Y. \tag{4.5}$$

Proof. With (4.4), $(u, p) \mapsto \langle p, p^*\rangle + L(u, p^*)$ is convex and l.s.c. on $\mathscr{A} \times Y$, $\forall p^* \in \mathscr{B}$. The same therefore holds for Φ. With (4.1) and (4.3) we can then see that Φ is convex and l.s.c. on the whole of the space $V \times Y$. ■

We now turn to the optimization problem (3.4), i.e.

$$\operatorname*{Inf}_{u \in \mathscr{A}} \operatorname*{Sup}_{p \in \mathscr{B}} L(u, p). \tag{4.6}$$

Clearly, this problem is now identical to problem III(1.1)–(1.2), i.e.

$$\operatorname*{Inf}_{u \in V} \Phi(u, 0). \tag{4.7}$$

The dual of problem (4.6) is the problem

$$\operatorname*{Sup}_{p \in \mathscr{B}} \operatorname*{Inf}_{u \in \mathscr{A}} L(u, p^*), \tag{4.8}$$

while the dual problem of (4.7) can be written as

$$\text{(4.9)} \qquad \operatorname*{Sup}_{p^* \in Y^*} [-\Phi^*(0, p^*)].$$

Proposition 4.1. *We assume that*

$$\text{(4.10)} \qquad \mathscr{B} \subset Y^* \textit{ is closed and convex,}$$

(4.11) $\forall u \in \mathscr{A}$, *the function* $p^* \to L(u, p^*)$ *is concave and u.s.c. on* $\mathscr{B}$.

Then problems (4.8) *and* (4.9) *are the same and the two concepts of duality are identical to each other.*

Proof. From (4.1) and (4.2)

$$\Phi^*(0, p^*) = \operatorname*{Sup}_{\substack{u \in \mathscr{A} \\ p \in Y}} [\langle p^*, p \rangle - \Phi(u, p)].$$

Then, with Chapter I:

$$\text{(4.12)} \qquad \begin{aligned} \Phi^*(0, p^*) &= \operatorname*{Sup}_{u \in \mathscr{A}} [-L(u, p^*)] \\ -\Phi^*(0, p^*) &= \operatorname*{Inf}_{u \in \mathscr{A}} L(u, p^*), \end{aligned}$$

which demonstrates the identity of problems (4.8) and (4.9). ∎

In the case where hypotheses (4.3), (4.4), (4.12) and (4.13) are satisfied, there is identity between the two dualities and the comparison of the results is easy. Proposition 1.1 is identical to Proposition III.1.1. The identity between Propositions 2.3 and III.2.2 results from the following lemma.

Lemma 4.3. *Under hypotheses* (4.10) *and* (4.11) *the following two conditions are equivalent to each other*

(4.13) $\exists u_0 \in V$, *such that the function* $p \mapsto \Phi(u_0, p)$ *is continuous at* 0.

(4.14) $\exists u_0 \in V$, *such that* $\{p^* \in Y^* | L(u_0, p^*) \geqslant \rho\}$ *is bounded in* Y^*, $\forall \rho \in \mathbf{R}$.

Proof. This is a direct consequence of Proposition I.4.3.

Remark 4.1. Under the conditions of Proposition 2.3, condition (4.14) is equivalent to a condition of the type (2.10), i.e.

$$\text{(4.15)} \qquad \exists u_0 \in V, \qquad L(u_0, p^*) \to -\infty, \text{ si } p^* \in Y^*, \quad \|p^*\|_{Y^*} \to +\infty.$$

4.2. New proof of the Arrow–Hurwicz and Kuhn–Tucker theorems

By the intermediary of Section 4.1, we can obtain the results of Section III.5 (Propositions III.5.3 and III.5.4) as consequences of the minimax theorems of Sections 1 and 2. Here we shall give a much shorter direct proof of these results.

The notation used is that of Section III.5, and the hypotheses are those of Proposition III.5.3.

The optimization problem III(5.14), i.e.,

$$\inf_{\substack{u \in \mathscr{A} \\ Bu \leqslant 0}} J(u), \tag{4.16}$$

can be written as

$$\inf_{u \in \mathscr{A}} \sup_{p^* \in \mathscr{C}^*} [J(u) + \langle p^*, Bu \rangle], \tag{4.17}$$

where $\mathscr{C}^* \subset Y^*$ is the polar cone of $\mathscr{C}$.

We denote as

$$L(u, p) = J(u) + \langle p^*, Bu \rangle, \tag{4.18}$$

the Lagrangian function defined on $\mathscr{A} \times \mathscr{C}^*$. Obviously, the dual problem of (4.17) can be written as

$$\sup_{p^* \in \mathscr{C}^*} \inf_{u \in \mathscr{A}} L(u, p^*). \tag{4.19}$$

Proposition 4.2. *We have*

$$\inf_{u \in \mathscr{A}} \sup_{p^* \in \mathscr{C}^*} L(u, p^*) = \max_{p^* \in \mathscr{C}^*} \inf_{u \in \mathscr{A}} L(u, p^*). \tag{4.20}$$

Proof. This is a direct consequence of Proposition 2.3 and of the following lemma. ■

Lemma 4.4. *Hypothesis* III(5.24)[1] *is equivalent to*

$$\lim_{\substack{p^* \in \mathscr{C}^* \\ \|p^*\|_{Y^*} \to +\infty}} \langle p^*, Bu_0 \rangle = -\infty. \tag{4.21}$$

Proof. Let $\Psi \in \Gamma(Y^*)$ be the function defined by

$$\Psi(p^*) = \left| \begin{array}{ll} -\langle p^*, Bu_0 \rangle, & \text{if } p^* \in \mathscr{C}^*, \\ +\infty & \text{otherwise.} \end{array} \right. \tag{4.22}$$

[1] I.e.—$Bu_0 \in \mathring{\mathscr{C}}$ = the interior of $\mathscr{C}$.

Condition (4.21) is equivalent to

(4.23) $$\{ p^* \in Y^* | \ \Psi(p^*) \leqslant \rho \} \quad \text{is bounded,} \quad \forall \rho \in \mathbf{R}.$$

By virtue of Proposition I.4.3, the latter condition is equivalent to

(4.24) $$\Psi^* \in \Gamma(Y) \text{ is continuous in } 0.$$

Now

(4.25) $$\Psi^*(p) = \begin{vmatrix} 0 & \text{if} \quad -p - Bu_0 \in \mathscr{C}, \\ +\infty & \text{otherwise,} \end{vmatrix}$$

and therefore (4.24) is equivalent to the qualification condition

(4.26) $$-Bu_0 \in \mathring{\mathscr{C}}.$$

Remark 4.2. Lemma 4.4 gives an interpretation, which appears to be new, of the qualification hypothesis (4.26).

Remark 4.3. Proposition III.5.3 (and hence III.5.4) *is now a direct consequence* of Proposition 4.2.

CHAPTER VII

Other Applications of Duality

Introduction

In this chapter, we shall develop other applications of duality:

in numerical analysis, where duality yields in certain cases algorithms suitable for the numerical solution of problems;

in optimal control theory;

in mechanics, where duality allows us to describe precisely the relationship between different energy principles which govern certain non-linear problems;

lastly in economics, which was the source of many theories of duality and where this concept linked with that of price plays an essential role.

This chapter is in no way a complete study of these applications of duality. We shall in fact limit ourselves to some simple remarks and to some significant examples.

1. NUMERICAL ALGORITHMS BASED ON DUALITY

Orientation

In this section we shall describe two optimization algorithms whose basic idea depends on duality. We shall describe these algorithms and study their convergence under relatively restricted hypotheses. These algorithms, however, may also be applied to many other situations.

1.1. Uzawa's algorithm

Let V and Z be two Hilbert spaces and

(1.1) $$\mathscr{A} \subset V \quad \text{a non-empty closed convex set,}$$

(1.2) $$\mathscr{B} \subset Z \quad \text{a non-empty closed convex set.}$$

Let L be a function of $\mathscr{A} \times \mathscr{B}$ into $\mathbf{R}$. We assume that the optimization problem is given in the form envisaged in Chapter VI,

(1.3) $$\inf_{u \in \mathscr{A}} \{ \sup_{p \in \mathscr{B}} L(u, p) \}.$$

In addition to the usual hypotheses (convexity, continuity, coerciveness), which we shall specify, we make the restrictive hypothesis that

the function $p \to L(u, p)$ is affine continuous,

that is,

(1.4) $$L(u, p) = J(u) + ((p, \Phi(u)))_Z$$

where J is a mapping of $\mathscr{A}$ into $\mathbf{R}$, and Φ is a mapping of $\mathscr{A}$ into Z (not necessarily linear).

We assume that

(1.5) $$\mathscr{B} \text{ is bounded,}$$

(1.6) $$J \text{ is Gâteaux-differentiable from } \mathscr{A} \text{ into } \mathbf{R},$$

(1.7) $$((J'(u) - J'(v), u - v))_V \geqslant \alpha \|u - v\|_V^2, \qquad \alpha > 0, \quad \forall u, v \in \mathscr{A},$$

(1.8) $\forall p \in \mathscr{B}$, the function $v \to ((p, \Phi(v)))_Z$ is convex and l.s.c. on $\mathscr{A}$,

(1.9) Φ is lipschitzian from $\mathscr{A}$ into Z, i.e.,

$$\|\Phi(u) - \Phi(v)\|_Z \leqslant c \|u - v\|_V^2, \qquad \forall u, v \in \mathscr{A}.$$

By virtue of (1.6) and (1.7):

Lemma 1.1. *J is a convex and l.s.c. mapping of $\mathscr{A}$ into* $\mathbf{R}$, *and*

(1.10) $$J(u) \geqslant J(v) + ((J'(v), u - v))_V + \frac{\alpha}{2} \|u - v\|_V^2, \qquad \forall u, v \in \mathscr{A},$$

(1.11) $$\lim_{\substack{u \in \mathscr{A} \\ \|u\|_V \to \infty}} J(u) = +\infty.$$

Proof. The convexity follows from Proposition I.5.4. From the inequality I(5.15)

$$J(v) \geqslant J(u) + ((J'(u), v - u))_V, \qquad \forall u, v \in \mathscr{A}$$

and if therefore $v \to u$ either weakly or strongly in V, we have

$$\liminf_{v \to u} J(v) \geqslant J(u),$$

which proves the semi-continuity.

Finally for $u, v \in \mathscr{A}$ and $t \in [0, 1]$, we have from (1.7)

$$((J'(v + t(u - v)), u - v))_V \geqslant ((J'(v), u - v))_V + \alpha t \|u - v\|_V^2).$$

By integrating with respect to t over $[0, 1]$, we obtain (1.10) and then (1.11) is a consequence of this inequality. ■

Before describing the algorithm, we prove

Lemma 1.2. *Under hypotheses* (1.1) (1.2) *and* (1.4) *to* (1.9) *the function* L *possesses at least one saddle point* $(\bar{u}, \bar{p})$ *on* $\mathscr{A} \times \mathscr{B}$: *the first component* $\bar{u}$ *is uniquely determined and it is the solution of the optimization problem* (1.3).

Proof. We will prove the existence of a saddle point using Proposition VI.2.2 (*cf.* also Remark 2.1 and (2.25)). We shall show VI(2.10): if p_0 is any element of $\mathscr{B}$ we have,

$$L(u, p) \geqslant J(u) + ((\rho_0, \Phi(u)))_Z.$$

Since Φ is lipschitzian, if u_0 is a fixed element of $\mathscr{A}$,

$$\begin{aligned} \|\Phi(u) - \Phi(u_0)\|_Z &\leqslant c\,\|u - u_0\|_V, \\ \|\Phi(u)\|_Z &\leqslant \|\Phi(u_0)\|_Z + c\,\|u - u_0\|_V. \end{aligned} \tag{1.12}$$

Then

$$L(u, p) \geqslant J(u) - \|p_0\|_Z\,\|\Phi(u_0)\|_Z - c\,\|p_0\|_Z\,\|u - u_0\|_V \tag{1.13}$$

and VI(2.10) results from this inequality together with (1.10).

The uniqueness of the first component $\bar{u}$ of the saddle point results from Proposition VI.1.5 as $u \to L(u, p)$ is strictly convex, $\forall p \in \mathscr{B}$. Also, it is clear that since $\{\bar{u}, \bar{p}\}$ is a saddle point of L on $\mathscr{A} \times \mathscr{B}$, $\bar{u}$ is the solution of the optimization problem (1.3). ■

Our aim is to approximate the solution $\bar{u}$ of the optimization problem (1.3). Since $\{\bar{u}, \bar{p}\}$ is a saddle point, we have:

$$J(\bar{u}) + ((\bar{p}, \Phi(\bar{u})))_Z \leqslant J(v) + ((\bar{p}, \Phi(v)))_Z, \qquad \forall v \in \mathscr{A}, \tag{1.14}$$

and

$$J(\bar{u}) + ((q, \Phi(\bar{u})))_Z \leqslant J(\bar{u}) + ((\bar{p}, \Phi(\bar{u})))_Z, \qquad \forall q \in \mathscr{B}, \tag{1.15}$$

whence

$$((q - \bar{p}, \Phi(\bar{u})))_Z \leqslant 0, \qquad \forall q \in \mathscr{B}, \tag{1.16}$$

which, from II(2.17), is equivalent to:

$$\bar{p} = \Pi_{\mathscr{B}}(\bar{p} + \rho\Phi(\bar{u})), \qquad \forall \rho > 0, \tag{1.17}$$

where

$$\Pi_{\mathscr{B}} = \text{the projection in } Z, \text{ on } \mathscr{B}. \tag{1.18}$$

Description of the algorithm

Making use of (1.14) and (1.17), Uzawa's algorithm is based on the construction of two sequences of elements $u^n \in \mathscr{A}$, $p^n \in \mathscr{B}$, defined in the following way: we start with any

(1.19) $$p^0 \in \mathscr{B}$$

we calculate u^0, then p^1, u^1, etc.

(1.20) p^n being known, we determine u^n as the element of $\mathscr{A}$ which minimizes $J(v) + ((p^n, \Phi(v)))_Z$.

Then we define

(1.21) $$p^{n+1} = \Pi_{\mathscr{B}}(p^n + \rho_n \Phi(u^n)),$$

where $\rho_n > 0$ will be chosen later on.

Remark 1.1. The dual problem of (1.3) can be written as

$$\operatorname*{Sup}_{p \in \mathscr{B}} \operatorname*{Inf}_{u \in \mathscr{A}} \{ J(u) + ((p, \Phi(u)))_Z \}.$$

With regularity hypotheses which are far stronger than previously and which are very restrictive, we can show that the function

$$p \to \operatorname*{Inf}_{u \in \mathscr{A}} \{ J(u) + ((p, \Phi(u)))_Z \}$$

is differentiable, with differential $\Phi(u)$. Uzawa's algorithm appears then as the standard gradient algorithm of optimization theory applied to the dual problem (*cf.* among others Céa [1], Polak [1]). ∎

Proposition 1.1. *Under hypotheses* (1.1), (1.2) *and* (1.4) *to* (1.9), *the algorithm defined by* (1.19)–(1.21) *is convergent in the following sense*:

(1.22) $$u^n \to \bar{u} \text{ in } V,$$

where $\bar{u}$ is the solution of the problem (1.3) *provided the ρ_n satisfy*

(1.23) $$0 < \rho_* \leqslant \rho_n \leqslant \rho'_*, \qquad \rho'_* \text{ sufficiently small.}^{(1)}$$

Proof. From (1.17), (1.21) and the property II(3.11) we have

(1.24) $$\|r^{n+1}\|_Z \leqslant \|r^n + \rho_n(\Phi(u^n) - \Phi(\bar{u}))\|_Z$$

where

(1.25) $$r^n = p^n - \bar{p}.$$

(1) An estimate of ρ'_* is provided in the course of the proof.

From II(2.2), (1.20) is equivalent to

$$(1.26)\quad ((J'(u^n), v - u^n))_V + ((p^n, \Phi(v) - \Phi(u^n)))_Z \geqslant 0, \qquad \forall v \in \mathscr{A},$$

and (1.14) is equivalent to

$$(1.27)\quad ((J'(\bar{u}), v - \bar{u}))_V + ((\bar{p}, \Phi(v) - \Phi(\bar{u})))_Z, \qquad \forall v \in \mathscr{A}.$$

Taking $v = \bar{u}$ in (1.26), $v = u^n$ in (1.27) and adding the equalities we have obtained, we find that

$$(1.28)\quad ((J'(u^n) - J'(\bar{u}), u^n - \bar{u}))_V + ((p^n - \bar{p}, \Phi(u^n) - \Phi(\bar{u})))_Z \leqslant 0.$$

With (1.7) and (1.9) this implies

$$(1.29)\quad \begin{aligned} &\alpha \|u^n - \bar{u}\|_V^2 + ((p^n - \bar{p}, \Phi(u^n) - \Phi(\bar{u})))_Z \leqslant 0 \\ &\alpha \|u^n - \bar{u}\|_V^2 \leqslant \|p^n - \bar{p}\|_Z \|\Phi(u^n) - \Phi(\bar{u})\|_Z \\ &\qquad \leqslant c \|p^n - \bar{p}\|_Z \|u^n - \bar{u}\|_Z. \end{aligned}$$

By virtue of (1.5), $\|p^n - \bar{p}\|_Z$ is bounded and we deduce that

(1.30) the sequence u^n is bounded in V.

From (1.24)

$$\begin{aligned} \|r^{n+1}\|_Z^2 &= \|r^n\|_Z^2 + 2\rho_n((r^n, \Phi(u^n) - \Phi(\bar{u})) + \rho_n^2 \|\Phi(u^n) - \Phi(\bar{u})\|_Z^2 \\ &\leqslant \text{(with (1.9) and (1.29))} \\ &\leqslant \|r^n\|_Z^2 - 2\alpha\rho_n \|u^n - \bar{u}\|_V^2 + c^2\rho_n^2 \|u^n - \bar{u}\|^2. \end{aligned}$$

Let us assume that ρ_* and ρ_*' are chosen in such a way that

$$(1.31)\quad 2\alpha\rho_n - c^2\rho_n^2 \geqslant \beta > 0 \quad \text{if} \quad \rho_n \in [\rho_*, \rho_*'].$$

Then

$$(1.32)\quad \|r^{n+1}\|_Z^2 + \beta \|u^n - \bar{u}\|_V^2 \leqslant \|r^n\|_Z^2.$$

Thus $\|r^n\|_Z^2$ decreases with n, and converges for $n \to \infty$ to a limit ℓ; (1.32) then implies

$$\beta \|u^n - \bar{u}\|_V^2 \to 0,$$

whence (1.22). ■

Remark 1.2. We can replace (1.5) by

(1.5′) $\mathscr{A}$ is bounded.

However, we must also assume the existence of a saddle point $(\bar{u}, \bar{p})$ which in this case does not follow from the hypotheses. The proof of Proposition 4.1 can be developed in an identical fashion up to (1.28). Property (1.30) is a direct result of (1.5′). The reasoning subsequent to (1.30) is valid as it stands. ■

1.2. Arrow–Hurwicz's algorithm

Uzawa's algorithm is not completely determined since in (1.20) we have a choice of the algorithm for calculating u^n. The Arrow–Hurwicz algorithm is a variation of this algorithm which specifies the algorithm for calculating u^n.

We shall make this study with more precise hypotheses than in Section 1.1 and which we now specify: V and Z denote two Hilbert spaces and let:

(1.33) $\mathscr{B} \subset Z$ be a non-empty closed convex set.

Let $\Phi \in \mathscr{L}(V, Z)$ and $A \in \mathscr{L}(V, V')$ (V' is the dual space of V) satisfying

$$A = A^* \tag{1.34}$$

$$\langle Av, v \rangle \geqslant \alpha \|v\|_V^2, \qquad \alpha > 0. \tag{1.35}$$

Denoting by ℓ an element of V', we set

$$J(v) = \langle Av, v \rangle - 2\langle \ell, v \rangle.. \tag{1.36}$$

As in (1.3), we are interested in the approximation of the problem

$$\operatorname*{Inf}_{u \in V} \{ \operatorname*{Sup}_{p \in \mathscr{B}} L(u, p) \}. \tag{1.37}$$

Since all the hypotheses of Lemma 1.2 are satisfied, $L(u, p)$ possesses at least one saddle point $(\bar{u}, \bar{p})$ on $V \times \mathscr{B}$ and $\bar{u}$ is the unique solution of (1.37).

In the present case relation (1.14) is equivalent to

$$\langle J'(\bar{u}), v \rangle + ((\bar{p}, \Phi v))_Z = 0, \qquad \forall v \in V,$$

or

$$J'(\bar{u}) + \Phi^* \bar{p} = 0 \tag{1.38}$$

$$A\bar{u} - l + \Phi^* \bar{p} = 0, \tag{1.39}$$

while (1.16) and (1.17) remain unchanged:

$$((q - \bar{p}, \Phi \bar{u}))_Z \geqslant 0, \qquad \forall q \in \mathscr{B} \tag{1.16}$$

$$\bar{p} = \Pi_{\mathscr{B}}(\bar{p} + \rho \Phi \bar{u}), \qquad \forall p > 0. \tag{1.17}$$

Description of the algorithm

We start with an arbitrary p^0,

(1.40) $$p^0 \in \mathscr{B};$$

we then calculate u^1, then p^1, u^2 etc.

(1.41) $$\left| \begin{array}{l} p^n \text{ being known, we determine } u^{n+1} \text{ by} \\ u^{n+1} = u^n - \rho_1 S^{-1}(Au^n - \ell + \Phi^* p) \end{array} \right.$$

where $\rho_1 > 0$ will be stated precisely later on and

(1.42) $S =$ the canonical isomorphism of V onto V'.

Next we define

(1.43) $$p^{n+1} = \Pi_{\mathscr{B}}(p^n + \rho_2 \Phi u^{n+1}).$$

Proposition 1.2. *Under hypotheses* (1.33) *to* (1.36), *we can choose*[1] ρ_1 *and* $\rho_2 > 0$ *so that the algorithm* (1.40)–(1.43) *converges in the following sense*:

(1.44) $$u^n \to \bar{u} \quad \textit{in } V,$$

where $\bar{u}$ *is the solution of the problem* (1.37).

Proof. We set

(1.45) $$w^n = u^n - \bar{u},$$

(1.46) $$r^n = p^n - \bar{p}.$$

As in (1.24), we have

(1.47) $$\begin{aligned} &\|r^{n+1}\|_Z^2 \leqslant \|r^n + \rho_2 \Phi w^{n+1}\|_Z^2 \\ &\|r^{n+1}\|_Z^2 \leqslant \|r^n\|_Z^2 + 2\rho_2((r^n, \Phi w^{n+1}))_Z + \rho_2 \| \Phi w^{n+1} \|_Z^2. \end{aligned}$$

From (1.39) and (1.41)

$$w^{n+1} = w^n - \rho_1 S^{-1}(Aw^n + \Phi^* r^n)$$

and taking the scalar product in V with w^{n+1}, we find that

(1.48) $$\|w^{n+1}\|_V^2 = (((I - \rho_1 S^{-1}A)w^n, w^{n+1}))_V - \rho_1((S^{-1}\Phi^* r^n, w^{n+1}))_V.$$

For ρ_1 sufficiently small,

$$\|I - \rho_1 S^{-1}A\|_{\mathscr{L}(V)} \leqslant \beta < 1,$$

[1] The conditions on ρ_1 and ρ_2 are stated precisely in the proof.

and hence

$$(((I - \rho_1 S^{-1}A)w^n, w^{n+1}))_V \leqslant \beta \, \|w^n\|_V \, \|w^{n+1}\|_V \quad .$$

We also have

$$((S^{-1}\Phi^* r^n, w^{n+1}))_V = \langle \Phi^* r^n, w^{n+1} \rangle = ((r^n, \Phi w^{n+1}))_Z \quad .$$

After this (1.48) yields

$$\begin{aligned} (1.49) \qquad \rho_1((r^n, \Phi w^{n+1}))_Z &\leqslant \beta \, \|w^n\|_V \, \|w^{n+1}\| \quad - \|w^{n+1}\|_V^2 \\ &\leqslant \frac{\beta}{2} \|w^n\|_V^2 + \left(\frac{\beta}{2} - 1\right) \|w^{n+1}\|_V^2 \quad . \end{aligned}$$

Substituting this majoration in (1.47), we obtain

$$\begin{aligned} \|r^{n+1}\|_Z^2 \leqslant \|r^n\|_Z^2 &+ \lambda\beta(\|w^n\|_V^2 - \|w^{n+1}\|_V^2) \\ &+ (\lambda^2 \rho_1^2 \, \|\Phi\|^2 + 2\lambda(\beta - 1)) \, \|w^{n+1}\|_V^2 \end{aligned}$$

where $\lambda = \rho_2/\rho_1$.

But for λ sufficiently small, ρ_1 having already been selected,

$$\lambda^2 \rho_1^2 \, \|\Phi\|^2 + 2\lambda(\beta - 1) \leqslant -\gamma < 0,$$

so that

$$(1.50) \quad (\|r^{n+1}\|_Z^2 + \lambda\beta \, \|w^{n+1}\|_V^2) - (\|r^n\|_Z^2 + \lambda\beta \, \|w^n\|_V^2) \leqslant -\gamma \, \|w^{n+1}\|_V^2 .$$

Thus the sequence $\|r^n\|_Z^2 + \lambda\beta\|w^n\|_V^2$ decreases with n and converges to a limit ℓ, and it then results from (1.50) that $\|w^{n+1}\|_V^2 \to 0$ when $n \to \infty$.

Remark 1.3. We can show for the two preceding algorithms that any cluster point p of the sequence p^n is such that $(\bar{u}, \bar{p})$ is a saddle point of L on $\mathscr{A} \times \mathscr{B}$ (resp. $V \times \mathscr{B}$). ■

2. ANOTHER EXAMPLE IN NUMERICAL ANALYSIS

Orientation

In Remark 1.1 we pointed out that Uzawa's algorithm was equivalent to the application of the gradient algorithm to the dual problem. This idea can be developed in different contexts: it may be that a dual problem $\mathscr{P}^*$ is easier to tackle than the primal problem $\mathscr{P}$. We can then think of approximating $\mathscr{P}^*$, for instance with a gradient method, and then using the primal–dual relationship to deduce from this an approximation for $\mathscr{P}$. This is the method of approach which we shall develop in two examples in this section and the next.

2.1. The exact problem. Properties

Let Ω be a bounded open simply connected set in $\mathbf{R}^2$, and $f \in L^2(\Omega)$ and $a \in \mathscr{C}^0(\Omega)$, with

$$a(x) \geqslant \eta > 0, \qquad \forall x \in \Omega. \tag{2.1}$$

We set

$$\mathscr{C} = \{\, v \in H_0^1(\Omega) \mid |\operatorname{grad} v(x) \leqslant 1 \text{ almost everywhere} \}.$$

We are concerned with the following optimization problem, which is related to the Maxwell equations

$$\operatorname*{Inf}_{u \in \mathscr{C}} \left[- \int_\Omega a(1 - |\operatorname{grad} u|^2)^{1/2} \, \mathrm{d}x + \int_\Omega f u \, \mathrm{d}x \right]. \tag{2.2}$$

We have

Proposition 2.1. *Problem* (2.2) *possesses a unique solution* $\bar{u}$.

Proof. We apply Proposition II.1.2 in the space $V = H_0^1(\Omega)$, with

$$J(u) = - \int_\Omega a(1 - |\operatorname{grad} u|^2)^{1/2} \, \mathrm{d}x + \int_\Omega f u \, \mathrm{d}x. \tag{2.3}$$

The set $\mathscr{C}$ is bounded and the functional J is continuous and strictly convex since

$$s \to \sqrt{1 - s^2}$$

is strictly concave on $[0, 1]$. The result follows. ∎

The dual problem

Problem (2.2) is difficult to tackle numerically, due to the nature of the convex set $\mathscr{C}$ and the non-differentiability of the functional J. We shall now specify the dual problem of (2.2) which is more amenable.

In order to put the problem in the context of Chapter III(4.16), we set

$$V = H_0^1(\Omega), \quad V^* = H^{-1}(\Omega) = \text{the dual of } H_0^1(\Omega),$$
$$Y = Y^* = L^2(\Omega)^2, \quad \Lambda = \text{grad},$$

and

(2.4) $$F(v) = \int_\Omega f v \, \mathrm{d}x, \qquad \forall v \in V,$$

and for $p \in Y$

(2.5) $$G(p) = \begin{cases} -\displaystyle\int_\Omega a(1 - |p|^2)^{1/2} \, \mathrm{d}x & \text{if} \quad |p(x)| \leqslant 1 \quad \text{a.e.} \\ +\infty \text{ otherwise.} \end{cases}$$

It is easily verified that F (or G) is l.s.c. and convex on V and that the problem

(2.6) $$\operatorname*{Inf}_{v \in V} \{ F(v) + G(\Lambda v) \}$$

is identical to problem (2.2). If we therefore call F^* (or G^*) the conjugate functional of F (resp. G) which is l.s.c. and convex from V^* (resp. Y) into $\overline{\mathbf{R}}$, then the dual problem to (2.3) can be written as (*cf.* III(4.18)):

(2.7) $$\operatorname*{Sup}_{p^* \in Y^*} [- F^*(\Lambda^* p^*) - G^*(- p^*)].$$

We have

Lemma 2.1.

(2.8) $$F^*(\Lambda^* p^*) = \begin{cases} 0 & \textit{if } \operatorname{div} p^* = f, \\ +\infty & \textit{otherwise} \end{cases}$$

(2.9) $$G^*(p^*) = \int_\Omega [\, |a(x)|^2 + |p^*(x)|^2]^{1/2} \, \mathrm{d}x.$$

Proof. By definition of Λ

$$\begin{aligned} F^*(\Lambda^* p^*) &= \operatorname*{Sup}_{v \in V} [\langle \Lambda^* p^*, v \rangle - F(v)] \\ &= \operatorname*{Sup}_{v \in H_0^1(\Omega)} \langle f - \operatorname{div} p^*, v \rangle, \end{aligned}$$

whence (2.8).

Similarly

$$G^*(p^*) = \operatorname*{Sup}_{p \in Y} [\langle p^*, p \rangle - G(p)]$$

$$= \operatorname*{Sup}_{\substack{p \in L^2(\Omega)^2 \\ |p(x)| \leqslant 1 \text{ a.e.}}} \left[\int_\Omega [p^*(x) \cdot p(x) + |a(x)| \sqrt{1 - |p(x)|^2}\,] \, dx \right].$$

From Proposition III.1.2 it suffices for the calculation of this supremum to determine for almost all x in Ω the supremum

$$\operatorname*{Sup}_{\substack{\xi \in \mathbb{R}^2 \\ |\xi| \leqslant 1}} [p^*(x) \cdot \xi + |a(x)| \, (1 - |\xi|^2)^{1/2}]. \tag{2.10}$$

This supremum is attained at

$$\xi = \frac{p^*(x)}{[\,|p^*(x)|^2 + |a(x)|^2]^{1/2}} \tag{2.11}$$

and its value is

$$[|p^*(x)|^2 + |a(x)|^2]^{1/2}. \tag{2.12}$$

From this we deduce (2.9). ∎

Problem (2.7) (or $\mathscr{P}^*$) the dual of (2.2) (problem $\mathscr{P}$) can be written

$$\operatorname*{Sup}_{\substack{p^* \in L^2(\Omega)^2 \\ \operatorname{div} p^* = f}} \left[- \int_\Omega [\,|a(x)|^2 + |p^*(x)|^2]^{1/2} \, dx \right]. \tag{2.13}$$

If $\varphi \in H_0^1(\Omega)$ satisfies

$$-\Delta\varphi = f, \tag{2.14}$$

then setting $p^* = q^* - \nabla\varphi$ we can also write (2.13) in the form

$$\operatorname*{Sup}_{\substack{q^* \in L^2(\Omega)^2 \\ \operatorname{div} q^* = 0}} \left[- \int_\Omega [\,|a(x)|^2 + |q^*(x) - \Lambda\varphi(x)|^2]^{1/2} \, dx \right]. \tag{2.15}$$

Primal–dual relations

The relations between $\mathscr{P}$ and $\mathscr{P}^*$ are as follows:

Proposition 2.2.

$$\operatorname{Inf} \mathscr{P} = \operatorname{Sup} \mathscr{P}^* \quad \textit{and this number is finite.} \tag{2.16}$$

Problem $\mathscr{P}^*$ possesses at most one solution $\bar{p}^*$ and if this solution exists it is linked with the solution $\bar{u}$ of $\mathscr{P}$ by the extremality relation

$$(2.17) \qquad \nabla\bar{u}(x) = \frac{\bar{p}^*(x)}{[|\bar{p}^*(x)|^2 + |a(x)|^2]^{1/2}} \qquad \text{a.e. } x \in \Omega.$$

Proof. Let us assume (2.16) for the time being (see below). If $\mathscr{P}^*$ possesses a solution $\bar{p}^*$, this is necessarily unique by the strict convexity of G^*; we also know from Proposition III.4.1 and III(4.23) that $\bar{p}^*$ and $\bar{u}$ are linked by the extremality relation

$$(2.18) \qquad G(\Lambda\bar{u}) + G^*(-\bar{p}^*) = -\langle \bar{p}^*, \Lambda\bar{u} \rangle.$$

From the proof of Lemma 2.1 (see in particular (2.11)), the relation (2.18) is equivalent to (2.17). ■

Proof of (2.16). Criterion III(4.21) which usually ensures that (2.16) is true is not directly applicable here, either to problem (2.2) or to the dual problem (2.13). Nevertheless, we shall see that it applies to a modified form of these problems.

Since the open set Ω is simply connected, if $\operatorname{div} q^* = 0$ and $q^* \in L^2(\Omega)^2$, there exists $\sigma \in H^1(\Omega)$ such that

$$\partial\sigma/\partial x_1 = -q_2^*, \qquad \partial\sigma/\partial x_2 = q_1^*.$$

Problem (2.15) after a change of sign then becomes

$$(2.19) \qquad \operatorname*{Inf}_{\sigma\in H^1(\Omega)} \int_\Omega \left[a^2 + \left(\frac{\partial\sigma}{\partial x_2} - \frac{\partial\varphi}{\partial x_1}\right)^2 + \left(\frac{\partial\sigma}{\partial x_1} + \frac{\partial\varphi}{\partial x_2}\right)^2\right]^{1/2} \mathrm{d}x.$$

Setting $\tilde{V} = H^1(\Omega)$, $\tilde{V}^* = (H^1(\Omega))'$, $\tilde{Y} = \tilde{Y}^* = L^2(\Omega)^2$, $\Lambda = \text{grad}$, $\tilde{F} = 0$,

$$\tilde{G}(\pi) = \int_\Omega [a^2 + (\pi_1 + \partial\varphi/\partial x_2)^2 + (\pi_2 - \partial\varphi/\partial x_1)^2]^{1/2}\, \mathrm{d}x, \qquad \forall\pi \in \tilde{Y}.$$

The problem

$$(2.20) \qquad \operatorname*{Inf}_{\sigma\in V} [\tilde{F}(\sigma) + \tilde{G}(\tilde{\Lambda}\sigma)]$$

is thus identical to problem (2.19) and we can verify that its dual problem

$$(2.21) \qquad \operatorname*{Sup}_{\pi^*\in Y^*} [-\tilde{F}^*(\tilde{\Lambda}^*\pi^*) - \tilde{G}^*(-\tilde{\pi}^*)]$$

can be written

(2.22)

$$\operatorname*{Sup}_{\substack{\pi^*\in L^2(\Omega)^2\\ |\pi^*(x)|\leqslant 1 \text{ a.e.}\\ \operatorname{div}\pi^*=0\\ \pi^*\nu=0 \text{ on } \partial\Omega}} \left[+ \int_\Omega \left(- \pi_1^* \cdot \frac{\partial\varphi}{\partial x_2} + \pi_2^* \frac{\partial\varphi}{\partial x_1} + |E_3| (1 - |\pi^*|^2)^{1/2} \right) \mathrm{d}x \right].$$

If $\pi^* \in L^2(\Omega)^2$ and div $\pi^* = 0$, there exists $v \in H^1(\Omega)$ such that

$$+ \pi_1^* = \partial v/\partial x_2, \qquad - \pi_2^* = \partial v/\partial x_1,$$

and$\pi^* \cdot \nu = 0$ on $\partial\Omega$ implies that $\partial v/\partial\tau = 0$ on $\partial\Omega$, that is $v \in H_0^1(\Omega)$ since v is only defined to within an additive constant.

We then verify with (2.14) that

$$\int_\Omega (- \pi_1^* \, \partial\varphi/\partial x_2 + \pi_2^* \, \partial\varphi/\partial x_1) \, \mathrm{d}x = - \int_\Omega (\nabla v \cdot \nabla\varphi) \, \mathrm{d}x$$
$$= - \int_\Omega f\varphi \, \mathrm{d}x$$

so that problem (2.2) can be written as

$$(2.23) \qquad \operatorname*{Sup}_{\substack{v\in H_0^1(\Omega)\\ |\nabla v|\leqslant 1 \text{ a.e.}}} \left[\int_\Omega |E_3| (1 - |\nabla v|^2)^{1/2} - f(\rho) \, \mathrm{d}x \right],$$

which is indeed problem (2.13) except for the sign.

We now note that there exists $\sigma_0 \in H^1(\Omega)$ such that $\tilde{F}$ is finite at σ_0 and $\tilde{G}$ finite and continuous at $\tilde{\Lambda}\sigma_0$: for example $\sigma_0 = 0$. From Theorem III.4.1:

$$\text{Sup for (2.22)} = \text{Inf for (2.19)},$$

which is equivalent to (2.16). ∎

Remark 2.1. We do not know if the dual problem (2.13) effectively possesses a solution as it is not coercive (or it is only coercive in the non-reflexive space $L^1(\Omega)^2$). The form (2.19) of the problem leads us to conjecture, by comparison with the results of Section V.4.1 that the problem possesses a solution.[1]

Remark 2.2. The function u is not exactly the unknown function in the "physical" problem. The physical unknown is the vector E whose components E_1, E_2 are defined as follows

[1] *Cf.* also Remark 2.2 below.

(2.24)
$$E_1 = \frac{|a|}{(1 - |\nabla \bar{u}|^2)^{1/2}} \frac{\partial \bar{u}}{\partial x_2}$$

(2.25)
$$E_2 = -\frac{|a|}{(1 - |\nabla \bar{u}|^2)^{1/2}} \frac{\partial \bar{u}}{\partial x_1}.$$

When (2.15) possesses a solution $\bar{p}^*$, the vector E can then be expressed simply as a function of $\bar{p}^*$:

(2.26)
$$\begin{aligned} E_1 &= \bar{p}_2^* \\ E_2 &= -\bar{p}_1^*. \end{aligned}$$

Before tackling the approximation of $\bar{u}$, we must note that no regularity property can be attributed *a priori* to E_1 and E_2 since nothing is known about the set of x's where $|\text{grad}\,\bar{u}(x)| = 1$. It is thus useful to take note of the following result:

Proposition 2.3. *E_1 and E_2, defined by* (2.24) *and* (2.25), *are two functions belonging to $L^1(\Omega)$.*

Proof. Let p_n^* be a maximizing sequence of the dual problem $\mathscr{P}^*$ (i.e. (2.13)). We have:

$$p_n^* \in L^2(\Omega)^2, \qquad \text{div}\, p_n^* = f, \quad \forall n,$$

and

(2.27)
$$-\int_\Omega (a^2 + |p_n^*|^2)^{1/2}\, dx = \text{Sup}\, \mathscr{P}^* - \rho_n,$$

where $\rho_n \geqslant 0\ \forall n$, and

(2.28)
$$\rho_n \to 0, \qquad n \to \infty.$$

By virtue of (2.16), we can write (2.27) in the form

$$\begin{aligned} -F^*(\Lambda^* p_n^*) - G^*(-p_n^*) &= \text{Inf}\, \mathscr{P} - \rho_n \\ &= F(\bar{u}) + G(\Lambda \bar{u}) - \rho_n, \end{aligned}$$

whence

$$\begin{aligned} &\{ F(u) + F^*(\Lambda^* p_n^*) - \langle \Lambda^* p_n^*, u \rangle \} + \\ &\qquad + \{ G(\Lambda \bar{u}) + G^*(-p_n^*) - \langle -p_n^*, \Lambda \bar{u} \rangle \} = \rho_n. \end{aligned}$$

As each expression between braces is positive, we deduce that

(2.29)
$$\left| \begin{aligned} &0 \leqslant F(\bar{u}) + F^*(\Lambda^* p_n^*) - \langle \Lambda^* p_n^*, \bar{u} \rangle \leqslant \rho_n \\ &0 \leqslant G(\Lambda \bar{u}) + G^*(-p_n^*) - \langle -p_n^*, \Lambda \bar{u} \rangle \leqslant \rho_n \end{aligned} \right.$$

or

$$(2.30)\qquad \left|\begin{array}{l} \Lambda^* p_n^* \in \partial_{\rho_n} F(\bar{u}), \\ -\ p_n^* \in \partial_{\rho_n} G(\Lambda\bar{u}). \end{array}\right.$$

We can apply Theorem I.6.2 on ε-subdifferentials: we obtain the existence of q_n and q_n^* such that

$$(2.31)\qquad q_n, q_n^* \in L^2(\Omega)^2,$$

$$(2.32)\qquad \|q_n^* - p_n^*\|_{L^2(\Omega)^2} \leqslant \sqrt{\rho_n}, \qquad \|q_n - \Lambda\bar{u}\|_{L^2(\Omega)^2} \leqslant \sqrt{\rho_n},$$

and

$$(2.33)\qquad -\ q_n^* \in \partial G(q_n).$$

From the proof of Lemma 2.1 (*cf.* also the proof of (2.17)), the relation (2.33) means that

$$q_n(x) = \frac{q_n^*(x)}{(|a(x)|^2 + |q_n^*(x)|^2)^{1/2}} \qquad \text{a.e. } x \in \Omega,$$

which is equivalent to

$$(2.34)\qquad |q_n(x)| < 1 \qquad \text{a.e. } x \in \Omega,$$

and

$$(2.35)\qquad q_n^*(x) = \frac{|a(x)q_n(x)}{(1 - |q_n(x)|^2)^{1/2}} \qquad \text{a.e. } x \in \Omega.$$

From (2.32), there exists a sequence extracted from n (still denoted by n) such that

$$q_n(x) \to \Lambda\bar{u}(x) \qquad \text{a.e. } x \in \Omega.$$

Hence

$$(2.36)\qquad q_n^*(x) \to \frac{|a(x)|\,\Lambda\bar{u}(x)}{[1 - |\Lambda\bar{u}(x)|^2]^{1/2}} \qquad \text{a.e. } x \in \Omega.$$

As p_n^* is a maximizing sequence of (2.13), this sequence is bounded in $L^1(\Omega)^2$ and the same is true of the sequence q_n^*, from (2.32). With this remark and Fatou's lemma, we deduce that:

$$\frac{|a|\,|\Lambda\bar{u}|}{(1 - |\Lambda\bar{u}|^2)^{1/2}} \in L^1(\Omega),$$

and since the measurability of the functions E_1 and E_2 is obtained,

$$\frac{|a|\, \Lambda \bar{u}}{(1 - |\Lambda \bar{u}|^2)^{1/2}} \in L^1(\Omega)^2,$$

which proves the proposition. ∎

Remark 2.3. If we set $p_1^* = -E_2$, $p_2^* = E_1$, then from Remark 2.2 and the preceding proposition we may think that p^* is the solution in $L^1(\Omega)^2$ of problem (2.13). In this case it remains to be proved that $\operatorname{div} p^* = f$. This is the difficulty, by no means insurmountable, which we meet in the proof of the existence of a solution for (2.13).

However, this is of little importance since we have the existence and uniqueness of a solution for the primal problem which is of greater interest to us. ∎

2.2. Approximation of problem (2.3)

As the direct approximation of problem (2.3) is not easy, we shall show how a gradient algorithm applied to the dual problem (2.13) enables us to approximate the solution of problem (2.3).

To simplify the notation, we shall denote by $p, q, \ldots,$ the elements of $Y^* = Y = L^2(\Omega)^2$, and we shall set:

(2.37) $$J = G^*.$$

Differential of J

Lemma 2.2. *The functional J is two times Gâteaux differentiable in $L^2(\Omega)^2$ and*

(2.38) $$\|J'(p)\|_{L^\infty(\Omega)^2} \leqslant 1,$$

(2.39) $$\|J''(p)\|_{\mathscr{L}(Y,Y)} \leqslant c_1 = c_1(\eta, \Omega).^{(1)}$$

Proof. By applying Lebesgue's theorem, we verify that for $p, q \in L^2(\Omega)^2$

$$\frac{J(p + \lambda q) - J(p)}{\lambda} \to \int_\Omega \frac{p \cdot q}{(a^2 + |p|^2)^{1/2}} \, dx;$$

this means that

(2.40) $$J'(p) = \frac{p}{(|a|^2 + |p|^2)^{1/2}} \in L^2(\Omega)^2,$$

(1) $c_1 = c_1(\eta, \Omega) =$ constant only dependent on η and Ω.

and hence

(2.41) $$|J'(p)(x)| \leqslant 1 \qquad \text{a.e. } x \in \Omega.$$

By a new application of Lebesgue's theorem we show that

$$\frac{J'(p + \lambda q) - J'(p)}{\lambda} \to J''(p) \cdot q, \quad \text{for} \quad \lambda \to 0,$$

with

(2.42) $$J''(p) \cdot q = \int_\Omega \frac{q}{(a^2 + p^2)^{1/2}}\,\mathrm{d}x - \int \frac{p \cdot (p \cdot q)}{(|a|^2 + |p|^2)^{3/2}}\,\mathrm{d}x.$$

We have:

(2.43) $$J''(p) \in \mathscr{L}(L^2(\Omega)^2, L^2(\Omega)^2)$$

and

$$\|J''(p)\| \leqslant \left(\int_\Omega \frac{\mathrm{d}x}{|a|^2 + |p|^2}\right)^{1/2} + \left(\int_\Omega \frac{|p|^4}{(|a|^2 + |p|^2)^3}\,\mathrm{d}x\right)^{1/2}.$$

With (2.1)

(2.44) $$\|J''(p)\| \leqslant \frac{\text{mes}\,\Omega}{\eta} + c_0(\eta) \quad \text{mes}\,\Omega = c(\eta, \Omega),$$

where $c_0(\eta)$ represents the maximum on $[0, +\infty[$, of the function $s^2/(\eta^2 + s^2)^{3/2}$.

Description of the algorithm

Let ρ_n be a sequence of positive numbers for which we shall specify the hypotheses later on. Starting with

(2.45) $$p^0 \in Y = L^2(\Omega)^2$$

satisfying

(2.46) $$\operatorname{div} p^0 = f,$$

we define by recurrence the elements p^n of Y, setting

(2.47) $$p^{n+1} = p^n + \rho_n(J'(-p^n) + \nabla\theta^n),$$

where θ^n is the solution in $H^1(\Omega)$ of the Dirichlet problem

$$(2.48)\qquad \Delta\theta^n = -\operatorname{div} J'(-p^n) = \operatorname{div}\left(\frac{p^n}{(|E_3|^2 + |p^n|^2)^{1/2}}\right).$$

We shall study the behaviour of this sequence p^n, when $n \to \infty$.

It is easy to verify that $J'(-p^n) + \nabla\theta^n$ is the projection in Y of $J'(-p^n)$ on the subspace of the p's such that

$$(2.49)\qquad \operatorname{div} p = f.$$

Effectively, from (2.48)

$$(2.50)\qquad \operatorname{div}(J'(-p^n) + \nabla\theta^n) = 0$$

and from the generalized Green's formula (*cf.* Lions and Magenes [1])

$$(2.51)\qquad (\nabla\theta^n, q) = 0, \qquad \forall q \in L^2(\Omega)^2 \text{ such that } \operatorname{div} q = 0.$$

From (2.45) and (2.46) we thus obtain

$$(2.52)\qquad \operatorname{div} p^n = f, \qquad \forall n \geqslant 0,$$

Lemma 2.3.

$$(2.53)\qquad J(-p^{n+1}) \leqslant J(-p^n) + \left(\frac{c_1}{2} - \frac{1}{\rho_n}\right)\|p^{n+1} - p^n\|_Y^2.$$

Proof. We write

$$J(-p^{n+1}) - J(-p^n) = J(-p^n + \lambda(-p^{n+1} + p^n))]_{\lambda=0}^{\lambda=1}\cdot$$

$$= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\lambda}[J(-p^n + \lambda(-p^{n+1} + p^n))]\,\mathrm{d}\lambda$$

$$= \int_0^1 ((J'(-p^n + \lambda(-p^{n+1} + p^n)), (-p^{n+1} + p^n)))_Y\,\mathrm{d}\lambda$$

$$= ((J'(-p^n), (-p^{n+1} + p^n)))_Y$$

$$+ \int_0^1 (([J'(-p^n + \lambda(-p^{n+1} + p^n)) - J'(-p^n)], [-p^{n+1} + p^n]))_Y\,\mathrm{d}\lambda.$$

From (2.47), (2.51) and (2.52),

$$((J'(-p^n), (-p^{n+1} + p^n)))_Y = -\frac{1}{\rho_n}\|p^{n+1} - p^n\|_Y^2.$$

From (2.39)

$$\|J'(p) - J'(q)\|_Y \leqslant c_1 \|p - q\|_Y, \qquad \forall p, q \in Y, \tag{2.54}$$

so that

(2.55)

$$\|J'(-p^n + \lambda(-p^{n+1} + p^n)) - J'(-p^n)\|_Y \leqslant c_1\lambda \|p^{n+1} - p^n\|_Y.$$

From this we obtain

$$\int_0^1 ((J'(-p^n + \lambda(-p^{n+1} + p^n)) - J'(-p^n), -p^{n+1} + p^n))_Y \,\mathrm{d}\lambda$$
$$\leqslant c_1\lambda \|p^{n+1} - p^n\|_Y \int_0^1 \lambda \,\mathrm{d}\lambda,$$

whence finally (2.53). ■

Lemma 2.4. *We assume that the sequence ρ_n satisfies*

$$0 < c_2 \leqslant \rho_n \leqslant c_3 < +\infty, \qquad \forall n. \tag{2.56}$$

Then $\{-p^n\}$ is a maximizing sequence of the dual problem (2.13).

Proof. As the sequence $J'(-p^n)$ is bounded in Y and the sequence ρ_n is bounded in $\mathbf{R}$, there exists a subsequence still denoted by n such that

$$J'(-p^n) \to \chi \quad \text{weakly in } Y, \tag{2.57}$$

$$\rho_n \to \rho, \tag{2.58}$$

where ρ satisfies

$$c_2 \leqslant \rho \leqslant c_3. \tag{2.59}$$

From (2.52) and (2.53)

$$-J(-p^n) \leqslant -J(-p^{n+1}) \leqslant \sup \mathscr{P}^*,$$

which shows that the sequence $-J(-p^n)$ is convergent towards a limit $\ell \leqslant \sup \mathscr{P}^*$.

Let us assume that $\ell < \sup \mathscr{P}^*$. Then there would exist $p \in Y$ such that $\operatorname{div} p = f$ and

$$\ell < -J(-p) < \sup \mathscr{P}^*. \tag{2.60}$$

As J is convex, we have

$$J(-p) \geqslant J(-p^m) + (J'(-p^m), -p + p^m)_Y. \tag{2.61}$$

We set $\varepsilon = -[\ell + J(-p)]/2$. Since from (2.57)

$$((J'(-p^m), p))_Y \to ((\chi, p))_Y$$

when $m \to \infty$, there exists m^* such that $m \geqslant m_*$ implies that

$$-\varepsilon \leqslant ((J'(-p^m) - \chi, p))_Y \leqslant +\varepsilon$$

and hence from (2.61)

$$J(-p) \geqslant J(-p^m) - ((\chi, p))_Y + ((J'(-p^m), p^m))_Y - \varepsilon.$$

Since $J(-p^m) \geqslant \ell$, we find that for $m \geqslant m_*$

$$((J'(-p^m), -p^m))_Y \geqslant -((\chi, p))_Y + \varepsilon. \tag{2.62}$$

Now from (2.47), (2.51) and (2.52)

$$((p^{n+1} - p^n - \rho_n J'(-p^n), p^{n+1}))_Y = 0.$$

Whence

$$\begin{aligned} ((p^{n+1} - p^n, p^{n+1}))_Y - ((p^{n+1} - p^n, p))_Y - \rho_n((J'(-p^n), p^{n+1}))_Y &= \\ = -\rho_n((J'(-p^n), p))_Y&. \end{aligned} \tag{2.63}$$

From (2.53) and (2.56)

$$\|p^{n+1} - p^n\|_Y \to 0, \qquad n \to \infty. \tag{2.64}$$

Thus

$$((J'(-p^n), p^{n+1}))_Y = ((J'(-p^n), p^n))_Y + o(1)$$

and

$$\|p^{n+1} - p^n\|_Y = o(1).$$

We then find in (2.63) that

$$\|p^{n+1}\|_Y^2 - \|p^n\|_Y^2 + \|p^{n+1} - p^n\|_Y^2 - 2\rho_n((J'(-p^n), p^n))_Y \\ = -2\rho((\chi, p))_Y + o(1).$$

For $n \geqslant m_*$, we have, using (2.62)

$$\|p^{n+1}\|_Y^2 - \|p^n\|_Y^2 - 2\rho_n((\chi, p))_Y + 2\rho_n\varepsilon \leqslant -2\rho((\chi, p))_Y + o(1).$$

Let us sum this latter inequality from m_* to $m(m > m_*)$. After dividing by $m - m_*$ we obtain:

$$-\frac{2}{m-m_*}\left(\sum_{n=m_*}^{m} \rho_n\right)((\chi, p))_Y + \frac{2}{m-m_*}\left(\sum_{n=m_*}^{m} \rho_n\right)\varepsilon \leqslant -\rho((\chi, p))_Y \\ + \frac{1}{m-m_*}\sum_{n=m}^{m} o(1).$$

From Cesaro's theorem and (2.58), the passage to the limit in this latter inequality gives

$$-2\rho((\chi, p))_Y + 2\rho\varepsilon \leqslant -2\rho((\chi, p))_Y,$$

which is impossible since $\varepsilon > 0$.

It is then impossible that $\ell < \sup \mathscr{P}^*$, and the lemma is thus proved. ∎

Lemma 2.5. *When $n \to \infty$,*

$$\frac{p^n}{(a^2 + |p^n|^2)^{1/2}} \to \Lambda\bar{u} \quad \text{in } L^2(\Omega)^2. \tag{2.65}$$

Proof. We repeat the proof of Proposition 2.3, with the sequence p^n replaced by the sequence p_n^*, because of Lemma 2.4. We also note that

$$\left|\frac{p^n}{(a^2 + |p^n|^2)^{1/2}} - \frac{q_n^*}{(a^2 + |q_n^*|^2)^{1/2}}\right| \leqslant c_1 \|p^n - q_n^*\|_Y \to 0.$$

Since the limit is independent of the chosen subsequence in (2.57), (2.58), the whole sequence is convergent. ∎

The convergence result

The above results are summarized in the following proposition.

Proposition 2.4. *We consider the sequence p^n defined by* (2.47), *the gradient method for* (2.13), *and we assume that the ρ_n's satisfy* (2.56).

Then for $n \to \infty$ the sequence p^n approximates $\bar{u}$ in the following sense

(2.66) $$\frac{p^n}{(a^2 + |p^n|^2)^{1/2}} \to \Lambda\bar{u} \quad \textit{in } L^2(\Omega)^2. \quad \blacksquare$$

3. APPLICATION TO A PROBLEM IN OPTIMAL CONTROL THEORY

In this section we describe the application of duality to a problem of optimal control theory for a system governed by partial differential equations with a *constraint on the state*. ■

3.1. The control problem and its dual

Let Ω be an open set in $\mathbf{R}^n$, $T > 0$, $Q = \Omega \times]0, T[$, and $y_0 \in L^2(\Omega)$ and $f \in L^2(Q)$be given. For all $u \in L^2(Q)$, there exists a unique y, $y = y(x, t, u)$(1) belonging to $L^2([0, T]; H_0^1(\Omega))$ and satisfying in a weak sense (*cf.* Lions and Magenes [1]):

(3.1) $$\partial y/\partial t - \Delta y = f + u \quad \text{in } Q$$

(3.2) $$y = 0 \quad \text{on } \Sigma = \partial\Omega \times]0, T[$$

(3.3) $$y(0, x, u) = y_0(x) \qquad x \in \Omega.$$

We now define the cost function

(3.4) $$J(u) = \int_Q |y(u) - y_d|^2 \, \mathrm{d}x \, \mathrm{d}t.$$

($\nu > 0$) and we are concerned with the following problem:

(3.5) $$\text{Minimization of } J(u)$$

amongst those u's such that

(3.6) $$u \in L^2(Q),$$

(3.7) $$|\text{grad } y(x, t, u)| \leqslant 1 \quad \text{a.e.} \qquad (x, t) \in Q.$$

(1) We shall occasionally denote this by $y(u)$ when only the dependence on u plays an interesting role.

Problem (3.5) (3.6) does not necessarily admit of a solution in $L^2(Q)$ but by enlarging the class of admissible controls we obtain

Proposition 3.1. *Let us assume that there exists admissible controls u such that condition* (3.6) *is satisfied.*[1]

Then problem (3.5) (3.6) *has a unique solution in a functional space V containing* $L^2(Q)$.[2]

Proof. By setting

$$\varphi = y(0) \quad (y \text{ for } u = 0) \tag{3.8}$$

and

$$z(u) = y(u) - y(0) \tag{3.9}$$

$$z_d = y_d - y(0). \tag{3.10}$$

Then $z = z(u)$ is a weak solution of

$$\frac{\partial z}{\partial t} - \Delta z = u \quad \text{in } Q, \tag{3.11}$$

$$z = 0 \quad \text{on } \Sigma, \tag{3.12}$$

$$z(0, x; u) = 0. \tag{3.13}$$

The mapping $u \to z(u)$ from $L^2(Q)$ into $L^2([0, T]; H_0^1(\Omega))$ is linear and continuous and from Lions and Magenes [1], *it is an isomorphism of V into* $L^2(Q)$.

The function $u \to J(u)$ is thus a strictly *convex continuous function of V* into $\mathbf{R}$. From Proposition II.1.2, if the set

$$\mathscr{C} = \{ u \in V | \ |\operatorname{grad} y(x, t; u)| \leqslant 1 \quad \text{a.e.}\} \tag{3.14}$$

is non-empty, the function J attains its minimum on $\mathscr{C}$ in a unique point $\bar{u}$.[3] ∎

The dual problem

We return to the situation of Chapter III(4.16) and so set V and V^* as in footnote [2]

[1] This holds in particular if the norms of y_0 and of f are not "too large".

[2] In the notation of Lions and Magenes [1] (*cf.* Vol. II, Chap. IV, Remark 12.3 especially), we have $V = \Xi^{-2,-1}(Q)$ which is the dual of $V^* = \Xi^{2,1}(Q)$. The precise definition of these spaces is not very important here. Let us only recall that both are Hilbert spaces and that $L^2(Q) \subset V$, $V^* \subset L^2(Q)$, with continuous and dense injection. Furthermore, V is a distribution space (and it is obvious that V^* is a space of functions).

[3] We must notice that the solution $\bar{u}$ of this problem is very easy to determine, at least formally: $\mathscr{C}$ is indeed a closed convex subset of $L^2(Q)$ and $y(\bar{u})$ is the projection of y_d on $\mathscr{C}$ in $L^2(Q)$; then $\bar{u}$ is given by (3.1). However, the techniques used in this section apply as well to similar "non-trivial" control problems; *cf.* Remark 3.4.

$$Y = Y^* = Y_1 \times Y_2 = L^2(Q) \times L^2(Q)^n,$$
$$\Lambda u = (\Lambda_1 u, \Lambda_2 u),$$
$$\Lambda_1 u = z(u), \qquad \Lambda_2 u = \operatorname{grad} z(u),$$

where $z = z(u)$ is defined by (3.11)–(3.13); the elements of Y are denoted by (p_1, p_2), $p_1 \in Y_1 = L^2(Q)$, $p_2 \in Y_2 = L^2(Q)^n$.

We write

$$F(u) = 0, \qquad \forall u \in V, \tag{3.15}$$

$$G(p) = G_1(p_1) + G_2(p_2), \tag{3.16}$$

$$G_1(p_1) = \frac{1}{2}\int_Q |p_1 - z_d|^2 \,\mathrm{d}x\,\mathrm{d}t, \qquad \forall p_1 \in Y_1, \tag{3.17}$$

$$G_2(p_2) = \begin{vmatrix} 0 & \text{if } |(p_2 + \operatorname{grad}\varphi)(x, t)| \leqslant 1 \quad \text{a.e.,} \\ +\infty & \text{otherwise.} \end{vmatrix} \tag{3.18}$$

It is easily verified that F and G are convex, l.s.c. and proper and that the problem

$$\operatorname*{Inf}_{v \in V} \{ F(v) + G(\Lambda v) \} \tag{3.19}$$

is indeed the problem (3.5), (3.6).

We easily get

$$G_1^*(p_1) = \int_Q (p_1 z_d + \tfrac{1}{2}|p_1|^2)\,\mathrm{d}x\,\mathrm{d}t \tag{3.20}$$

$$G_2^*(p_2) = \int_Q [|p_2| - p_2 \operatorname{grad}\varphi]\,\mathrm{d}x\,\mathrm{d}t. \tag{3.21}$$

For F^* we have

Lemma 3.1.

$$F^*(\Lambda^* p) = \begin{vmatrix} 0 & \textit{if } p_1 = \operatorname{div} p_2, \\ +\infty & \textit{otherwise.} \end{vmatrix} \tag{3.22}$$

Proof. For $p \in Y^* = Y$,

$$F^*(\Lambda^* p) = \operatorname*{Sup}_{u \in V} \langle p, \Lambda u \rangle. \tag{3.23}$$

But

$$\langle p, \Lambda u \rangle = \int_Q [p_1 z(u) + p_2 \operatorname{grad} z(u)] \, dx \, dt$$

and since $z(u) = 0$ on Σ, this is equal, by the generalized Green's formula, to

$$\int_Q (p_1 - \operatorname{div} p_2) z(u) \, dx \, dt,$$

(*cf.* Lions and Magenes [1]).

Now let Ψ be the unique solution in $L^2([0, T]; H_0^1(\Omega))$ of the parabolic problem

$$-\frac{\partial \Psi}{\partial t} - \Delta \Psi = p_1 - \operatorname{div} p_2 \quad \text{in } Q, \tag{3.24}$$

$$\Psi = 0 \quad \text{on } \Sigma, \tag{3.25}$$

$$\Psi(x, T) = 0 \quad \text{in } \Omega. \tag{3.26}$$

We have, from integration by parts and from (3.11)–(3.13) and (3.24)–(3.26)

$$\begin{aligned}\langle p, \Lambda u \rangle &= \int_Q (p_1 - \operatorname{div} p_2) z(u) \, dx \, dt = \\ &= \int_Q \left(-\frac{\partial \Psi}{\partial t} - \Delta \Psi \right) z(u) \, dx \, dt \\ &= \int_Q \Psi u \, dx \, dt.\end{aligned}$$

Then the supremum (3.23) is equal to

$$\operatorname*{Sup}_{u \in \mathscr{V}} \int_Q \Psi u \, dx \, dt,$$

and this supremum takes the value 0 if $\Psi = 0$ and $+\infty$ otherwise. Now from (3.24)–(3.26), $\Psi = 0$ is equivalent to $p_1 = \operatorname{div} p_2$. ∎

We are now in a position to state explicitly the dual problem of (3.19):

$$\operatorname*{Sup}_{p \in Y} \{ -F^*(\Lambda^* p) - G^*(-p) \}. \tag{3.27}$$

It can be written as

$$\text{(3.28)} \qquad \operatorname*{Sup}_{\substack{p_1 \in L^2(Q) \\ p_2 \in L^2(Q)^n \\ p_1 = \operatorname{div} p_2}} \left[\int_Q \left[-\tfrac{1}{2}|p_1|^2 + p_1 z_d - |p_2| - p_2 \operatorname{grad} \varphi \right] \mathrm{d}x\, \mathrm{d}t \right]$$

where φ is defined by (3.8).

Proposition 3.2.

$$\text{(3.29)} \qquad \operatorname{Inf} \mathscr{P} = \sup \mathscr{P}^* \quad \textit{and this number is finite.}$$

If the problem $\mathscr{P}^*$ possesses a solution $\bar{p}^* = (\bar{p}_1^*, \bar{p}_2^*)$ then it is linked to the solution $\bar{u}$ of the problem $\mathscr{P}$ by the extremality relations:

$$\text{(3.30)} \qquad \bar{p}_1^* = y(\bar{u}) - y_d,$$

$$\text{(3.31)} \qquad |\bar{p}_2| + \bar{p}_2 \operatorname{grad} y(\bar{u}) = 0 \quad \textit{a.e.}$$

that is, since $|\operatorname{grad} y(\bar{u})| \leqslant 1$ *almost everywhere,*

$$\operatorname{grad} y(\bar{u}) = -\frac{\bar{p}_2}{|\bar{p}_2|} \quad \textit{a.e. when} \quad |\bar{p}_2(x)| \neq 0.$$

Proof. As in the example described in Section 2, criterion III(4.21) which usually ensures (3.29) applies here neither to the problem (3.5) (3.6) nor to the dual problem (3.28). *We will prove* (3.29) *in Section* 3.2 with the help of a passage to the limit, the problems $\mathscr{P}$ and $\mathscr{P}^*$ being approximated by the problems $\mathscr{P}_\varepsilon$ and $\mathscr{P}_\varepsilon^*$ with

$$\begin{aligned} &\inf \mathscr{P}_\varepsilon = \sup \mathscr{P}_\varepsilon^* \\ &\inf \mathscr{P}_\varepsilon \to \inf \mathscr{P}, \qquad \sup \mathscr{P}_\varepsilon^* \to \sup \mathscr{P}. \end{aligned}$$

We cannot assert that problem (3.28) possesses a solution. But if it does admit of a solution $\bar{p}^*$, then by provisionally allowing (3.29), Proposition III.4.1 (*cf.* III(4.27)) permits us to write the extremality relations as

$$\text{(3.32)} \qquad G_1(z(\bar{u})) + G_1^*(-\bar{p}_1) = -\langle z(\bar{u}), \bar{p}_1^* \rangle,$$

$$\text{(3.33)} \qquad G_2(\operatorname{grad} z(\bar{u})) + G_2^*(-\bar{p}_2) = -\langle \operatorname{grad} z(\bar{u}), \bar{p}_2 \rangle$$

which implies precisely (3.30) and (3.31). ∎

3.2. Approximated problems

We shall approximate the problems $\mathscr{P}$ and $\mathscr{P}^*$, by the problems $\mathscr{P}_\varepsilon$ and $\mathscr{P}_\varepsilon^*$ respectively, in order to prove (3.29) and also to give a numerical approximation procedure for the control problem.

The functional frame being the same as in (3.15)–(3.18), we set, for $\varepsilon > 0$ fixed,

$$G_\varepsilon(p) = G_1(p_1) + G_{2\varepsilon}(p_2), \tag{3.34}$$

where

$$G_{2\varepsilon}(p_2) = \frac{1}{2\varepsilon}\int_Q (|p_2 + \operatorname{grad}\varphi| - 1)_+^2 \, \mathrm{d}x \, \mathrm{d}t. \tag{3.35}$$

It is easily seen that the conjugate function of

$$\xi \to \frac{1}{2\varepsilon}(|\xi + \operatorname{grad}\varphi(x)| - 1)_+^2$$

is the function

$$\xi \to \frac{\varepsilon}{2}|\xi^*|^2 + |\xi^*| - \xi^* \operatorname{grad}\varphi(x),$$

and hence the conjugate function of $G_{2\varepsilon}$ can be written as:

$$G_{2\varepsilon}^*(p_2) = \int_Q \left[\frac{\varepsilon}{2}|p_2^*|^2 + |p_2| - p_2 \operatorname{grad}\varphi\right] \mathrm{d}x \, \mathrm{d}t. \tag{3.36}$$

We now define the problem $\mathscr{P}_\varepsilon$; it is

$$\operatorname*{Inf}_{v\in V}\{ F(v) + G_\varepsilon(\Lambda v) \} \tag{3.37}$$

or, alternatively

$$\operatorname*{Inf}_{u\in V}\left[J(u) + \frac{1}{\varepsilon}\int_Q (|\operatorname{grad} y(u)| - 1)_+^2 \, \mathrm{d}x \, \mathrm{d}t\right].$$

Its dual is the problem $\mathscr{P}_\varepsilon^*$

$$\operatorname*{Sup}_{p\in Y^* = Y}\{ - F^*(\Lambda^* p) - G_\varepsilon^*(- p) \} \tag{3.38}$$

or, alternatively

$$\text{(3.39)} \qquad \sup_{\substack{p_1 \in L^2(Q) \\ p_2 \in L^2(Q)^n \\ p_1 = \operatorname{div} p_2}} \left[\int_Q \left(-\frac{1}{2}|p_1|^2 + p_1 z_d - \frac{\varepsilon}{2}|p_2|^2 - |p_2| - p_2 \operatorname{grad} \varphi \right) \mathrm{d}x\, \mathrm{d}t \right].$$

Remark 3.1. The problems $\mathscr{P}_\varepsilon$ are penalized forms of the problem $\mathscr{P}$: we have penalized the constraint (3.6) with the help of the penalization function $G_{2\varepsilon}$. See e.g. Courant [1], Lions [4] for penalization.

The problems $\mathscr{P}_\varepsilon^*$ are the regularized forms of the problem $\mathscr{P}^*$: they are more regular since they are coercive in Y which ensures that they possess a solution (see below).

In this special case we can verify a completely general property: *penalization and regularization are mutually dual methods of approximation* (for a different application of this, see Bensoussan and Kenneth [1]. ∎

Proposition 3.3.

$$\text{(3.40)} \qquad \inf \mathscr{P}_\varepsilon = \sup \mathscr{P}_\varepsilon^* \quad \textit{and this number is finite.}$$

Problem $\mathscr{P}_\varepsilon$ possesses a unique solution $\bar{u}_\varepsilon \in V$, problem $\mathscr{P}_\varepsilon^*$ possesses a unique solution $\bar{p}_\varepsilon$ and

$$\text{(3.41)} \qquad \bar{p}_{1\varepsilon} = y(\bar{u}_\varepsilon) - y_d$$

$$\text{(3.42)} \qquad \left| \begin{array}{ll} |\nabla y(\bar{u}_\varepsilon)| = 1 & \textit{if } |\bar{p}_{2\varepsilon}| = 0 \\ \nabla y(\bar{u}_\varepsilon) = \dfrac{\bar{p}_{2\varepsilon}}{|\bar{p}_{2\varepsilon}|}(1 + \varepsilon|\bar{p}_{2\varepsilon}|) & \textit{if } |\bar{p}_{2\varepsilon}| \neq 0. \end{array} \right.$$

Proof. The existence and uniqueness of a solution of $\mathscr{P}_\varepsilon$ can be proved as in Proposition 3.1.

Theorem III.4.1 and Remark III.4.2 apply, condition III(4.21) being satisfied. We obtain (3.40) and the existence of a solution of $\mathscr{P}_\varepsilon^*$. Since $G_{2\varepsilon}^*$ is strictly convex, the solution of $\mathscr{P}_\varepsilon^*$ is unique. Finally (3.41) and (3.42) express the extremality relations

$$G_1(z(\bar{u}_\varepsilon)) + G_1^*(-\bar{p}_{1\varepsilon}) = -\langle z(\bar{u}_\varepsilon), \bar{p}_{1\varepsilon} \rangle$$
$$G_{2\varepsilon}(\operatorname{grad} z(\bar{u}_\varepsilon)) + G_{2\varepsilon}(-\bar{p}_{2\varepsilon}) = -\langle \operatorname{grad} z(\bar{u}_\varepsilon), \bar{p}_{2\varepsilon} \rangle. \quad ∎$$

The behaviour of the approximating problems when $\varepsilon \to 0$ is given by

Proposition 3.4. *If problem $\mathscr{P}$ possesses a solution $\bar{u}$, then when $\varepsilon \to 0$*

$$\inf \mathscr{P}_\varepsilon \to \inf \mathscr{P}, \tag{3.43}$$

$$\sup \mathscr{P}^*_\varepsilon \to \sup \mathscr{P}^*, \tag{3.44}$$

$$\bar{u}_\varepsilon \to \bar{u} \quad \textit{strongly in } V, \tag{3.45}$$

$$\bar{p}_{1\varepsilon} \to y(\bar{u}) - y_1 \quad \textit{strongly in } L^2(Q). \tag{3.46}$$

Proof. By definition of $\bar{u}_\varepsilon$,

$$G_\varepsilon(\Lambda \bar{u}_\varepsilon) \leqslant G_\varepsilon(\Lambda \bar{u}) = G(\Lambda \bar{u}), \tag{3.47}$$

and hence the sequence $y(\bar{u}_\varepsilon)$ is bounded in $L^2(Q)$ and the sequence $\bar{u}_\varepsilon$ is bounded in V.

By extracting a subsequence, we obtain

$$\bar{u}_\varepsilon \to u \quad \text{weakly in } V, \tag{3.48}$$

$$y(\bar{u}_\varepsilon) \to y(u) \quad \text{weakly in } L^2(Q), \tag{3.49}$$

(as Λ_1 is weakly continuous from V into $L^2(Q)$ weakly).

From (3.47)

$$\int_Q (|\text{grad } y(\bar{u}_\varepsilon)| - 1)_+^2 \, \mathrm{d}x \, \mathrm{d}t \leqslant 2\varepsilon G(\Lambda \bar{u}),$$

which implies by lower semi-continuity:

(3.50)

$$\int_Q (|\text{grad } y(u)| - 1)_+^2 \, \mathrm{d}x \, \mathrm{d}t \leqslant \liminf_{\varepsilon \to 0} \int_Q (|\text{grad } y(\bar{u}_\varepsilon)| - 1)_+^2 \, \mathrm{d}x \, \mathrm{d}t = 0.$$

Thus $y(u)$ satisfies (3.7). On the other hand,

$$\begin{aligned} J(u) = G_1(\Lambda_1 \bar{u}) \leqslant \liminf_{\varepsilon \to 0} G_1(\Lambda_1 \bar{u}_\varepsilon) &\leqslant \text{(by (3.47))} \\ &\leqslant G_1(\Lambda_1 \bar{u}) = G(\Lambda \bar{u}) \end{aligned} \tag{3.51}$$

and hence u is the solution of the problem (3.5) (3.6). As this solution is unique, $u = \bar{u}$ and the convergences (3.48) (3.49) hold for the whole sequence ε.

With (3.47) and (3.51) we again have

$$G_\varepsilon(\Lambda \bar{u}_\varepsilon) = \inf \mathscr{P}_\varepsilon \to G_1(\Lambda \bar{u}) = \inf \mathscr{P}$$

which yields (3.43) and, with (3.50),

$$G_{2\varepsilon}(\Lambda_2 \bar{u}_\varepsilon) \to 0,$$

$$G_1(\Lambda_1 \bar{u}_\varepsilon) = |y(\bar{u}_\varepsilon) - y_d|^2_{L^2(Q)} \to G_1(\Lambda_1 u) = |y(u) - y_d|_{L^2(Q)}. \tag{3.52}$$

With (3.49) and (3.52), we obtain

$$y(\bar{u}_\varepsilon) \to y(\bar{u}) \quad \text{strongly in } L^2(Q), \tag{3.53}$$

and as Λ_1 is an isomorphism of V in $L^2(Q)$, we deduce (3.45) from this.

Since the property (3.46) is a direct consequence of (3.41) and (3.53), it only remains for us to prove (3.44). To do so, we note that

$$\begin{aligned} \operatorname{Sup} \mathscr{P}^* &\geqslant - G^*(-\bar{p}_\varepsilon) \geqslant (\text{for } G_\varepsilon^* \geqslant G^*) \\ &\geqslant - G_\varepsilon^*(-\bar{p}_\varepsilon) = \operatorname{Sup} \mathscr{P}_\varepsilon^* \geqslant - G_\varepsilon^*(q) \end{aligned}$$

for $\forall q \in Y$ such that $q_1 = \operatorname{div} q_2$. When $\varepsilon \to 0$ we easily obtain $-G_\varepsilon^*(q) \to -G^*(q)$, and thus

$$\operatorname{Sup} \mathscr{P}^* \geqslant \limsup_{\varepsilon \to 0} [- G^*(-p_\varepsilon)] \geqslant \liminf_{\varepsilon \to 0} [- G^*(-p_\varepsilon)] \geqslant - G(q)$$

$$\operatorname{Sup} \mathscr{P}^* \geqslant \liminf_{\varepsilon \to 0} [\sup \mathscr{P}_\varepsilon^*] \geqslant \limsup_{\varepsilon \to 0} [\sup \mathscr{P}_\varepsilon^*] \geqslant - G(q),$$

for $\forall q \in Y$, $q_1 = \operatorname{div} q_2$. Then by taking the suprema for the q's under consideration, we obtain

$$- G^*(-p_\varepsilon) \to \operatorname{Sup} \mathscr{P}^*, \tag{3.54}$$

$$\operatorname{Sup} \mathscr{P}_\varepsilon = - G_\varepsilon^*(-p_\varepsilon) \to \operatorname{Sup} \mathscr{P}^*, \tag{3.55}$$

which is none other than (3.44). ∎

Remark 3.2. With (3.40), (3.43) and (3.44), we have proved (3.29) as we stated in Proposition 3.2. ∎

3.3. Approximation of the control problem

Apart from a change of sign, problem (3.28) can be written as

$$\operatorname*{Inf}_{r\in L^2(Q)^n} \int_Q \left[\frac{1}{2}|\operatorname{div} r|^2 - \operatorname{div} r \,.\, z_d + \frac{\varepsilon}{2}|r|^2 + |r| + r \,.\, \operatorname{grad} \varphi\right] \mathrm{d}x\, \mathrm{d}t,$$

where we have set $r = p_2$. We can also write

$$\operatorname*{Inf}_{r\in L^2(Q)^n} \quad \operatorname*{Sup}_{\substack{\lambda\in L^2(Q)^n \\ |\lambda(x)|\leqslant 1 \text{ a.e.}}} L(r, \lambda), \tag{3.56}$$

where

(3.57)

$$L(r, \lambda) = \int_Q \left[\frac{1}{2}|\operatorname{div} r|^2 - \operatorname{div} r \,.\, z_d + \frac{\varepsilon}{2}|r|^2 + r \operatorname{grad} \varphi + r \,.\, \lambda\right] \mathrm{d}x\, \mathrm{d}t.$$

By replacing V by $\{r \in L^2(Q)^n | \operatorname{div} r \in L^2(Q)\}$ and $Z = L^2(Q)$, we satisfy the conditions for the application of Uzawa's algorithm of Section 1;[1] we set r and λ instead of u and p, $\mathscr{A} = V$

$$\mathscr{B} = \{\, \lambda \in L^2(Q)^n \mid |\lambda(x)| \leqslant 1 \ \text{ a.e.}\}.$$

We start with $\lambda^0 \in \mathscr{B}$; when λ^m is known, we determine r^m as the element of $L^2(Q)^n$, such that $\operatorname{div} r \in L^2(Q)$ and which minimizes $L(r, \lambda^m)$. Next we set

$$\lambda^{m+1} = \Pi_{\mathscr{B}}(\lambda^m + \rho_m r^m).$$

As the selected ρ_n satisfy

$$0 < \rho_* < \rho_m < \rho'_*, \qquad \text{with suitable } \rho_*, \rho'_*$$

we have from Proposition 1.1 and the above:

$$\begin{aligned} &r^m \to \bar{r} = \bar{p} \quad \text{ in } \quad L^2(Q)^n, \\ &\operatorname{div} r^m \to \operatorname{div} \bar{r} = \operatorname{div} \bar{p}_{2\varepsilon} = \bar{p}_{1\varepsilon} \quad \text{ in } \quad L^2(Q), \end{aligned}$$

when $m \to \infty$.

Then for m large and ε sufficiently small, $\operatorname{div} r^m$ tends to $y(\bar{u}) - y_d$ in the norm of $L^2(Q)$. When we know $y(\bar{u})$, "the optimal state", we deduce from (3.1) the optimal control $\bar{u}$.

[1] Here we use the fact that $\varepsilon > 0$. Condition (1.7) is not satisfied if $\varepsilon = 0$.

The process is clearly very simple since the calculation of r^m and λ^{m+1} is very economical.

Remark 3.3. We can show that problem $\mathscr{P}^*$ allows of a generalized solution $\bar{p}$, with $\bar{p}_1 \in L^2(Q)$, $\bar{p}_2$ = bounded measure on Ω with values in $\mathbf{R}^n$, $\bar{p}_1 = \operatorname{div} \bar{p}_2$. When $\varepsilon \to 0$, we have $\bar{p}_{1\varepsilon} \to \bar{p}_1$ strongly in $L^2(Q)$, any limit point $\bar{p}_{2\varepsilon}$ being a possible value of $\bar{p}_2$ ($\bar{p}_2$ is not unique, $\bar{p}_1 = \operatorname{div} \bar{p}_2$ is unique).

Remark 3.4. Many applications of duality to optimal control problems for systems governed by partial differential equations are studied in J. Mossino [1]. For the numerical approximation of these problems, see J. Mossino [2].

4. APPLICATIONS OF DUALITY IN MECHANICS

Orientation

We shall show by some simple examples that the fundamental principles of elasticity termed principle of potential energy and principle of complementary energy are in duality. The primal problem has as its solution the displacement field. The solution of the dual problem is the field of constraints. ■

In the following section we shall apply the considerations of Chapter III, Section 4. With the usual functional context, the primal problem can be written as

$$\operatorname*{Inf}_{u \in V} \{ F(u) + G(\Lambda u) \}. \tag{4.0}$$

The results of Section III.4 were obtained by considering the perturbation function

$$\Phi(u, p) = F(u) + G(\Lambda u - p).$$

We can equally well consider the perturbation function

$$\Phi(u, p) = F(u) + G(\Lambda u + p).$$

This only results in a few changes of sign. The dual problem of (4.0) relative to these perturbations can be written as

$$\operatorname*{Sup}_{p^* \in Y^*} \{ - F^*(- \Lambda^* p^*) - G^*(p^*) \} \tag{4.1}$$

and when they have been satisfied, the extremality relations can be written as

$$F(\bar{u}) + F^*(-\Lambda^*\bar{p}^*) = -\langle \Lambda^*\bar{p}^*, \bar{u} \rangle \tag{4.2}$$

$$G(\Lambda\bar{u}) + G^*(\bar{p}^*) = \langle \bar{p}^*, \Lambda\bar{u} \rangle. \tag{4.3}$$

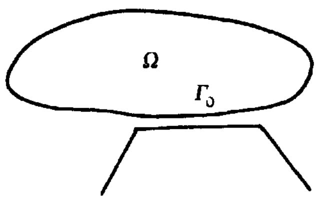
Theorems III.4.1 and III.4.2 and Proposition III.4.1 are easily adapted to this situation.

Owing to this modification, the solution of the dual problem will have a more standard mechanical meaning. ■

4.1. Equilibrium of an elastic body on a rigid base

We consider an elastic body filling a set Ω of $\mathbf{R}^3$ and in contact with a rigid base on a subset Γ_0 of $\partial\Omega$. We assume the contact to be bilateral, which means that the contact takes place effectively over all Γ_0. We denote by $\Gamma_1 = \Gamma - \Gamma_0$, $\Gamma = \partial\Omega$.

The body is subject to some forces applied to the boundary Γ_1 denoted g (e.g. $g \in L^2(\Gamma_1)^3$) and to the body forces f ($f \in L^2(\Omega)^3$).



The space V is the space of displacement fields v, *one of the unknowns of the problem being the displacement field $\bar{u} \in V$ for the equilibrium position of the body Ω.* Here we shall take

$$V = \{ v \in H^1(\Omega)^3 \mid \gamma_0 v = 0 \ \text{sur}\ \Gamma_0 \}, \tag{4.4}$$

which is a Hilbert space for the usual norm.

For a given displacement $v \in V$, the internal energy of the body is $G(\Lambda v)$ where

$$\Lambda v = \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right) \tag{4.5}$$

and

(4.6) $\sigma \mapsto G(\sigma)$ is a strictly convex l.s.c. function coercive on $L^2(\Omega)^2$, finite and continuous at 0.

The energy corresponding to the applied forces is

$$-(f, v)_{L^2(\Omega)^3} - (g, \gamma_0 v)_{L^2(\Gamma_1)^3} \tag{4.7}$$

and the total potential energy can be written as

$$\{ F(v) + G(\Lambda v) \}. \tag{4.8}$$

According to the principle of Virtual Work, the displacements from equilibrium realizes the minimum of the potential energy among the v's of V: this is the primal problem.

Let us specify the dual problem (4.1). We set

$$Y = Y^* \subset L^2(\Omega)^6,$$

and the elements of Y are the symmetric matrices

$$\sigma = (\sigma_{ij}), \qquad \sigma_{ij} = \sigma_{ji} \in L^2(\Omega).$$

The space Y is a closed subspace of $L^2(\Omega)^6$.

Lemma 4.1. *For $\sigma \in Y$*

$$F^*(-\Lambda^*\sigma) = \left| \begin{array}{l} 0 \quad \textit{if} \;\; \dfrac{\partial \sigma_{ij}}{\partial x_j} + f_i = 0, \text{[1]} \\ +\infty \quad \textit{otherwise}. \end{array} \right. \tag{4.9}$$

Proof.

$$\begin{aligned} F^*(-\Lambda^*\sigma) &= \operatorname*{Sup}_{v \in V} [-\langle \sigma, \Lambda v \rangle - (f, v) - (g, \gamma_0 v)] \\ &\geqslant \operatorname*{Sup}_{v \in \mathscr{D}(\Omega)^3} [-\langle \sigma, \Lambda v \rangle - (f, v)]. \end{aligned}$$

Now

$$-\langle \sigma, \Lambda v \rangle - (f, v) = \langle \frac{\partial \sigma_{ij}}{\partial x_j} - f_i ; v_i \rangle.$$

The latter supremum (and *a fortiori* the former) equals $+\infty$ unless

$$\frac{\partial \sigma_{ij}}{\partial x_j} + f_i = 0. \tag{4.10}$$

[1] In this section we use the repeated index summation convention.

When (4.7) is satisfied, the vector $\theta_i = (\sigma_{i1}, \sigma_{i2}, \sigma_{i3}) \in L^2(\Omega)^3$ and div $\theta \in L^2(\Omega)$; we can thus define the trace on Γ of

(4.11) $$\theta_i \nu = \sigma_{ij}\nu_j \in H^{-1/2}(\Gamma),$$

and we have Green's theorem (*cf.* Lions and Magenes [1])

(4.12)
$$\int_\Omega \operatorname{div} \theta_i \, . \, w \, dx + \int_\Omega \theta_i \operatorname{grad} w \, dx = \langle \theta_i \nu, \gamma_0 w \rangle, \qquad \forall w \in H^1(\Omega).$$

Then for σ satisfying (4.7)

$$\langle \sigma, \Lambda v \rangle = \langle \sigma_{ij}\nu_j, v_i \rangle - (f, v),$$

and

$$F^*(-\Lambda^*\sigma) = \operatorname*{Sup}_{v \in V} \left[- \langle \sigma_{ij}\nu_j, v_i \rangle - \langle g_i, v_i \rangle \right]$$

and this supremum equals $+\infty$ unless

(4.13) $$\sigma_{ij}\nu_j + g_i = 0 \text{ on } \Gamma_1$$

in which case the supremum is zero (we note that $v = 0$ on Γ_0). ∎

The dual problem can be written as

(4.14) $$\operatorname{Sup} \{ -G^*(\sigma) \}$$

among the $\sigma \in Y$ satisfying (4.10) and (4.13). This problem coincides with the principle of complementary energy: in equilibrium the constraints achieve the maximum complementary energy among all the fields of admissible kinematic constraints (i.e. satisfying (4.7) and (4.10)).

Proposition 4.1. *With the above hypotheses (and in particular hypothesis* (4.6) *for G), problem* (4.8) *possesses a unique solution $\bar{u}$. The dual problem* (4.14) *possesses at least one solution $\bar{\sigma}$; we have*

$$\inf \mathscr{P} = \sup \mathscr{P}^*$$

and $\bar{u}$ and $\bar{\sigma}$ are linked by the extremality relation:

(4.15) $$G(\Lambda\bar{u}) + G^*(\bar{\sigma}) = \langle \bar{\sigma}, \Lambda\bar{u} \rangle.$$

Proof. The existence and uniqueness of the solution of (4.8) is easily shown: the minimizing functional on V is strictly convex, l.s.c. and coercive; the coerciveness results from Korn's inequality (see Lions and Magenes [1]):

$$\|v\|_V \leqslant c \, \|\Lambda v\|_Y \quad , \qquad \forall v \in V. \tag{4.16}$$

We have $\inf \mathscr{P} = \sup \mathscr{P}^*$ and the existence of $\bar{\sigma}$ because of (4.6) and of Theorem III.4.1. Finally, (4.15) is none other than (4.3).

Remark 4.1. In the special case of linear elasticity,

$$G(\varepsilon) = \frac{1}{2} \int_\Omega a_{ijhk} \varepsilon_{ij} \varepsilon_{hk} \, \mathrm{d}x$$

where $_{ijhk} = a_{ijkh} = a_{hkij}, \in L^\infty(\Omega)$ and

$$a_{ijhk} \varepsilon_{ij} \varepsilon_{hk} \geqslant \alpha \varepsilon_{ij} \varepsilon_{ij} \, .$$

In this case relation (4.15) is

$$\bar{\sigma}_{ij} = a_{ijhk} \varepsilon_{hk}(\bar{u}), \tag{4.17}$$

where

$$\varepsilon_{hk}(\bar{u}) = \Lambda \bar{u},$$

and (4.17) is the elastic law for the behaviour of the material.

Generally, relation (4.15) when solved for $\bar{\sigma}$,

$$\bar{\sigma} = \Psi(\Lambda \bar{u}), \tag{4.18}$$

is the relation for the behaviour of the material. ■

Remark 4.2. We can associate with problems $\mathscr{P}$ and $\mathscr{P}^*$ different Lagrangian and different saddle point problems for which $(\bar{u}, \bar{\sigma})$ is the solution. For this see Fremond [2]. ■

4.2. Case of a unilateral contact

When the contact between Ω and the base is unilateral, which means that it can be broken on a subset of Γ_0, the principle of Virtual Work tells us that the field of displacements at the equilibrium minimize the potential energy among the $v \in H^1(\Omega)^3$ satisfying

$$v \, . \, \nu \leqslant 0 \quad \text{on } \Gamma_0. \tag{4.19}$$

We now take $V = H^1(\Omega)^3$, $Y = Y^*$ as before,

$$F(v) = - \langle f, v \rangle - \langle g, \gamma_0 v \rangle + \chi_K(v),$$

where χ_K is the indicator function of the convex set

$$K = \{ v \in H^1(\Omega)^3 \mid v \, . \, \nu \leqslant 0 \text{ on } \Gamma_0 \} ;$$

Λ and G are unchanged.

Lemma 4.2. *For $\sigma \in Y$*

$$(4.20) \qquad F^*(-\Lambda^*\sigma) = \left| \begin{array}{ll} 0 & \textit{if } (4.10) \textit{ and } (4.13) \textit{ hold and} \\ \hat{\sigma}_\tau = 0 & \textit{and} \quad \hat{\sigma}_\nu \leqslant 0 \textit{ on } \Gamma_0, \\ +\infty & \textit{otherwise}, \end{array} \right.$$

where $\hat{\sigma}$ is the vector $\hat{\sigma}_{ij}\nu_j$, $\hat{\sigma}_\nu$ and $\hat{\sigma}_\tau$ its normal and tangential components.

Proof.

$$F^*(-\Lambda^*\sigma) = \operatorname*{Sup}_{\substack{v \in H^1(\Omega) \\ v . \nu < 0 \text{ on } \Gamma_0}} [- \langle \sigma, \Lambda v \rangle - (f, v) - (g, \gamma_0 v)].$$

As for Lemma 4.1 this supremum is $+\infty$ if (4.10) and (4.13) are not satisfied. When σ satisfies (4.10) and (4.13), there remains

$$F^*(-\Lambda^*\sigma) = \operatorname*{Sup}_{\substack{v \in H^1(\Omega) \\ v . \nu \leqslant 0 \text{ on } \Gamma_0}} [- \langle \hat{\sigma}_\tau, v_\tau \rangle - \langle \hat{\sigma}_\nu, v_\nu \rangle]$$

where $v_\nu = (v \cdot \nu)$ and $v_\tau = v - v_\nu$. The Lemma follows. ∎

The primal problem being,

$$(4.21) \qquad \operatorname*{Inf}_{v \in V} \{ F(v) + G(\Lambda v) \}$$

the dual problem can be written as

$$(4.22) \qquad \operatorname{Sup} \{ - G^*(\sigma) \}$$

for $\sigma \in Y$ and satisfying (4.10), (4.13) and (4.23)

$$(4.23) \qquad \hat{\sigma}_\tau = 0, \qquad \hat{\sigma}_\nu \leqslant 0 \quad \text{on } \Gamma_0.$$

It is the principle of complementary energies which is expressed by (4.22).

For the existence of a solution, since Korn's inequality is not true for $\forall v \in H^1(\Omega)$, we have to make the following hypothesis (*cf.* Fremond [1]):

$$(4.24) \qquad \text{Let} \quad U = \{ v \in V \mid v = a + b\Lambda x, \gamma v \leqslant 0, v \neq 0 \}.$$

We have

$$\langle f, v \rangle + \langle g, \gamma_0 v \rangle \leqslant 0, \qquad \forall v \in U.$$

We can then prove

Proposition 4.2. *With the above hypotheses, especially* (4.6) *and* (4.24), *problem* (4.21) *possesses a unique solution* $\bar{u}$, *and problem* (4.22) *has a solution* $\bar{\sigma}$; *we have*

$$\inf \mathscr{P} = \sup \mathscr{P}^* \tag{4.25}$$

and $\bar{u}$ *and* $\bar{\sigma}$ *are linked by the extremality relation*

$$G(\Lambda\bar{u}) + G^*(\bar{\sigma}) = \langle \bar{\sigma}, \Lambda\bar{u} \rangle. \tag{4.26}$$

Proof. The only new feature is the coerciveness of $F(v) + G(\Lambda v)$ on K; for this we refer to Fremond [1]. ∎

5. APPLICATIONS IN ECONOMICS

We shall be interested in economic activities concerning goods. The amount of each article is measured by a real number so that to each vector $x = (x_1, \ldots, x_m)$ of $\mathbf{R}^m_+$ there is associated the quantity x_i of article i. The *price* of article i is fixed at π^*_i, so that the sum to be paid to acquire $x \in \mathbf{R}^m_+$ is

$$\langle \pi^*, x \rangle = \sum_{i=1}^{m} \pi^*_i x_i.$$

Firms produce certain of these articles to the exclusion of others; in this context, where they do not influence the price and when they are sure of being able to dispose of all their production, they merely seek to maximize their profit. We are in the domain of micro-economics.

5.1. Micro-economic theory of the factory

For all $x \in \mathbf{R}^m$ we term x^+ the vector with components $x^+_i = \max\{x_i, 0\}$ and x^- the vector with components $x^-_i = -\min\{x_i, 0\}$ so that:

$$x = x^+ - x^-. \tag{5.1}$$

We shall characterize a factory by its production set $Y \subset \mathbf{R}^m$. To say that $y \in Y$ means that the factory is in a position to produce articles y^+ while consuming articles y^-. Thus, for example, to say that $0 \in Y$ means that inaction is possible, and to say that $\mathbf{R}^m_- \subset Y$ means that it is possible to destroy any quantity of goods. We usually make the following hypotheses:

$$Y \text{ is a closed convex set containing } \mathbf{R}^m_-. \tag{5.2}$$

To a product $y \in Y$ costing $\sum_{i=1}^m \pi^* y^-$ and giving as return $\sum_{i=1}^m \pi^* y_i^+$ is attached a profit:

$$\sum_{i=1}^m \pi_i^* y_i^+ - \sum_{i=1}^m \pi_i^* y_i^- = \langle \pi^*, y \rangle . \tag{5.3}$$

Let us assume that c_i is the maximum quantity of article i that the company can buy on the market at price π_i^*. The vector $c = (c_1, \ldots, c_m)$ thus constitutes the *initial resources*. The company seeks that production $\bar{y}$ which allows it to maximize its profit without exceeding its initial resources. This is the solution, if it exists, of the optimization problem:

$$(\mathscr{P}) \quad \left| \begin{array}{l} \sup\limits_{y \in Y} \langle \pi^*, y \rangle \\ y_i \geqslant -c_i \quad \text{for} \quad i = 1, \ldots, m. \end{array} \right.$$

Let us recall the method of Chapter III. To each $p = (p_1, \ldots, p_m) \in \mathbf{R}^m$ we associate the perturbation problem.[1]

$$(\mathscr{P}_p) \quad \left| \begin{array}{l} \sup\limits_{y \in Y} \langle \pi^*, y \rangle \\ y_i \geqslant -c_i - p_i , \end{array} \right.$$

and the concave function:

$$h(p) = \sup (\mathscr{P}_p). \tag{5.4}$$

Thus $h(p)$ is the new profit which the company can realize if its initial resources are improved by p. Proposition 5.1 of Chapter III then shows us that if

$$Y \cap \{ -c + \mathring{\mathbf{R}}^m_+ \} \neq \varnothing \tag{5.5}$$

then the problem $(\mathscr{P})$ is stable with respect to the perturbations under consideration. Condition 5.5 is satisfied if, for example, $c_i > 0$ for all i, that is all the articles are initially disposable in non-zero quantities.

The dual problem can be written as (Chapter III(5.23) and note [1] before):

$$(\mathscr{P}^*) \qquad \operatorname*{Inf}_{p^* \geqslant 0} \operatorname*{Sup}_{y \in Y} [\langle \pi^* + p^*, y \rangle + \langle p^*, c \rangle].$$

[1] Here, as in Section 4 before (see (4.1)–(4.3)), we change p in $-p$, so that the "prices" appearing as the solutions of the dual problem will be $\geqslant 0$.

Under hypothesis (5.5) problem ($\mathscr{P}$) is stable, and thus the problem ($\mathscr{P}^*$) possesses solutions and these solutions $\bar{p}^*$ are the over-gradients of h at 0 (Prop. 2.2, Chap. III)

$$-\ p^* \in \partial(-\ h)(0). \tag{5.6}$$

The economic interpretation can be developed more easily by assuming that h is Gâteaux-differentiable at 0 and that the problem ($\mathscr{P}$) has an optimal solution $\bar{y}$. Equation (5.6) then becomes:

$$p_i^* = \frac{\partial h}{\partial p_i}(0) \qquad \text{for} \quad i = 1, ..., n. \tag{5.7}$$

In other terms, $-p_i^*$ is the marginal cost of the ith constraint: if it changed from c_i to $c_i + p_i$, the profit would be improved by $p_i^* p_i$ to first order. In an equivalent way, if a small additional quantity p_i of article i can be disposed of at price π_i^*, the profit will be increased by $\pi_i^* p_i$. Let us state this latter point more explicitly.

The company produces $\bar{y}^+$ while consuming $\bar{y}^-$, and this activity uses to the best advantage the resources c which are available at price π^*. Let us then assume that it is offered to it additional quantities of article i at the price $\pi_i^* + \lambda_i^*$. Should it be interested in acquiring a quantity $p_i > 0$? This reduces to introducing the additional quantity p_i in the full market against a payment of $\lambda_i^* p_i$. The maximum profit realizable will be $h(p_i) - \lambda_i^* p_i$, which must be compared with $h(0)$. Because of (5.7) and the concavity of h, we obtain:

$$\exists p_i > 0 : h(p_i) - \lambda_i^* p_i > h(0) \Leftrightarrow \lambda_i^* < p_i^*. \tag{5.8}$$

The company will only buy if $\lambda_i^* < p_i^*$. Thus, $(\pi_i^* + p_i^*)$ represents the threshold below which the company proceeds to buy article i. We immediately deduce that:

if $\bar{y}_i > -c_i$, there remain in the market some unused quantities of article i at the price π_i^*. Therefore π_i^* is greater than or equal to the intervention price $\pi_i^* + p_i^*$, and so $p_i^* = 0$;

if $p_i^* > 0$, the market price is lower than the intervention price $\pi_i^* + p_i^*$. Then the company wishes to buy additional quantities of article i at price π_i^*. If it does not do so, that is if $\bar{y}_i$ is optimal, that means that there is a shortage, and so $\bar{y}_i = c_i$.

We recognize the extremality relations which have been recovered by purely economic reasoning. The economic interpretation of the multiplier $\bar{p}^*$ does not stop there: there is also a decentralization effect which we shall see better by enlarging the framework.

5.2. Decentralized management of a firm

Let us consider a firm which comprises N factories. The production set of enterprise n is Y_n, and the profit which it takes from the product $y \in Y_n$ is $u_n(y)$. Unlike in (5.3), we assume that various phenomena (saturation, taxation, costs) prevent the linear growth of profit with production. Our hypotheses are that, for all $n \in \{1,\ldots,N\}$:

(5.9) $\quad Y_n$ is a closed convex set containing $\mathbf{R}^n_-$,

(5.10) $\quad$ the function π_n: $Y_n \to \mathbf{R}$ is concave and u.s.c.

As previously, we denote by $c = (c_1,\ldots,c_n)$ the initial resources; thus the quantity of article i disposable on the market is limited to c_i. But some enterprises of the firm can produce article i and supply it to others. In total, the firm can assign to the enterprise n the product $y(n) \in Y_n$ under the condition that, globally, these constraints are satisfied:

$$(5.11) \qquad \sum_{n=1}^{N} y_i(n) \geqslant -c_i \qquad \text{for} \qquad i = 1,\ldots,m.$$

The firm seeks to determine $\bar{y}(1),\ldots,\bar{y}(N)$ so as to maximize its total profit, that is the sum of the profits of its members. We must solve the optimization problem:

$$(\mathscr{P}) \qquad \left| \begin{array}{l} \operatorname*{Sup}_{y(n)\in Y_n} [u_1(y(1)) + \ldots + u_N(y(N))] \\ \sum_{n=1}^{N} y_i(n) \geqslant -c_i \quad \text{for} \quad i = 1,\ldots,m. \end{array} \right.$$

With each $p = (p_1,\ldots,p_m) \in \mathbf{R}^m$ we associate the perturbed problem

$$(\mathscr{P}_p) \qquad \left| \begin{array}{l} \operatorname*{Sup}_{y(n)\in Y_n} [u_1(y(1)) + \ldots + u_N(y(N))] \\ \sum_{n=1}^{N} y_i(n) \geqslant -c_i - p_i \quad \text{for} \quad i = 1,\ldots,m. \end{array} \right.$$

From Proposition 5.1 of Chapter III, if

$$(5.12) \qquad \sum_{n=1}^{N} Y_n \cap \{-c + \mathring{\mathbf{R}}^m\} \neq \emptyset$$

(if, for example, $c_i > 0$ for all i), then the problem $(\mathscr{P})$ is stable relative to the perturbations under consideration. The dual problem can be written as

$$(\mathscr{P}^*) \qquad \operatorname*{Inf}_{p^*\geqslant 0} \operatorname*{Sup}_{\substack{y(n)\in Y_n \\ 1\leqslant n\leqslant N}} \left[\sum_{n=1}^{N} u_n(y(n)) + \left\langle p^*, \sum_{n=1}^{N} y(n) + c \right\rangle \right].$$

Under hypothesis (5.12) the problem ($\mathscr{P}^*$) possesses some solutions $\bar{p}^*$. From Proposition 5.3 of Chapter III, the problem ($\mathscr{P}$) and the problem

$$\operatorname*{Sup}_{\substack{y(n)\in Y_n\\ 1\leqslant n\leqslant N}} \left[\sum_{n=1}^{N} u_n(y(n)) + \left\langle \bar{p}^*, \sum_{n=1}^{N} y(n) + c \right\rangle \right]$$

have the same solutions. To simplify the interpretation, we assume that there exists a unique solution $(\bar{y}(1),\ldots,\bar{y}(n))$ (which takes place if, for example, the u_n are strictly concave or the Y_n are strictly convex).

We rewrite the latter problem in a more convenient form by omitting the constant $\langle p^*, c\rangle$ which does not interfere with the optimization:

$$(Q) \qquad \operatorname*{Sup}_{\substack{y(n)\in Y_n\\ 1\leqslant n\leqslant N}} \sum_{n=1}^{N} [u_n(y(n)) + \langle \bar{p}^*, y(n) \rangle].$$

But the problem (Q) is the sum of N independent problems; it can be also written:

$$(Q) \qquad \sum_{n=1}^{N} \operatorname*{Sup}_{y(n)\in Y_n} [u_n(y(n)) + \langle \bar{p}^*, y(n) \rangle].$$

The problems ($\mathscr{P}$) and (Q) have the same solution. This means that two attitudes can be adopted by the firm:

it can determine for itself $(\bar{y}(1),\ldots,\bar{y}(n))$ by solving the problem ($\mathscr{P}$) directly;

it can restrict itself to calculating $\bar{p}^*$, and let enterprise n have the responsibility of determining $\bar{y}(n)$ by solving the problem:

$$(Q_n) \qquad \operatorname*{Sup}_{y(n)\in Y_n} [u_n(y(n)) + \langle \bar{p}^*, y(n) \rangle].$$

In this second case, the decision is decentralized to the extent where each factory determines for itself its own production without having to concern itself either with the other factories or with the initial resources. The firm achieves this decentralization by an accounting artifice: it imposes on its factories shadow prices $\bar{p}_i^*$ for the articles i, $1 \leqslant i \leqslant m$. The production (or the consumption) of one unit of article i must produce a shadow income of $\bar{p}_i^*$ (or a shadow expenditure of $-\bar{p}^*$), to be added to the real profit. The object of factory n is to maximize the sum:

$$\left| \begin{array}{l} \text{real profit} + \text{shadow profit} \\ u_n(y(n)) + \langle p^*, y(n) \rangle \end{array} \right.$$

which leads to the problem (Q_n). It has a unique solution which is indeed that $\bar{y}(n) \in Y_n$. If the choice of shadow prices $\bar{p}^*$ has been properly made, that is if it is effectively a solution of $(\mathscr{P}^*)$, the individual components of the N enterprises contribute to the communal wealth, that is that the total demand for each article does not exceed the available assets ($\sum_{n=1}^N y_i(n) \geqslant -c_i$ for all i) and that the total (real) profit is at a maximum.

It is well known that the difficulty consists in effectively calculating the shadow prices $\bar{p}^*$. We can, for example, proceed in the following way. The firm states a system of shadow prices $p^{*0} \in \mathbf{R}^m$. Each enterprise n then states that its production will be $y(n)$ at price p^{*0}. The firm then calculates $\sum_{n=1}^N y(n)$; the negative components constitute the total demand which the firm will compare with the available assets c. If

$$- \sum_{n=1}^{N} y_i(n) - c_i \geqslant 0,$$

there is excess demand over supply and the firm increases the shadow price p_i^{*0} by a quantity proportional to this excess. If

$$- \sum_{n=1}^{N} y_i(n) - c_i \leqslant 0,$$

there is an excess of supply over demand and the firm diminishes the shadow price p_i^{*0} by a quantity proportional to this excess—always taking care not to make the price negative. The firm then states the new pricing system p^{*1} (we denote by ρ a constant > 0):

$$\text{if } \sum_{n=1}^{N} y_i(n) + c_i \geqslant 0, \qquad p_i^{*1} = \operatorname{Sup}\left[0, p_i^{*0} - \rho\left(c_i + \sum_{n=1}^{N} y_i(n)\right)\right]$$

$$\text{if } \sum_{n=1}^{N} y_i(n) + c_i \leqslant 0, \qquad p_i^{*1} = p_i^{*1} - \rho\left(c_i + \sum_{n=1}^{N} y_i(n)\right).$$

And we start again. Thus a dialogue arises between the centre (the firm) and the periphery (the factories), as a result of which we obtain a sequence p^{*0}, $p^{*1}, \ldots, p^{*n}, \ldots$ of systems of shadow prices. This procedure is standard in economic theory under the name of Walras "tâtonnements". The reader will have recognized Uzawa's algorithm; we set out in Section 1 of this chapter some conditions under which the sequence p^{*n} converges to $\bar{p}^*$, the solution of $(\mathscr{P}^*)$, which is the required system of shadow prices. ■

PART THREE

Relaxation and Non-convex Variational Problems

This page intentionally left blank

CHAPTER VIII

Existence of Solutions for Variational Problems

Orientation

In this chapter, we shall study *non-convex* problems. In Section 1, we shall introduce normal (not necessarily convex) integrands, an important class of functions of two variables without any convexity; we shall establish their main properties, including a measurable selection theorem, and we shall recall the characterization of weakly relatively compact subsets of L^1.

In Section 2, these results will be applied to the study of a non-convex optimization problem, and a sufficient condition for the existence of solutions will be given. The final sections will show that a number of problems in the calculus of variations (Section 3) and in optimal control (Section 4) can be put into the above form and from this we deduce theorems on the existence of solutions.

1. NON-CONVEX NORMAL INTEGRANDS

1.1. Definition and main property

We recall that the Borel σ-algebra of a topological space is the σ-algebra generated by the closed subsets; in other words, the Borel subsets are the sets obtained from open and closed subsets by denumerable union, denumerable intersection, complementation and by any denumerable combination of such operations. A mapping f into $\bar{\mathbf{R}}$ will be called a Borel function if $f^{-1}(F)$ is Borel for every closed set F. For instance, continuous functions are Borel.

Let Ω be an open subset of $\mathbf{R}^n$ provided with the Lebesgue measure. All the Borel subsets of Ω, especially the open and closed sets, are measurable. We recall that a mapping $f: \Omega \to \bar{\mathbf{R}}$ is measurable if the inverse image under f of every closed subset of $\bar{\mathbf{R}}$ is measurable. The Borel functions are measurable and, in general, we have at our disposal the following criterion for measurability:

Lusin Theorem. *A function $f: \Omega \to \bar{\mathbf{R}}$ is measurable if and only if, for every compact set $K \subset \Omega$ and all $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset K$ such that* $\operatorname{meas}(K - K_\varepsilon) \leqslant \varepsilon$ *for which the restriction of f to K_ε is continuous.*

The integrands used in the calculus of variations are functions of several variables playing different roles; by and large, they are measurable with respect to some variables and l.s.c. with respect to others.

Definition 1.1. *If B is a Borel subset of* $\mathbf{R}^p$, *a mapping f of $\Omega \times B$ into* $\overline{\mathbf{R}}$ *is termed a normal integrand if*:

(1.1) *for almost all $x \in \Omega$, $f(x, .)$ is l.s.c. on B*

(1.2) *there exists a Borel function $\tilde{f}: \Omega \times B \to \overline{\mathbf{R}}$ such that $\tilde{f}(x, .) = f(x, .)$ for almost all $x \in \Omega$.*

A first consequence of this definition is that for all $a \in B, f(., a)$ is measurable on Ω. Better still, if u is a measurable mapping of Ω into B, the function $x \mapsto f(x, u(x))$ is measurable on Ω. Indeed, it is almost everywhere equal to the function $x \to \tilde{f}(x, u(x))$, which is measurable since $\tilde{f}$ is Borel. We note that this property is no longer satisfied if, instead of assuming that f is a normal integrand, we merely assume it to be measurable in x and l.s.c. in a. Later on, we shall see that functions measurable in x and continuous in a are normal integrands. Furthermore, it follows from the definition that:

if f is a normal integrand, λf is a normal integrand for all $\lambda \in \mathbf{R}$;

if f and g are normal integrands, $(f + g)$ and $\inf(f, g)$ are normal integrands;

if $(f_n)_{n \in \mathbf{N}}$ is a denumerable family of normal integrands, $\sup_{n \in \mathbf{N}} \tilde{f}_n$ is a normal integrand.

The study of normal integrands depends on the following characterization, which is a Lusin theorem "uniform" in the second variable:

Theorem 1.1. *Let B be a Borel subset of* $\mathbf{R}^p$. *For $f: \Omega \times B \to \overline{\mathbf{R}}$ to be a normal integrand, it is necessary and sufficient that for every compact set $K_\varepsilon \subset \Omega$ and all $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset K$ such that meas$(K - K_\varepsilon) \leqslant \varepsilon$ for which the restriction of f to $K_\varepsilon \times B$ is l.s.c.*

Proof. The sufficient condition is obvious: we take $\varepsilon = 1/n$ and let $K' = \bigcup_{n \in \mathbf{N}} K_{1/n}$. Then meas$(K - K') = 0$, K' is Borel, f is Borel on $K' \times B$ and $f(x, .)$ is l.s.c. for all $x \in K'$.

To show the necessary condition, we begin by modifying f on a Borel set of Ω with null measure so that $f(x, .)$ becomes l.s.c. for all $x \in \Omega$ and f becomes Borel over all of $\Omega \times B$. Using if necessary an isomorphism of $\overline{\mathbf{R}}$ onto $[0, 1]$, we may assume that f takes its values in $[0, 1]$.

Since B is a subspace of $\mathbf{R}^p$, it possesses a denumerable basis $\mathcal{U}$ of open subsets. Let us denote by Φ the family of l.s.c. functions of B into $[0, 1]$ defined by

$\Phi = \{ k \; 1_u | \, u \in \mathcal{U}, k \in \mathbf{Q} \, ; \, 0 \leqslant k \leqslant 1 \}$, where $\mathbf{Q}$ is the set of rationals.

Clearly Φ is denumerable and for all l.s.c. functions $h: B \to [0,1]$ we have $h = \sup\{\varphi \mid \varphi \in \Phi \text{ and } \varphi \leqslant h\}$. Setting $\Phi = \{\varphi_n\}_{n\in\mathbf{N}}$, then:

$$E_n = \{ x \in \Omega \mid f(x, .) \geqslant \varphi_n(.) \},$$
$$G_n = \{ (x, a) \in \Omega \times B \mid f(x, a) < \varphi_n(a) \}.$$

Since f is Borelian and φ_n is l.s.c., G_n is a Borel subset of $\Omega \times B$. Its projection on Ω which is none other than the complement of E_n is thus measurable.[1] Hence E_n is measurable in Ω. Now for all $x \in \Omega, f(x, .)$ is a l.s.c. mapping of B into $[0,1]$

$$f(x, a) = \sup_{n\in\mathbf{N}} \varphi_n(a) \mathbf{1}_{E_n}(x).$$

If the compact set $K \subset \Omega$ and the number $\varepsilon > 0$ are given, we choose for all $n \in \mathbf{N}$, by virtue of Lusin's Theorem, a compact space $K_n \subset K$ such that $\operatorname{meas}(K - K_n) \leqslant \varepsilon 2^{-(n+1)}$ and that the restriction of $\mathbf{1}_{E_n}$ to K_n be continuous. Let $K_\varepsilon = \bigcup_{n\in\mathbf{N}} K_n$. Then $\operatorname{meas}(K - K_\varepsilon) \leqslant \varepsilon$ and the restriction of $\varphi_n \mathbf{1}_{\varepsilon_n}$ to $K_\varepsilon \times B$ is l.s.c. for all n. Since f is the least upper bound of the $\varphi_n \mathbf{1}_{\varepsilon_n}$ its restriction to $K_\varepsilon \times B$ is l.s.c. ∎

1.2. First example: indicator function of a varying closed set

We give here an important example of a *positive* normal integrand. Let C be a Borel subset of $\Omega \times B$; we shall assume that for almost all $x \in \Omega$, the trench $C_x = \{a \in B \mid (x,a) \in C\}$ is closed in B. Then the *indicator function* f of C:

$$f(x, a) = 0 \quad \text{if} \quad (x, a) \in C,$$
$$f(x, a) = +\infty \quad \text{if} \quad (x, a) \notin C,$$

is a positive normal integrand since it is Borelian and trivially satisfies (1.1). In particular, for all measurable mappings u of Ω into B the following three conditions are equivalent to each other:

(1.3) $$(x, u(x)) \in C \quad \text{a.e.}$$

(1.4) $$u(x) \in C_x \quad \text{a.e.}$$

(1.5) $$\int_\Omega f(x, u(x))\, dx < +\infty.$$

[1] We admit here the fact that the projection of a Borel set is measurable. This is a difficult result; it is a consequence of, e.g., Choquet's capacity theorem.

1.3. Second example: Carathéodory functions

Definition 1.2. *Let B be a Borel subset of* $\mathbf{R}^p$. *A mapping $f: \Omega \times B \to \bar{\mathbf{R}}$ is said to be a Carathéodory function if*

(1.6) *for almost all $x \in \Omega$, $f(x, .)$ is continuous on B,*

(1.7) *for all $a \in B$, $f(., a)$ is measurable on Ω.*

Proposition 1.1. *Every Carathéodory function is a normal integrand.*

Proof. By modifying f on a Borel set of Ω with null measure and by using an isomorphism of $\bar{\mathbf{R}}$ onto $[0, 1]$, we may assume that f is measurable in x for all a, is continuous in a for all x, and takes values in $[0, 1]$.

Once again we introduce a denumerable base of open sets $\mathscr{U}$ of B and the family Φ of l.s.c. functions of B into $[0,1]$ defined by $\Phi = \{k\mathbf{1}_u | u \in \mathscr{U}, k \in \mathbf{Q}, 0 \leqslant k \leqslant 1\}$. For all l.s.c. functions h of A into $[0,1]$ we have $h = \sup\{\varphi \in \Phi | \varphi \leqslant h\}$. We now introduce a dense denumerable family $\mathscr{B}$. Enumerating $\Phi = \{\varphi_n\}_{n \in \mathrm{N}}$ and for all $n \in \mathrm{N}$ and $a \in \mathscr{B}$ let us set:

$$E_{n,a} = \{ x \in \Omega | f(x, a) \geqslant \varphi_n(a) \}.$$

Since $f(., a)$ is measurable, $E_{n,a}$ is measurable and hence $E_n = \bigcap_{a \in \mathscr{B}} E_{n,a}$ is measurable

$$E_n = \{ x \in \Omega | f(x, a) \geqslant \varphi_n(a) \ \ \forall a \in \mathscr{B} \}.$$

Now $f(x, .)$ is continuous, φ_n is l.s.c. and $\mathscr{B}$ is everywhere dense. We deduce that

$$E_n = \{ x \in \Omega | f(x, a) \geqslant \varphi_n(a) \ \ \forall a \in B \}.$$

And by definition of the family Φ:

$$f(x, a) = \sup_{n \in \mathrm{N}} \varphi_n(a) \mathbf{1}_{E_n}(x).$$

For each $n \in \mathrm{N}$, there exists a Borel subset C_n of Ω such that $\mathbf{1}_{E_n} = \mathbf{1}_{C_n}$ almost everywhere. Let:

$$\tilde{f}(x, a) = \sup_{n \in \mathrm{N}} \varphi_n(a) \mathbf{1}_{C_n}(x).$$

The function $\tilde{f}$ is Borel on $\Omega \times B$ as the least upper bound of Borel functions, and $f(x, .) = \tilde{f}(x, .)$ for almost all x. Hypothesis (1.2) is thus verified; hypothesis (1.1) is covered by (1.6). ∎

In this particular case Theorem 1.1 takes the following form, known as the Scorza–Dragoni theorem:

Scorza–Dragoni Theorem. *A mapping $f: \Omega \times B \to \overline{\mathbf{R}}$ is a Carathéodory function if and only if for all compact sets $K \subset \Omega$ and all $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset K$ such that $meas(K - K_\varepsilon) \leqslant \varepsilon$ for which the restriction of f to $K_\varepsilon \times B$ is continuous.*

Proof. Let the compact subset K and $\varepsilon > 0$ be given. Since f is a normal integrand, there exists a compact set $K_+ \subset K$ such that $\text{meas}(K - K_+) \leqslant \varepsilon/2$ and that the restriction of f to $K_+ \times B$ is l.s.c. But $-f$ is also a normal integrand and we can thus find a compact set $K_- \subset K$ such that $\text{meas}(K - K_-) \leqslant \varepsilon/2$ and for which the restriction of $-f$ to $K_- \times B$ is l.s.c. If $K_\varepsilon = K_+ \cap K_-$, f will be l.s.c. and u.s.c. and hence continuous on $K_\varepsilon \times A$ and $\text{meas}(K - K_\varepsilon) \leqslant \varepsilon$. The converse is an easy consequence of Theorem 1.1. ■

1.4. A measurable selection theorem

We shall now assume that B is a compact subset of $\mathbf{R}^p$. For almost all $x \in \Omega$, there thus exists an $a(x) \in B$ where $f(x, .)$ attains its minimum. We will show that we can choose $a(x)$ in such a way that the mapping a defined from Ω into B is measurable.

Lemma 1.1. *Let B be a compact subset of $\mathbf{R}^p$ and g a normal integrand of $\Omega \times B$. We set*

$$(1.8) \qquad \left| \begin{aligned} &g_0(x, a) = 0 \quad \text{if } g(x, a) = \min_{b \in B} \{ g(x, b) \} \\ &g_0(x, a) = +\infty \quad \text{if } g(x, a) > \min_{b \in B} \{ g(x, b) \}. \end{aligned} \right.$$

Then g_0 is a normal integrand.

Proof. Using an isomorphism, we may assume that g takes its values in $[-1, 1]$. For every $\varepsilon > 0$ and every compact $K \subset \Omega$ we can find a compact subset $K_\varepsilon \subset K$ such that $\text{meas}(K - K_\varepsilon) \leqslant \varepsilon$ and that the restriction of g to $K_\varepsilon \times B$ is l.s.c. Define $\varphi: K_\varepsilon \to [-1, 1]$ by:

$$\varphi(x) = \min_{b \in B} \{ g(x, b) \}.$$

Let us show that φ is l.s.c. Let $(x_n)_{n \in \mathbf{N}}$ be a sequence of K_ε converging to $\bar{x}$. We extract a subsequence $x_{n'}$ such that:

$$(1.9) \qquad \varliminf_{n \to \infty} \varphi(x_n) = \lim_{n' \to \infty} \varphi(x_{n'}).$$

Since B is compact, there always exists $a_{n'} \in B$ such that $\varphi(x_{n'}) = g(x_{n'}, a_{n'})$. We can extract a subsequence $a_{n''}$ converging to $a \in B$. Since g is l.s.c. on $K_\varepsilon \times B$ we have:

$$\varliminf_{n'' \to \infty} g(x_{n''}, a_{n''}) \geqslant g(\bar{x}, \bar{a})$$

and *a fortiori*

$$\varliminf_{n'' \to \infty} \varphi(x_{n''}) \geqslant \varphi(\bar{x}). \tag{1.10}$$

Whence by comparing (1.9) and (1.10), the lower semi-continuity of φ follows. Let us then introduce $C \subset K_\varepsilon \times B$ defined by

$$C = \{ (x, a) \in K_\varepsilon \times B \,|\, g(x, a) = \varphi(x) \}.$$

This is the set of points where two l.s.c. functions coincide and is thus a Borel subset of $K_\varepsilon \times B$. Moreover, for all $x \in K_\varepsilon$, $C_x = \{a \in B \,|\, (x,a) \in C\}$ is closed since it is the set of points where an l.s.c. function attains its minimum. This means that g_0, which is none other than the indicator function of C, is a normal integrand of $K_\varepsilon \times B$.

In particular, we can find a compact subset $K_{2\varepsilon} \subset K_\varepsilon$ such that $\operatorname{meas}(K_\varepsilon - K_{2\varepsilon}) \leqslant 2\varepsilon$ and for which the restriction of g_0 to $K_{2\varepsilon}$ is l.s.c. Let us collect our results together: for all $\varepsilon > 0$ and all compact sets $K \subset \Omega$, we have found a compact set $K_{2\varepsilon} \subset K$ such that $\operatorname{meas}(K - K_{2\varepsilon}) \leqslant 2\varepsilon$ and for which the restriction of g_0 to $K_{2\varepsilon} \times B$ is l.s.c. Hence g_0 is a normal integrand of $\Omega \times B$. ∎

Theorem 1.2. *Let B be a compact subset of $\mathbf{R}^p$ and g a normal integrand of $\Omega \times B$. Then there exists a measurable mapping $\bar{u} : \Omega \to B$ such that for all $x \in \Omega$:*

$$g(x, \bar{u}(x)) = \min_{a \in B} \{ g(x, a) \}. \tag{1.11}$$

Proof. We define g_0 by formulae (1.8). We then take a sequence a_n, $n \geqslant 1$, which is dense in B. We now define by induction a sequence g_n, $n \geqslant 1$, of normal integrands in the following way:

$$h_n(x, a) = g_n(x, a) + |a - a_n|$$

$$\left| \begin{array}{l} g_{n+1}(x, a) = 0 \quad \text{if} \quad h_n(x, a) = \min\limits_{b \in B} \{ h_n(x, b) \} \\ g_{n+1}(x, a) = +\infty \quad \text{if} \quad h_n(x, a) > \min\limits_{b \in B} \{ h_n(x, b) \}. \end{array} \right.$$

Successive applications of Lemma 1.1 show that the g_n, $n \in \mathbf{N}$, are all normal integrands. Thus $\bar{g} = \sup_{n\in\mathbf{N}} g_n$ is also a normal integrand. But it can be easily verified that $\operatorname{dom}\bar{g}(x,.)$ has been reduced to a point for all $x \in \Omega$, i.e. $\bar{g}$ is of the form:

$$\left| \begin{array}{l} \bar{g}(x, a) = 0 \quad \text{if} \quad a = \bar{u}(x) \\ \bar{g}(x, a) = +\infty \quad \text{if} \quad a \neq \bar{u}(x). \end{array} \right.$$

The function $\bar{u}$ thus defined is measurable, since $\bar{g}$ is a normal integrand and satisfies (1.11) since $\bar{g} \geqslant g_0$.

In particular we deduce that the mapping $x \mapsto \min_{a\in B}\{g(x,a)\}$ of Ω into $\mathbf{R}$ is measurable. Theorem 1.2 includes as a special case the following measurable selection theorem.

Corollary 1.1. *If B is a compact subset of $\mathbf{R}^p$, if C is a Borel set of $\Omega \times B$ whose sections:*

$$C_x = \{ a \in B \,|\, (x, a) \in C \}$$

are closed and non-empty for almost all x, then there exists a measurable mapping $\bar{u}: \Omega \to B$ such that for almost all x

$$\bar{u}(x) \in C_x.$$

We say that $\bar{u}$ is a *measurable selection* of C. This corollary is proved by applying Theorem 1.2 to the indicator function of C.

1.5. Polars and bipolars of normal integrands

Let us now take $B = \mathbf{R}^p$, and consider a normal integrand $f: \Omega \times \mathbf{R}^p \to \overline{\mathbf{R}}$. For all fixed $x \in \Omega$, the polar of the function $f(x,.)$ will be a mapping of $\mathbf{R}^p$ into $\overline{\mathbf{R}}$ denoted by:

$$\xi^* \mapsto f^*(x\,;\xi^*).$$

Proposition 1.2. *If f is a normal integrand of $\Omega \times \mathbf{R}^p$, then f^* is a normal integrand of $\Omega \times \mathbf{R}^p$.*

Proof. For all $x \in \Omega$, $f^*(x,.)$ is a convex l.s.c. function. It only remains to verify hypothesis (1.2) for f^*. To do this we define:

$$\begin{array}{l} f_n(x, \xi) = +\infty \quad \text{if} \quad |\xi| > n \\ f_n(x, \xi) = f(x, \xi) \quad \text{if} \quad |\xi| \leqslant n. \end{array}$$

We have $f = \inf_{n \in \mathbf{N}} f_n$, hence (by I(4.6)), for all $x \in \Omega$:

$$f^*(x\,;.) = \sup_{n \in \mathbf{N}} f_n^*(x\,;.). \tag{1.12}$$

By definition

$$f_n^*(x\,;\xi^*) = \sup_{|\xi| \leqslant n} \{ \langle \xi, \xi^* \rangle - f_n(x, \xi) \}. \tag{1.13}$$

For all $x \in \Omega$, $f_n^*(x;.)$ will be either identically equal to $-\infty$, or a finite convex function of $\mathbf{R}^p$ into $\overline{\mathbf{R}}$.

Let us take $\xi^* \in \mathbf{R}^p$, and any compact set $K \subset \Omega$. For all $\varepsilon > 0$ there exists a compact subset $K_\varepsilon \subset K$ such that $\operatorname{meas}(K - K_\varepsilon) \leqslant \varepsilon$ and for which the restriction of f_n to $K_\varepsilon \times \mathbf{R}^p$ is l.s.c. Since the balls of $\mathbf{R}^p$ are compact, the family of mappings $-f_n(.,\xi) + \langle \xi, \xi^* \rangle$, for $|\xi| \leqslant n$, will be equi-l.s.c. on K_ε. From equation (1.13) we also deduce that $f_n^*(.\,;\xi^*)$ is u.s.c. on K_ε. Since $\varepsilon > 0$ is arbitrary, $f_n^*(.\,;\xi^*)$ is measurable on K. Since the compact set K is arbitrary $\tilde{f}_n^*(.\,;\xi^*)$ is measurable on Ω.

For all $n \in \mathbf{N}$, f is therefore a Carathéodory function, and *a fortiori*, a normal integrand. We thus have $f_n^*(x;.) = \tilde{f}_n^*(x;.)$, for almost all x, where f_n^* is Borel. From (1.12) we deduce that $f^*(x;.) = \tilde{f}^*(x;.)$ for almost all x, where $\tilde{f}^* = \sup_{n \in \mathbf{N}} \tilde{f}_n^*$ is Borel. Whence (1.2) for f^*. ∎

By repeating this operation, we arrive at the Γ-regularization of the function $f(x,.)$ which will be denoted by: $\xi \mapsto f^{**}(x;\xi)$. We at once obtain the following corollary of Proposition 1.2:

Proposition 1.3. *If f is a normal integrand of $\Omega \times \mathbf{R}^p$, then f^{**} is a normal integrand of $\Omega \times \mathbf{R}^p$.*

1.6. Lower semi-continuity of integrals

Let us now pass from the questions of measurability to those of integrability. Here we give an easy consequence of Fatou's lemma which will be extremely useful to us in the future:

Proposition 1.4. *Let f be a normal positive integrand of $\Omega \times \mathbf{R}^p$ and $(u_n)_{n \in \mathbf{N}}$ a sequence of measurable mappings of Ω into $\mathbf{R}^p$, converging almost everywhere to $\bar{u}$. Then we have:*

$$\int_\Omega f(x, \bar{u}(x))\, \mathrm{d}x \leqslant \varliminf_{n \to \infty} \int_\Omega f(x, u_n(x))\, \mathrm{d}x. \tag{1.14}$$

Proof. We have a sequence of positive measurable functions to which we can apply Fatou's Lemma:

$$\int_\Omega \varliminf_{n\to\infty} f(x, u_n(x))\,\mathrm{d}x \leqslant \varliminf_{n\to\infty} \int_\Omega f(x, u_n(x))\,\mathrm{d}x. \tag{1.15}$$

But since $f(x, .)$ is l.s.c. for almost all $x \in \Omega$:

$$f(x, \bar{u}(x)) \leqslant \varliminf_{n\to\infty} f(x, u_n(x)) \quad \text{a.e.} \tag{1.16}$$

Whence the result, on substituting (1.16) into (1.15). ∎

Corollary 1.2. *Let f be a normal positive integrand. The function*

$$F : u \mapsto \int_\Omega f(x, u(x))\,\mathrm{d}x,$$

is positive and l.s.c. of $L^\alpha(\Omega)$ into $\overline{\mathbf{R}}$, for all α, $1 \leqslant \alpha \leqslant \infty$.

1.7. Weak compactness in $L^1(\Omega)$

To conclude these preliminaries, let us recall the characterization of the weakly relatively compact subsets of $L^1(\Omega)$.

Theorem 1.3. *Let $\mathscr{F} \subset L^1(\Omega)$. The following statements are then equivalent to one another:*

(a) *from any sequence $(u_n)_{n\in\mathbf{N}}$ of $\mathscr{F}$, we can extract a subsequence which is weakly convergent in L^1;*

(b) *for all $\varepsilon > 0$, there exists $\lambda > 0$ such that:*

$$\forall u \in \mathscr{F}, \qquad \int_{\{|u| \geqslant \lambda\}} |u(x)|\,\mathrm{d}x \leqslant \varepsilon;$$

(c) *for all $\varepsilon > 0$, there exists $\delta > 0$ such that we have $\int_B |u(x)|\mathrm{d}x \leqslant \varepsilon$ for all $u \in \mathscr{F}$ and all measurable B of measure $\leqslant \delta$;*

(d) *there exists a positive Borel function Φ: $[0, \infty[\to \overline{\mathbf{R}}_+$ such that $\lim_{\to t\infty}(\Phi(t)/t = +\infty$ and*

$$\sup_{u\in\mathscr{F}} \int \Phi \circ |u| < +\infty.$$

Condition (b) is called *equi-integrability.* The equivalence (a)$\Leftrightarrow$(b)$\Leftrightarrow$(c) is *Dunford–Pettis' compactness criterion.* The equivalence (b)$\Leftrightarrow$(d) is due to de la Vallée–Poussin.

Lemma 1.2. *The function Φ introduced in (d) can be assumed to be convex, l.s.c. and increasing.*

Proof. To say that $\lim_{t\to\infty}\Phi(t)/t=+\infty$ means that for all $m\in\mathbf{R}$, Φ admits an affine minorant with slope m. Then so does Φ^{**}, and hence $\Phi^{**}(t)/t\to+\infty$ when $t\to+\infty$. The function Φ^{**} is convex and l.s.c. and attains its minimum $\bar{a}$ at $\bar{t}$:

$$\bar{a} = \inf\Phi = \min\Phi^{**} = \Phi^{**}(\bar{t}).$$

It decreases on $[0,\bar{t}]$, then increases over $[\bar{t},+\infty[$. Let us define a convex function $\tilde{\Phi}$, increasing and l.s.c. on $[0,+\infty[$ by

$$\begin{aligned}\tilde{\Phi}(t) &= \bar{a} \quad \text{on} \quad [0,\bar{t}]\\ \tilde{\Phi}(t) &= \Phi^{**}(t) \quad \text{on} \quad [\bar{t},+\infty[.\end{aligned}$$

We have $\tilde{\Phi}<\Phi$, and thus

$$\sup_{u\in\mathscr{F}}\int\tilde{\Phi}\circ|u| \leqslant \sup_{u\in\mathscr{F}}\int\Phi\circ|u| < \infty. \quad \blacksquare$$

Theorem 1.3 is a very deep result. We are going to use it to prove a generalization of Lebesgue's theorem which will be extremely useful in what follows:

Corollary 1.3. *Let $(u_n)_{n\in\mathbf{N}}$ be an equi-integrable sequence of $L^1(\Omega)$ such that $u_n(x)\to u(x)$ almost everywhere. Then u is integrable, and $u_n\to u$ in $L^1(\Omega)$.*

Proof. We begin by showing that, under the given hypotheses, the sequence u_n converges weakly to u in $L^1(\Omega)$. For this it is sufficient to show that we can extract from it a subsequence converging weakly in $L^1(\Omega)$ to u. Now there exists a subsequence $u_{n'}$ converging weakly in $L^1(\Omega)$ to a function v (Theorem 1.3, (b)$\Rightarrow$(a)). From Mazur's lemma, we can find a sequence of convex combinations $v_{n'}\in\overline{\mathrm{co}}\bigcup_{p'\geqslant n'}\{u_{p'}\}$ converging to v in $L^1(\Omega)$, and so we can extract a sequence $v_{n''}$ converging to v almost everywhere:

$$v_{n''}(x)\to v(x) \qquad \text{a.e.} \tag{1.17}$$

But by hypothesis:

$$v_{n'}(x)\in\overline{\mathrm{co}}\bigcup_{p'\geqslant n'}\{u_{p'}(x)\}\to u(x) \qquad \text{a.e.} \tag{1.18}$$

Comparing (1.17) and (1.18), we obtain $v = u$ almost everywhere, and the sequence u_n thus converges weakly to u.

We immediately deduce from this that it also converges in the norm: it is sufficient to apply the preceding result to the sequence of functions $|u_n - u|$. We can easily verify that it is equi-integrable and that it converges almost everywhere to zero. It thus converges weakly to zero, and in particular:

$$\int_\Omega |u_n(x) - u(x)|\, dx \to 0, \qquad \|u_n - u\|_{L^1} \to 0. \quad \blacksquare$$

If for example there exists $a \in L^1(\Omega)$ such that for all n, $u_n(x) \leqslant a(x)$ almost everywhere, the sequence $(u_n)_{n\in\mathbf{N}}$ will then be equi-integrable. We thus recover Lebesgue's theorem.

2. AN OPTIMIZATION PROBLEM

2.1. The integrand f: definition, first properties

Let $\Phi: [0,+\infty[\to \mathbf{R} \cup \{+\infty\}$ be a non-negative increasing, convex, l.s.c. function such that

$$\lim_{t\to\infty} \frac{\Phi(t)}{t} = +\infty. \tag{2.1}$$

Let us consider a mapping f of $\Omega \times (\mathbf{R}^\ell \times \mathbf{R}^m)$ into $\overline{\mathbf{R}}$. We shall assume that it is a normal integrand, i.e. that it satisfies (1.1) and (1.2) with $B = \mathbf{R}^\ell \times \mathbf{R}^m$, and that we have the estimate

$$\Phi(|\xi|) \leqslant f(x, s, \xi) \tag{2.2}$$

In particular, f is non-negative, $f(x,.,.)$ is l.s.c. for almost all $x \in \Omega$, and $f(.,s,\xi)$ is measurable for all $(s,\xi) \in \mathbf{R}^\ell \times \mathbf{R}^m$. By epi$f(x,s)$ we shall denote the epigraph of the function $f(x,s,.)$ in $\mathbf{R}^m \times \mathbf{R}$. We shall now prove a continuity property of the multi-valued mapping $(x,s) \to \overline{\text{co}}\,\text{epi}f(x,s)$.

Lemma 2.1. *Let E be a metric space and φ a l.s.c. mapping of $E \times \mathbf{R}^m$ into $\overline{\mathbf{R}}$, satisfying*

$$\forall e \in E, \qquad \Phi(|\xi|) \leqslant \varphi(e, \xi). \tag{2.3}$$

Then, for all $\bar{e} \in E$, we have

$$\bigcap_{\varepsilon>0} \overline{\text{co}} \bigcup_{|e-\bar{e}|\leqslant\varepsilon} \text{epi}\,\Phi(e,.) = \overline{\text{co}}\,\text{epi}\,\Phi(\bar{e},.). \tag{2.4}$$

Proof. It is trivial that the right-hand side is included in the left-hand side. For the converse, take any:

$$(2.5) \qquad (\bar{\xi}, \bar{a}) \notin \overline{\text{co}}\ \text{epi}\ \varphi(\bar{e}, .).$$

There exists an affine hyperplane of $\mathbf{R}^m \times \mathbf{R}$ which strictly separates $(\bar{\xi},\bar{a})$ from $\overline{\text{co}}\ \text{epi}\ \varphi(\bar{e}, .)$. If this hyperplane is non-vertical, it is the graph of an affine function ℓ over $\mathbf{R}^m$ such that:

$$(2.6) \qquad \bar{a} < \ell(\bar{\xi})$$

$$(2.7) \qquad \forall \xi \in \mathbf{R}^m, \qquad \ell(\xi) < \varphi(\bar{e}, \xi).$$

If the hyperplane in question, denoted by $\mathscr{H}$, is vertical, there exists an affine function ℓ' over $\mathbf{R}^m$ such that $\ell'(\xi) = 0$ for $(\xi, a) \in \mathscr{H}, \ell'(\bar{\xi}) > 0$ and $\ell'(\xi) < 0$ for $\xi \in \text{dom}\varphi(\bar{e},.)$. But, from (2.3), φ is non-negative; the function $\ell = c\ell'$, for $c > 0$ sufficiently large, will thus also satisfy (2.6) and (2.7).

From (2.3) we deduce that there exists $M > 0$ such that

$$(2.8) \qquad |\xi| \geqslant M \Rightarrow \Phi(|\xi|) \geqslant \ell(\xi).\text{(*)}$$

Since the balls, of $\mathbf{R}^m$ are compact, the family $\Phi(.,\xi)$, for $|\xi| \leqslant M$, of mappings from E into $\overline{\mathbf{R}}$, is equi-l.s.c. Let ρ be an increasing homeomorphism of $\overline{\mathbf{R}}$ onto $[-1, 1]$.(1) We then set:

$$(2.9) \qquad m = \min_{|\xi| \leqslant M} \{ \rho \circ \varphi(\bar{e}, \xi) - \rho \circ \ell(\xi) \}.$$

It is the minimum of a l.s.c. function on a compact set. It is thus attained and we have $m > 0$ from (2.7). By equi-lower-semi-continuity, there exists $\bar{\varepsilon} > 0$ such that $|e - \bar{e}| \leqslant \bar{\varepsilon}$ and $|\xi| \leqslant M$ imply that:

$$(2.10) \qquad \rho \circ \varphi(e, \xi) \geqslant \rho \circ \varphi(\bar{e}, \xi) - m$$

$$(2.11) \qquad \rho \circ \varphi(e, \xi) \geqslant \rho \circ \ell(\xi)$$

by (2.9).

Finally, let:

$$(2.12) \qquad |e - \bar{e}| \leqslant \bar{\varepsilon} \text{ and } |\xi| \leqslant M \Rightarrow \varphi(e, \xi) \geqslant \ell(\xi).$$

Grouping together (2.8) and (2.12) and taking (2.3) into account, we see that:

$$(2.13) \qquad |e - \bar{e}| \leqslant \bar{\varepsilon} \Rightarrow \forall \xi \in \mathbf{R}^m, \qquad \varphi(e, \xi) \geqslant \ell(\xi).$$

(1) For example $\rho(s) = (2/\pi) \arctan s$.

(*) See an erratum concerning the next few lines at the end of this chapter.

But (2.6) means that $(\bar{\xi},\bar{a})$ belongs to the lower open half-space determined by ℓ, and (2.13) means that $\bigcup_{|e-\bar{e}|\leqslant\bar{\varepsilon}} \operatorname{epi}\varphi(e,.)$ belongs to the upper closed half-space. But then $\overline{\text{co}} \bigcup_{|e-\bar{e}|\geqslant\bar{\varepsilon}} \operatorname{epi}\varphi(e,.)$ also belongs to the upper closed half-space, and hence:

$$(\bar{\xi},\bar{a}) \notin \overline{\text{co}} \bigcup_{|e-\bar{e}|\leqslant\bar{\varepsilon}} \text{epi } \varphi(e,.).$$

A fortiori:

$$(\bar{\xi},\bar{a}) \notin \bigcap_{\varepsilon>0} \overline{\text{co}} \bigcup_{|e-\bar{e}|\leqslant\varepsilon} \text{epi } \varphi(e,.). \tag{2.14}$$

The fact that (2.5) implies (2.14) gives us the desired inclusion:

$$\overline{\text{co}} \text{ epi } \varphi(\bar{e},.) \subset \bigcap_{\varepsilon>0} \overline{\text{co}} \bigcap_{|e-e|\leqslant\varepsilon} \text{epi } \varphi(e,.). \quad \blacksquare$$

Corollary 2.1. *If f is a normal integrand of $\Omega \times \mathbf{R}^{\ell} \times \mathbf{R}^m$ satisfying* (2.2), *we have for almost all $x \in \Omega$,*

$$\bigcap_{\varepsilon>0} \overline{\text{co}} \bigcup_{|s-\bar{s}|\leqslant\varepsilon} \text{epi } f(x,s) = \overline{\text{co}} \text{ epi } f(x,\bar{s}). \tag{2.15}$$

Proof. It is sufficient to apply Lemma 2.1 to the function $f(x.,.,.)$ over $\mathbf{R}^{\ell} \times \mathbf{R}^m$, at any point $x \in \Omega$ which makes it l.s.c.

2.2. A lower semi-continuity result

We are now in a position to state the fundamental result of this chapter which yields a property of lower semi-continuity:

Theorem 2.1. *Let f be a normal integrand of $\Omega \times (\mathbf{R}^{\ell} \times \mathbf{R}^m)$, such that:*

$$\Phi(|\xi|) \leqslant f(x,s,\xi) \tag{2.2}$$

where $\Phi: \mathbf{R}_+ \mapsto \mathbf{R}_+$ is a convex increasing l.s.c. function satisfying (2.1) *and,*

$$\forall (x,s) \in \Omega \times \mathbf{R}^{\ell}, \quad f(x,s,.) \textit{ is convex over } \mathbf{R}^m. \tag{2.16}$$

Let $(p_n)_{n\in\mathbf{N}}$ be a sequence converging weakly to $\bar{p}$ in $L^1(\Omega)^m$ and $(u_n)_{n\in\mathbf{N}}$ a sequence of measurable functions converging almost everywhere to $\bar{u}$. Then:

$$\int_{\Omega} f(x,\bar{u}(x),\bar{p}(x))\,\mathrm{d}x \leqslant \underline{\lim}_{n\to\infty} \int_{\Omega} f(x,u_n(x),p_n(x))\,\mathrm{d}x. \tag{2.17}$$

Proof. Since f is a normal non-negative integrand, all the written integrals have the same sign. If the right-hand side of (2.17) takes the value $+\infty$, the inequality is trivial. Otherwise, by extracting a subsequence, we may assume that:

$$\lim_{n \to \infty} \int_\Omega f(x, u_n(x), p_n(x))\, \mathrm{d}x = c < +\infty. \tag{2.18}$$

We now apply Mazur's lemma (*cf.* Chap. I, Section 1) to the sequence p_n, which is weakly convergent in L^1. There exists a sequence of convex combinations $\sum_{k=n'}^N \alpha_k p_k$, with $\alpha_k > 0$ and $\sum_{k=n'}^N \alpha_k = 1$, converging to $\bar{p}$ in L^1. We can thus extract a subsequence $\sum_{k=n'}^N \alpha_k p_k$ converging almost everywhere to $\bar{p}$:

$$\sum_{k=n'}^N \alpha_k p_k(x) \to \bar{p}(x) \quad \text{a.e. as} \quad n' \to \infty. \tag{2.19}$$

Let us take a point $x \in \Omega$ where the convergence (2.19) occurs and where $u_n(x) \to \bar{u}(x)$ as $n \to \infty$. *A fortiori*, $\bar{u}_{n'}(x) \to \bar{u}(x)$ when $n' \to \infty$. In $\mathbf{R}^m \times \mathbf{R}$, for all n', we have:

$$\Big(\sum_{k=n'}^N \alpha_k p_k(x), \sum_{k=n'}^N \alpha_k f(x, u_k(x), p_k(x))\Big) \in \operatorname{co} \bigcup_{k=n'}^N \operatorname{epi} f(x, u_k(x)) \tag{2.20}$$

and *a fortiori*:

$$\Big(\sum_{k=n}^N \alpha_k p_k(x), \sum_{k=n'}^N \alpha_k f(x, u_k(x), p_k(x))\Big) \in \operatorname{co} \bigcup_{q \geqslant n'} \operatorname{epi} f(x, u_q(x)). \tag{2.21}$$

We now take $\varepsilon > 0$. From (2.13), there exists n_0' sufficiently large to give us $|u_2(x) - \bar{u}(x)| \leqslant \varepsilon$ for all $q \geqslant n_0'$. Hence for all $n' \geqslant n_0'$, we have:

$$\Big(\sum_{k=n'}^N \alpha_k p_k(x), \sum_{k=n'}^N a_k f(x, u_k(x), p_k(x))\Big) \in \operatorname{co} \bigcup_{|s - \bar{u}(x)| \leqslant \varepsilon} \operatorname{epi} f(x, s). \tag{2.22}$$

By making n' go to infinity, we obtain from (2.19):

$$\Big(\bar{p}(x), \varliminf_{n' \to \infty} \sum_{k=n'}^N \alpha_k f(x, u_k(x), p_k(x))\Big) \in \overline{\operatorname{co}} \bigcup_{|s - \bar{u}(x)| \leqslant \varepsilon} \operatorname{epi} f(x, s) \tag{2.23}$$

and as this is true for all $\varepsilon > 0$:

$$\Big(\bar{p}(x), \varliminf_{n' \to \infty} \sum_{k=n'}^N \alpha_k f(x, u_k(x), p_k(x))\Big) \in \bigcap_{\varepsilon > 0} \overline{\operatorname{co}} \bigcup_{|s - \bar{u}(x)| \leqslant \varepsilon} \operatorname{epi} f(x, s). \tag{2.24}$$

From Corollary 2.1, we have:

$$\bigcup_{\varepsilon > 0} \overline{\mathrm{co}} \bigcup_{|s - \bar{u}(x)| \leqslant \varepsilon} \operatorname{epi} f(x, s) = \overline{\mathrm{co}} \operatorname{epi} f(x, \bar{u}(x)). \tag{2.25}$$

But because of hypothesis (2.16), the function $f(x, \bar{u}(x), .)$ is convex and l.s.c., and its epigraph is therefore closed and convex. Finally (2.24) can be written as:

$$\left(\bar{p}(x), \varliminf_{n' \to \infty} \sum_{k=n'}^{N} \alpha_k f(x, u_k(x), p_k(x))\right) \in \operatorname{epi} f(x, \bar{u}(x)) \tag{2.26}$$

which by definition means that:

$$f(x, \bar{u}(x), \bar{p}(x)) \leqslant \varliminf_{n' \to \infty} \sum_{k=n'}^{N} \alpha_k f(x, u_k(x), p_k(x)). \tag{2.27}$$

We now integrate both sides over Ω:

$$\int_\Omega f(x, \bar{u}(x), \bar{p}(x))\,\mathrm{d}x \leqslant \int_\Omega \varliminf_{n' \to \infty} \sum_{k=n'}^{N} \alpha_k f(x, u_k(x), p_k(x))\,\mathrm{d}x. \tag{2.28}$$

But all the integrands are positive, which allows us to apply Fatou's lemma:

$$\int_\Omega \varliminf_{n' \to \infty} \sum_{k=n'}^{N} \alpha_k f(x, u_k(x), p_k(x))\,\mathrm{d}x \leqslant \varliminf_{n' \to \infty} \sum_{k=n'}^{N} \alpha_k \int_\Omega f(x, u_k(x), p_k(x)\,\mathrm{d}x \tag{2.29}$$

$$\int_\Omega f(x, \bar{u}(x), \bar{p}(x))\,\mathrm{d}x \leqslant \varliminf_{n' \to \infty} \sum_{k=n'}^{N} \alpha_k \int_\Omega f(x, u_k(x), p_k(x))\,\mathrm{d}x. \tag{2.30}$$

But from (2.18) it is easy to deduce that the right-hand side of (2.30) is equal to c. Indeed, for all $\varepsilon > 0$, we can find n_0 such that

$$\forall n \geqslant n_0, \qquad c - \varepsilon \leqslant \int_\Omega f(x, u_n(x), p_n(x))\,\mathrm{d}x \leqslant c + \varepsilon \tag{2.31}$$

$$c - \varepsilon \leqslant \sum_{k=n'}^{N} \alpha_k \int_\Omega f(x, u_k(x), p_k(x))\,\mathrm{d}x \leqslant c + \varepsilon, \tag{2.32}$$

whence finally

$$\lim_{n' \to \infty} \sum_{k=n'}^{N} \alpha_k \int_\Omega f(x, u_k(x), p_k(x))\,\mathrm{d}x = c. \tag{2.33}$$

Returning to (2.30) and (2.18), we obtain the desired result:

$$\int_\Omega f(x, \bar{u}(x), \bar{p}(x))\,\mathrm{d}x \leqslant c = \lim_{n \to \infty} \int f(x, u_k(x), p_k(x))\,\mathrm{d}x. \quad \blacksquare$$

2.3. The case where f is not convex in ξ

We now return to the general case where f is a normal integrand satisfying (2.2) but no longer satisfying (2.16). Then, of course, we no longer have Theorem 2.1 at our disposal. It is natural to introduce $f^{**}(x,s;.)$, the Γ-regularization of the function $f(x,s,.)$. From the results of Section 1, this is a normal integrand of $(\Omega \times \mathbf{R}^\ell) \times \mathbf{R}^m$, but the question arises as whether it is a normal integrand of $\Omega \times (\mathbf{R}^\ell \times \mathbf{R}^m)$.(1)

Proposition 2.1. *If f is a normal integrand of $\Omega \times (\mathbf{R}^\ell \times \mathbf{R}^m)$ satisfying*

$$\Phi(|\xi|) \leqslant f(x, s, \xi) \tag{2.2}$$

where Φ answers to the above description, then f^{**} is also a normal integrand of $\Omega \times (\mathbf{R}^\ell \times \mathbf{R}^m)$ and satisfies

$$\Phi(|\xi|) \leqslant f^{**}(x, s; \xi). \tag{2.34}$$

Proof. By (2.2), the convex l.s.c. function $\Phi(|.|)$ is everywhere less than $f(x,s,.)$. Taking the Γ-regularization of both sides, we obtain (2.34).

Let us now take any compact subset $K \subset \Omega$, and $\varepsilon > 0$. Since f is a normal integrand, we can find a compact subset $K_\varepsilon \subset K$ such that $\mathrm{meas}(K - K) \leqslant \varepsilon$ and such that the restriction of f to $K_\varepsilon \times \mathbf{R}^\ell \times \mathbf{R}^m$ is l.s.c. (Theorem 1.1). Moreover we know that $\mathrm{epi} f^{**}(\bar{x},\bar{s}) = \overline{\mathrm{co}}\,\mathrm{epi} f(\bar{x},\bar{s})$. From Lemma 2.1 applied to f over $(K_\varepsilon \times \mathbf{R}^\ell) \times \mathbf{R}^m$, we have

$$\mathrm{epi}\ f^{**}(\bar{x}, \bar{s}) = \bigcap_{\varepsilon > 0} \overline{\mathrm{co}} \bigcup_{\substack{|x - \bar{x}| \leqslant \varepsilon \\ |s - \bar{s}| \leqslant \varepsilon}} \mathrm{epi}\ f(x, s). \tag{2.35}$$

(1) We already know that $f^{**}(x, s; .)$ is l.s.c. in ξ, but we are not yet in a position to state that $f^{**}(x, .; .)$ is l.s.c. in (s, ξ).

But it is clear that:

$$\overline{\text{co}} \bigcup_{\substack{|x-\bar{x}| \leqslant \varepsilon \\ |s-\bar{s}| \leqslant \varepsilon}} \text{epi } f(x, s) \supset \overline{\bigcup_{\substack{|x-\bar{x}| \leqslant \varepsilon \\ |s-\bar{s}| \leqslant \varepsilon}} \overline{\text{co}} \text{ epi } f(x, s)} \tag{2.36}$$

which enables us to write (2.35) in the form:

$$\text{epi } f^{**}(\bar{x}, \bar{s}) \supset \bigcap_{\varepsilon > 0} \overline{\bigcup_{\substack{|x-\bar{x}| \leqslant \varepsilon \\ |s-\bar{s}| \leqslant \varepsilon}} \overline{\text{co}} \text{ epi } f(x, s)} \tag{2.37}$$

or again, replacing $\overline{\text{co}}\,\text{epi} f(x,s)$ by $\text{epi} f^{**}(x,s)$,

$$\text{epi } f^{**}(\bar{x}, \bar{s}) \supset \bigcap_{\varepsilon > 0} \overline{\bigcup_{\substack{|x-\bar{x}| \leqslant \varepsilon \\ |s-\bar{s}| \leqslant \varepsilon}} \text{epi } f^{**}(x, s)}. \tag{2.38}$$

Let us now take a sequence (x_n, s_n, ξ_n), $n \in \mathbf{N}$, converging to $(\bar{x}, s, \xi)$ in $K_\varepsilon \times \mathbf{R}^\ell \times \mathbf{R}^m$. We obtain:

$$(\bar{x}, \bar{s}, \bar{\xi}, \underline{\lim}\, f^{**}(x_n, s_n; \xi_n)) \in \bigcap_{\varepsilon > 0} \overline{\bigcup_{\substack{|x-\bar{x}| \leqslant \varepsilon \\ |s-\bar{s}| \leqslant \varepsilon}} \text{epi } f^{**}(x, s)}, \tag{2.39}$$

and by virtue of (2.38):

$$(\bar{x}, \bar{s}, \bar{\xi}, \underline{\lim}\, f^{**}(x_n, s_n; \xi_n)) \in \text{epi } f^{**}(x, s) \tag{2.40}$$

which by definition means that:

$$f^{**}(\bar{x}, \bar{s}; \bar{\xi}) \leqslant \underline{\lim}\, f^{**}(x_n, s_n; \xi_n). \tag{2.41}$$

Thus we have shown that for any compact set $K \subset \Omega$ and for all $\varepsilon > 0$, we can find a compact subset $K_\varepsilon \subset K$ such that $\text{meas}(K - K) \leqslant \varepsilon$ for which the restriction of f^{**} to $K_\varepsilon \times \mathbf{R}^\ell \times \mathbf{R}^m$ is l.s.c.

By Theorem 1.1, f^{**} is thus a normal integrand of $\Omega \times (\mathbf{R}^\ell \times \mathbf{R}^m)$. ∎

We can now have a partial extension of Theorem 2.1 to the non-convex case, making use of f^{**}. The result is as follows:

Proposition 2.2. *Let f be a normal integrand of $\Omega \times (\mathbf{R}^\ell \times \mathbf{R}^m)$ satisfying*

$$\Phi(|\xi|) \leqslant f(x, s, \xi). \tag{2.2}$$

Let $(p_n)_{n\in\mathbf{N}}$ *be a sequence converging weakly to* $\bar{p}$ *in* $L^1_m(\Omega)$ *and* $(u_n)_{n\in\mathbf{N}}$ *a sequence of measurable functions converging almost everywhere to* $\bar{u}$. *Then*

$$\int_\Omega f^{**}(x, \bar{u}(x); \bar{p}(x))\,\mathrm{d}x \leqslant \varliminf_{n\to\infty} \int_\Omega f(x, u_n(x), p_n(x))\,\mathrm{d}x. \tag{2.42}$$

Proof. It is sufficient to apply Theorem 2.1 to f^{**}, and to make use of the inequality $f^{**} \leqslant f$. We obtain:

$$\int_\Omega f^{**}(x, \bar{u}(x); \bar{p}(x))\,\mathrm{d}x \leqslant \varliminf_{n\to\infty} \int_\Omega f^{**}(x, u_n(x); p_n(x))\,\mathrm{d}x$$

$$\leqslant \varliminf_{n\to\infty} \int_\Omega f(x, u_n(x), p_n(x))\,\mathrm{d}x.$$

2.4. Calculus of variations: existence of solutions by convexity

We shall now formulate an optimization problem, which embodies a large class of problems in the calculus of variations, and apply to it the preceding results. As before, we are given a l.s.c., convex, increasing function $\Phi : [0, +\infty[\to \overline{\mathbf{R}}_+$ which satisfies:

$$\lim_{t\to\infty} \frac{\Phi(t)}{t}\,\mathrm{d}t = +\infty. \tag{2.1}$$

By L^Φ_m we shall denote those (classes of) measurable mappings p from Ω into $\mathbf{R}^m$ (modulo equality almost everywhere) for which $\int_\Omega \Phi \circ |p| < +\infty$.[1]

From (2.1) it is clear that $L^\Phi_m \subset L^1(\Omega)^m$. If, for instance we took for Φ the function $t \mapsto t^\alpha$, with $1 < \alpha < \infty$, the set L^Φ_m would be none other than $L^\alpha(\Omega)^m$ which will also be denoted by L^α_m. If we take for Φ the indicator function of the interval $[0, 1]$, L^Φ_m coincides with the unit ball of $L^\infty(\Omega)^m$ (or L^∞_m).

We are given a normal integrand f of $\Omega(\mathbf{R}^\ell \times \mathbf{R}^m)$ into $\overline{\mathbf{R}}$, such that there exists a function $a \in L^1(\Omega)$ which satisfies:

$$a(x) + \Phi(|\xi|) \leqslant f(x, s, \xi). \tag{2.43}$$

[1] The L^Φ_m are Orlicz classes. For the theory of Orlicz spaces *see* M. A. Krasnosel'skii and Ruticki [1], A. Fougères [1].

We are also given a weakly closed subset $\mathscr{U}$ of L^1_m, and a mapping $\mathscr{G}$ of $L^\Phi_m \cap \mathscr{U}$ into L^β_l, with $1 \leqslant \beta \leqslant \infty$, which satisfy the following property:

(2.44) if a sequence (p_n) of $\mathscr{U}$ converges weakly to $\bar{p}$ in L^1_m and if $\sup \int_\Omega \Phi \circ |p_n| < +\infty$, then we can extract from the sequence $(\mathscr{G}p_n)$ a subsequence which converges almost everywhere to $\mathscr{G}\bar{p}$

Assumption (2.44) is satisfied if, for example, $\mathscr{G}$ maps the sequences (p_n) which converge weakly to $\bar{p}$ in L^1_n and are such that $\int \Phi \circ p_n \leqslant$ constant $\forall n$, into sequences $(\mathscr{G}p_n)$ which converge strongly to $\mathscr{G}\bar{p}$ in L^β_β: we then say that $\mathscr{G}$ is a (Φ, β)-compactifier. If $\Phi(t) = t^\alpha$, $1 < \alpha < \infty$, *this amounts to saying that $\mathscr{G}$ maps the bounded and weakly convergent sequences of L^α_m into strongly convergent sequences of L^β*; indeed, the topologies $\sigma(L^1, L^\infty)$ and $\sigma(L^\alpha, L^{\alpha'})$ coincide on the bounded subsets of L^α_m, since L^∞_m is dense in $L^{\alpha'}_m$, $1/\alpha + 1/\alpha' = 1$. We say that $\mathscr{G}$ is an (α, β)-compactifier. Here are the main examples:

Proposition 2.3. *Let $\ell < \alpha < \infty$. If $\mathscr{G}$ is a compact continuous linear mapping of L^α_m into L^β_l, with $\ell \leqslant \beta \leqslant \infty$, then $\mathscr{G}$ is an (α, β)-compactifier.*

We recall that, by definition, a continuous linear mapping of L^α_m into L^β_l is called compact if it maps the bounded subsets of L^α_m into relatively compact subsets of L^β_l. Proposition 2.3 follows directly.

Proposition 2.4. *If $\mathscr{G}$ is a continuous linear mapping of L^1_m into L^β_l, with $1 < \beta \leqslant \infty$, then $\mathscr{G}$ satisfies* (2.44).

Proof. Since Ω is bounded, $L^\beta_l(\Omega) \subset L^1_l(\Omega)$, and the bounded sets of $L^\beta_l(\Omega)$ are weakly relatively compact in $L^1_l(\Omega)$. The mapping $\mathscr{G}$ maps the bounded sets of L^1_m into weakly relatively compact sets of L^1_l; by Grothendieck [1], theorem V.4.2 it will map the weakly compact subsets of L^1_m into compact subsets of L^1_l.

Let us now take Φ into account. If the $\int \Phi \circ |p_n|$'s are uniformly bounded, the $(p_n)_{n\in\mathbb{N}}$ form a weakly relatively compact subset of L^1_m (Theorem 1.3, de la Vallée–Poussin's criterion). The $(\mathscr{G}p_n)_{n\in\mathbb{N}}$ thus form a relatively compact subset of L^1_l, and we can extract a subsequence $(\mathscr{G}p_{nk})_{k\in\mathbb{N}}$ converging in L^1_l. Since the p_n converge weakly to $\bar{p}$, the $\mathscr{G}p_n$ converge weakly to $\mathscr{G}\bar{p}$. The limit of the sequence $(\mathscr{G}p_{nk})_{k\in\mathbb{N}}$ in L^1_l can therefore only be $\mathscr{G}\bar{p}$. Finally, we can extract from $(\mathscr{G}p_{nk})_{k\in\mathbb{N}}$ a subsequence which converges almost everywhere to $\mathscr{G}\bar{p}$. ∎

At last we are in a position to state the optimization problem:

$$(\mathscr{P}) \qquad \operatorname*{Inf}_{p\in\mathscr{U}\cap L^\Phi_m} \int_\Omega f(x, \mathscr{G}p(x), p(x))\,\mathrm{d}x$$

which can be put into the equivalent form:

$$(\mathscr{P}) \qquad \operatorname*{Inf}_{\substack{u,p \\ p\in\mathscr{U}\cap L^{\Phi}_m \\ u=\mathscr{G}p}} \int_{\Omega} f(x, u(x), p(x))\, \mathrm{d}x .$$

From Theorem 2.1, we immediately deduce an existence criterion for solutions to problem $(\mathscr{P})$:

Theorem 2.2. *Let f be a normal integrand of $\Omega \times (\mathbf{R}^{\ell} \times \mathbf{R}^m)$ satisfying*

$$(2.43) \qquad a(x) + \Phi(|\xi|) \leqslant f(x, s, \xi), \quad \textit{with } a \in L^1(\Omega)$$

$$(2.45) \qquad \forall (x, s) \in \Omega \times \mathbf{R}^{\ell}, \qquad f(x, s, .) \textit{ is convex on } \mathbf{R}^m.$$

Let $\mathscr{G}$ be a mapping of L^{Φ}_m into L^{β}_{ℓ} and $\mathscr{U}$ a weakly closed subset of L^1 satisfying (2.44). Problem $(\mathscr{P})$ admits at least one solution.

Proof. We set $g(x,s,\xi) = f(x,s,\xi) - a(x)$. This is a normal integrand such that

$$(2.46) \qquad \Phi(|\xi|) \leqslant g(x, s, \xi)$$

$$(2.47) \qquad \forall (x, s) \in \Omega \times \mathbf{R}^{l}, \qquad g(x, s, .) \text{ is convex on } \mathbf{R}^m.$$

It is clear that:

$$(2.48) \qquad \int_{\Omega} f(x, u(x), p(x)) = \int_{\Omega} g(x, u(x), p(x)) + \int_{\Omega} a(x)\, \mathrm{d}x.$$

The last term is a constant. Let us therefore take a minimizing sequence $(p_n)_{n\in\mathbf{N}}$ of problem $(\mathscr{P})$; and let us set $u_n = \mathscr{G}p_n$. By definition, $p_n \in \mathscr{U}$ for all n, and:

$$(2.49) \qquad \int_{\Omega} g(x, u_n(x), p_n(x))\, \mathrm{d}x \to \inf(\mathscr{P}) - \int_{\Omega} a(x)\, \mathrm{d}x.$$

From (2.46) we deduce that:

$$(2.50) \qquad \int_{\Omega} \Phi \circ |p_n| \leqslant \text{constant}.$$

From Theorem 1.3, we can extract from $(p_n)_{n\in\mathbf{N}}$ a subsequence $p_{n'}$ which is weakly convergent to $\bar{p}$ in L^1_m. Since $\mathscr{G}$ satisfies (2.44) we can extract a subsequence $p_{n''}$ such that $p_{n''}$ converges weakly to $\bar{p}$ in L^1_m and $\mathscr{G}p_{n''}$ converges to $\mathscr{G}\bar{p}$ almost everywhere.

Applying Theorem 2.1:

$$\int_\Omega g(x, \bar{u}(x), \bar{p}(x))\, dx \leqslant \varliminf_{n\to\infty} \int g(x, u_n(x), p_n(x))\, dx. \tag{2.51}$$

Adding $\int_\Omega a(x) dx$ to both sides:

$$\int_\Omega f(x, u(x), p(x))\, dx \leqslant \varliminf_{n\to\infty} \int_\Omega f(x, u_n(x), p_n(x))\, dx. \tag{2.52}$$

That is, as the sequence (p_n) is minimizing:

$$\int_\Omega f(x, \bar{u}(x), \bar{p}(x))\, dx \leqslant \inf(\mathscr{P}). \tag{2.53}$$

But $\bar{p} \in \mathscr{U}$ is the weak limit of the $p_{n''} \in \mathscr{U}$'s. From (2.53) we conclude therefore that $\bar{u} = \mathscr{G}\bar{p}$ is the solution of $(\mathscr{P})$.

2.5. Calculus of variations: relaxation

In the case where we no longer assume $f(x,s,.)$ to be convex, problem $(\mathscr{P})$ in general no longer has a solution. We shall see that it is natural to associate with problem $(\mathscr{P})$ the following problem, termed the *relaxed problem*,

$$\operatorname*{Inf}_{p\in\mathscr{U}\cap L^\Phi_m} \int_\Omega f^{**}(x, \mathscr{G}p(x); p(x))\, dx \tag{$\mathscr{PR}$}$$

or again

$$\operatorname*{Inf}_{\substack{u,p\\ p\in\mathscr{U}\cap L^\Phi_m\\ u=\mathscr{G}p}} \int_\Omega f^{**}(x, u(x); p(x))\, dx.$$

From Theorem 2.2 we immediately deduce

Proposition 2.5. *Let f be a normal integrand of $\Omega(\mathbf{P}^\ell \times \mathbf{R}^m)$, satisfying*

$$a(x) + \Phi(|\xi|) \leqslant f(x, s, \xi), \quad \textit{with } a \in L^1(\Omega). \tag{2.43}$$

Let $\mathscr{G}$ be a mapping which is a (Φ,β)-compactifier from L^Φ_m into L^β_l, $1 \leqslant \beta \leqslant \infty$, and let $\mathscr{U}$ be a weakly closed set of L^1. Then the relaxed problem $(\mathscr{PR})$ has at least one solution.

Remark 2.1. Since $f^{**} < f$, we always have $\min(\mathscr{PR}) \leqslant \inf(\mathscr{P})$. The subsequent chapters will elaborate further on the comparison between $(\mathscr{P})$ and $(\mathscr{PR})$; in particular they will treat the case when $\inf(\mathscr{P}) = \min(\mathscr{PR})$. For the present, we shall consider some examples of situations in the calculus of variations where Theorem 2.2 applies.

Remark 2.2. The relaxed problem $(\mathscr{PR})$ arises in simple cases as problem $(\mathscr{P}^{**})$, the bidual of $(\mathscr{P})$ with respect to suitable perturbations.

3. EXAMPLES

We shall describe several examples of mappings $\mathscr{G}$, satisfying (2.44), and arising naturally in variational problems. We shall use these examples in the following chapter, where the relationship between problems $(\mathscr{P})$ and $(\mathscr{PR})$ will be stated more precisely.

Example 1

Let Ω be a very regular open subset of $\mathbf{R}^n$, and A the Laplace operator:

$$Au = \Delta u. \tag{3.1}$$

For p given in $L^2(\Omega)$, there exists a unique u in $H_0^1(\Omega)$ such that:

$$Au = p \quad \text{a.e.} \tag{3.2}$$

and the mapping $p \mapsto u$ is linear and continuous from $L^2(\Omega)$ into $H_0^1(\Omega)$. Since the injection of $H_0^1(\Omega)$ into $L^2(\Omega)$ is compact (*cf.* Lions and Magenes, [1]), the mapping $p \mapsto u$ is linear and compact from $L^2(\Omega)$ into itself. If we call this operator $\mathscr{G}$, it then satisfies property (2.44) trivially with $\Phi(s) = s^2$: it is a (2,2)-compactifier.

Example 2

In a much more general way, $\mathscr{G}$ can be the Green operator of any regular elliptic problem.

Let A be a differential operator of order $2m$ in a very regular open subset $\Omega \subset \mathbf{R}^n$:

$$Au = A(x, D)u = \sum_{|\alpha|, |\beta| \leqslant m} (-1)^{|\alpha|} D^\alpha (a_{\alpha\beta}(x) D^\beta u) \tag{3.3}$$

with

$$a_{\alpha\beta} \in \mathscr{C}^{\infty}(\bar{\Omega}) \tag{3.4}$$

and

(3.5) the operator A is uniformly elliptic in $\bar{\Omega}$.

We denote by $\partial^j u/\partial n^j$ the j-th normal derivative of a function u on the boundary $\partial\Omega$, as it is defined by the usual trace theorems ($\partial^j u/\partial n^j \in H^{r-j-\frac{1}{2}}(\partial\Omega)$ if $u \in H^r(\Omega)$).

Under those assumptions, for p given in $L^2(\Omega)$, there exists a unique u in $H^{2m}(\Omega)$ satisfying:

$$\begin{cases} Au = p \quad \text{a.e. in } \Omega & (3.6) \\ \dfrac{\partial^j u}{\partial n^j} = 0 \quad \text{on } \partial\Omega, \text{ for } 0 \leqslant j \leqslant m-1 & (3.7) \end{cases}$$

and the mapping $\mathscr{G}: p \mapsto u$ is linear and continuous from $L^2(\Omega)$ into $H^{2m}(\Omega)$, and thus linear and compact from $L^2(\Omega)$ into itself. It is a (2,2)-compactifier.

Example 3

A being defined as in the preceding example, denote by Q the cylinder $\Omega \times (0, T)$ of $\mathbf{R}^{n+1}$, with T a positive real number, and consider the *parabolic equation*:

$$\begin{cases} \dfrac{\partial u}{\partial t} + Au = p \quad \text{a.e. in } Q & (3.8) \\ \dfrac{\partial^j u}{\partial n^j} = 0 \quad \text{on } \partial\Omega \times (0, T), \text{ for } 0 \leqslant j \leqslant m-1 & (3.9) \\ u(x, 0) = 0 \quad \text{a.e. in } \Omega. & (3.10) \end{cases}$$

For every $p \in L^2(Q)$ there is a unique solution $u \in H^{2m,1}(Q)$, and the mapping $\mathscr{G}: p \mapsto u$ is linear and compact from $L^2(Q)$ into itself.

Example 4

Assume moreover that A is coercive on $H_0^m(\Omega)$, and symmetric:

$$\exists c > 0 = \langle A\varphi, \varphi \rangle_{L^2} \geqslant c \|\varphi\|_{H_0^m}$$
$$a_{\alpha\beta} = a_{\beta\alpha}.$$

Consider the *hyperbolic equation*:

$$\frac{\partial^2 u}{\partial t^2} + Au = p \quad \text{a.e. in } Q \tag{3.11}$$

$$\frac{\partial^j u}{\partial n^j} = 0 \quad \text{on } \partial\Omega \times (0, T), \text{ for } 0 \leqslant j \leqslant m-1 \tag{3.12}$$

$$u(x, 0) = \frac{\partial u}{\partial t}(x, 0) = 0 \quad \text{a.e. in } \Omega. \tag{3.13}$$

For every $p \in L^2(Q)$ there is a unique solution $u \in H^{m,1}(Q)$, and the mapping $\mathscr{G}: p \mapsto u$ is linear and compact from $L^2(Q)$ into itself.

Example 5

Let us now consider cases where $\mathscr{G}$ is non-linear.

If Ω is bounded and if $1 < \gamma < \infty$, for p given in $L^{\gamma'}(\Omega)$, where $1/\gamma + 1/\gamma' = 1$, we verify with the help of Theorem 3.1 and Remark 3.4 from Chapter II that there exists a unique u in $W_0^{1,\gamma}(\Omega)$ which satisfies:

$$Au = -\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(\left|\frac{\partial u}{\partial x_i}\right|^{\gamma-2} \frac{\partial u}{\partial x_i}\right) = p. \tag{3.14}$$

We term $\mathscr{G}$ the non-linear mapping which results from this and which sends $L^{\gamma'}(\Omega)$ into $L^{\gamma}(\Omega)$. We now have the following result:

Lemma 3.1. *The mapping $\mathscr{G}$ defined by* (3.14) *is a* (γ', γ)*-compactifier.*

Proof. We note that:

$$\sum_{i=1}^{n} \int_\Omega \left|\frac{\partial u}{\partial x_i}\right|^{\gamma} \mathrm{d}x = \int_\Omega pu \,\mathrm{d}x \leqslant \left(\int_\Omega |p|^{\gamma'} \,\mathrm{d}x\right)^{1/\gamma'} \left(\int_\Omega |u|^{\gamma} \,\mathrm{d}x\right)^{1/\gamma}$$

whence by Poincare's inequality it follows that:

$$\|u\|_{W_0^{1,\gamma}} \leqslant c\, \|p\|_{L^{\gamma'}}.$$

If now a sequence p_m converges to p weakly in $L^{\gamma'}(\Omega)$, it is bounded in $L^{\gamma'}(\Omega)$ and the sequence of $u_m = \mathscr{G}p_m$'s is bounded in $W_0^{1,\gamma}(\Omega)$; by extracting a subsequence we can assume that u_m converges to a limit u weakly in $W_0^{1,\gamma}(\Omega)$ and hence strongly in $L^\gamma(\Omega)$. Then:

$$\langle Au_m, u_m \rangle = \sum_{i=1}^{m} \int_\Omega \left(\frac{\partial u_m}{\partial x_i}\right)^\gamma \mathrm{d}x$$

$$= \int_\Omega p_m u_m \,\mathrm{d}x \to \int_\Omega pu \,\mathrm{d}x$$

which obviously implies that:

$$\langle Au_m, u_m - u \rangle \to 0$$

and we deduce, as in Lemma II.3.3 that:

$$Au = -\sum_{i=1}^{m} \frac{\partial}{\partial x_i}\left(\left|\frac{\partial u}{\partial x_i}\right|^{\gamma-2} \frac{\partial u}{\partial x_i}\right) = p.$$

Thus $u_m = \mathscr{G}p_m$ converges to $u = \mathscr{G}p$ strongly in $L^\gamma(\Omega)$, which proves the lemma. ■

We can generalize this to the more general situation of Theorem II.3.1. Taking all the hypotheses of this theorem, together with (3.26), we assume moreover that $V \subset L^\gamma(\Omega)$ with compact injection and dense image, so that the dual V' of V contains $L^{\gamma'}(\Omega)$. In this case, for p given in $L^{\gamma'}(\Omega)$ there exists a unique u in V satisfying equation (3.2) of Chapter II, and the mapping $\mathscr{G}: f = p \mapsto u$ from $L^{\gamma'}(\Omega)$ into $L^\gamma(\Omega)$ is a (γ', γ)-compactifier.

Remark 3.1. We could give a great many more examples of operators $\mathscr{G}$ which are (α, β)-compactifiers by considering evolution equations, as in Lions [1], and inhomogeneous boundary-value problems as in Lions and Magenes [1]: a large number of examples from this work would allow us to define in a similar way the operators $\mathscr{G}$.

4. OPTIMAL CONTROL

We shall now apply Theorem 2.2 to the optimal control of systems governed by ordinary differential equations. In this example we shall show how to pass from a formulation of the "optimal control" type to a formulation of the "calculus of variations" type.

4.1. The optimal control problem

The evolution equation. We take a number $T > 0$, a *metrizable compact set* K and a *continuous* mapping φ from $[0,T] \times \mathbf{R}^n \times K$ into $\mathbf{R}^n$ such that: (denoting the unit-ball of $\mathbf{R}^n$ by B)

$$\text{(4.1)} \quad \left| \begin{array}{l} \textit{for all } \rho \geqslant 0, \textit{ there exists } k \geqslant 0 \textit{ such that, for } 0 \leqslant t \leqslant T \textit{ and } \omega \in K: \\ \forall y, y' \in \ell B, \qquad |\varphi(t, y, \omega) - \varphi(t, y', \omega)| \leqslant k\,|y - y'|; \end{array} \right.$$

$$\text{(4.2)} \quad \left| \begin{array}{l} \textit{there exists a constant } \ell \geqslant 0 \textit{ such that} \\ \forall (t, y, \omega) \in [0, T] \times \mathbf{R}^n \times K, \qquad |y \cdot \varphi(t, y, \omega)| \leqslant \ell(1 + |y|^2). \end{array} \right.$$

Let us choose $y_0 \in \mathbf{R}^n$. The system is governed by the ordinary differential equation:

$$\text{(4.3)} \qquad \left| \begin{array}{l} \mathrm{d}y(t)/\mathrm{d}t = \varphi(t, y(t), \omega(t)) \\ y(0) = y_0. \end{array} \right.$$

Lemma 4.1. *For any measurable mapping* $\omega : [0,T] \to K$, *the differential equation* (4.3) *has a unique solution* $y : [0,T] \to \mathbf{R}^n$, *and we have, for* $0 \leqslant t \leqslant T$:

$$\text{(4.4)} \qquad |y(t)| \leqslant (|y_0|^2 + 2\ell T)^{1/2}\, \mathrm{e}^{\ell T}.$$

Proof. Since φ has been assumed to be continuous there exists $\tau > 0$ sufficiently small for the equation (4.3) to have a solution defined on $[0,\tau]$. By virtue of the inequality (4.2), we now have, for $0 \leqslant t \leqslant \tau$

$$\frac{\mathrm{d}|y(t)|^2}{\mathrm{d}t} = 2y(t) \cdot \frac{\mathrm{d}y(t)}{\mathrm{d}t} \leqslant 2\ell(1 + |y(t)|^2)$$

$$|y(t)|^2 \leqslant |y_0|^2 + 2\ell\tau + \int_0^t 2\ell |y(s)|^2\, \mathrm{d}s.$$

By Gronwall's inequality, for $0 \leqslant t \leqslant \tau$:

$$|y(t)|^2 \leqslant (|y_0|^2 + 2\ell\tau)\, \mathrm{e}^{2\ell\tau} \leqslant (|y_0|^2 + 2\ell T)\, \mathrm{e}^{2\ell\tau}.$$

Set $\rho = (|y|^2 + 2\ell T)\mathrm{e}^{2\ell T}$. It can be seen that the solution lies in the bounded set ρB, independently of τ. We know that in this case the solution can be extended to the whole of $[0,T]$.

If there were two solutions of (4.3) on $[0,T]$, they would both have values in ρB, and by applying the Lipschitz condition (4.1), we could show them to coincide. ■

We term *control* any measurable mapping of $[0,T]$ into K. Once the control has been chosen, the unique solution y of (4.3) is the *trajectory* of the system, and $y(t)$ is its *state* at the instant t.

Lemma 4.2. *The set of trajectories of the system is relatively compact in* $\mathscr{C}([0,T];\mathbf{R}^n)$.

Proof. It is sufficient to note that the estimate (4.4) is independent of the control ω: along all the trajectories, $|y(t)|$ is bounded above by a constant ρ. Let us now denote by μ the maximum of the continuous function φ over the compact set $[0,T] \times \rho B \times K$. By equation (4.3) we have:

$$\left|\frac{\mathrm{d}y(t)}{\mathrm{d}t}\right| \leqslant \mu. \tag{4.5}$$

This means that all the trajectories are μ-Lipschitzian. In particular, they are equicontinuous and so form a relatively compact subset of $\mathscr{C}([0,T];\mathbf{R}^n)$ (Ascoli's theorem). ■

Constraints. We take a *closed* subset E of $[0,T] \times \mathbf{R}^n \times K$. For (t,y) given in $[0,T] \times \mathbf{R}^n$, we denote by $E_{t,y}$ the section:

$$E_{t,y} = \{\, \omega \in K \mid (t, y, \omega) \in E \,\}. \tag{4.6}$$

We term a control ω and its corresponding trajectory y *admissible* if they are linked by:

$$\forall t \in [0, T], \qquad (t, y(t), \omega(t)) \in E. \tag{4.7}$$

Cost. We take a Carathéodory function f of $[0,T] \times (R^n \times K)$ into $[0,+\infty[$. We associate with an admissible control ω and its corresponding trajectory y the cost function

$$\int_0^T f(t, y(t), \omega(t))\, \mathrm{d}t. \tag{4.8}$$

An admissible control $\bar{\omega}$ will be called *optimal* if it minimizes (4.8); we shall also call the corresponding trajectory $\bar{y}$ optimal.

Let us gather all the data:

$$(\mathscr{P}) \qquad \left|\begin{array}{l} \text{to minimize } \displaystyle\int_0^T f(t, y(t), \omega(t))\, \mathrm{d}t \\ \mathrm{d}y(t)/\mathrm{d}t = \varphi(t, y(t), \omega(\mathrm{d}t))\, \mathrm{d}t \\ (t, y(t), \omega(t)) \in E \\ y(0) = y_0 \end{array}\right.$$

4.2. Compactness of the set of trajectories

For $0 \leqslant t \leqslant T$ and $y \in \mathbf{R}^n$, we denote by $\Gamma_{t,y}$ the set of *admissible speeds* at the instant t at the point y:

$$\begin{aligned} \Gamma_{t,y} &= \{ \varphi(t, y, \omega) \mid (t, y, \omega) \in E \} \\ &= \varphi(t, y, E_{t,y}). \end{aligned}$$

This is a compact set in $\mathbf{R}^n$. If moreover it is convex and non-empty it can be shown that the set of admissible trajectories is compact and non-empty in $\mathscr{C}([0,T];\mathbf{R}^n)$. We shall here be content with part of this result.

Proposition 4.1. *If $\Gamma_{t,y}$ is convex for all $t \in [0,T]$ and all $y \in \mathbf{R}^n$, the set of admissible trajectories is compact in $\mathscr{C}([0,T];\mathbf{R}^n)$.*

Proof. In $\mathscr{C}([0,T];\mathbf{R}^n)$, the set of trajectories is relatively compact (Lemma 4.2), and it suffices therefore to show that the set of admissible trajectories is closed.

Hence, let $(y_k)_{k\in\mathbf{N}}$ be a sequence of admissible trajectories which converges uniformly to a continuous function $\bar{y}$. For all $k \in \mathbf{N}$ we have

$$\mathrm{d}y_k(t)/\mathrm{d}t = \varphi(t, y_k(t), \omega_k(t)) \tag{4.9}$$

$$(t, y_k(t), \omega_k(t)) \in E \tag{4.10}$$

for $0 \leqslant t \leqslant T$, and $y_k(0) = y_0$. We must show that $\bar{y}$ also is an admissible trajectory; we have immediately that $\bar{y}(0) = y_0$.

From the estimate (4.5), $\|\mathrm{d}y_k/\mathrm{d}t\| \leqslant \mu$ for $k \in \mathbf{N}$. By extracting a subsequence we can thus assume that $\mathrm{d}y_k/\mathrm{d}t$ converges to $\mathrm{d}\bar{y}/\mathrm{d}t$ in the topology $\sigma(L^2,L^2)$. By Mazur's lemma there exists a sequence of convex combinations $\sum_{n=k}^{N} \alpha_n \mathrm{d}y_n/\mathrm{d}t$ which converges to $\mathrm{d}\bar{y}/\mathrm{d}t$ in $L^2_n(0,T)$ as $k \to \infty$. We can therefore extract a subsequence which converges simply to $d\bar{y}/\mathrm{d}t$ on $[0,T]$. We thus have for $t \in [0,T]$:

$$\forall k \in \mathbf{N}, \qquad \frac{\mathrm{d}\bar{y}}{\mathrm{d}t}(t) \in \overline{\mathrm{co}} \left\{ \frac{\mathrm{d}y_n(t)}{\mathrm{d}t} \mid n \geqslant k \right\}. \tag{4.11}$$

But we can summarize (4.9) and (4.10) by:

$$\frac{\mathrm{d}y_k(t)}{\mathrm{d}t} \in \Gamma_{t,y(t)} \quad \text{for} \quad 0 \leqslant t \leqslant T \tag{4.12}$$

and on substituting into (4.11):

$$(4.13) \qquad \forall k \in \mathrm{N}, \qquad \frac{\mathrm{d}\bar{y}(t)}{\mathrm{d}t} \in \overline{\mathrm{co}} \bigcup_{n \geqslant k} \Gamma_{t, y_n(t)} \quad \text{for} \quad 0 \leqslant t \leqslant T.$$

Let us fix $t \in [0, T]$ and take $\varepsilon > 0$. Since φ is continuous and K is compact there exists $\eta > 0$ such that:

$$|y - \bar{y}(t)| \leqslant \eta \Rightarrow |\varphi(t, y, \omega) - \varphi(t, \bar{y}(t), \omega)| \leqslant \varepsilon \qquad \forall \omega \in K.$$

Since y_k converges uniformly to $\bar{y}$, we can take $k \in \mathrm{N}$ large enough for:

$$\forall n \geqslant k, \qquad \Gamma_{t, y_n(t)} \subset \Gamma_{t, \bar{y}(t)} + \varepsilon B,$$

$$\bigcup_{n \geqslant k} \Gamma_{t, y_n(t)} \subset \Gamma_{t, \bar{y}(t)} + \varepsilon B.$$

The right-hand side is the sum of two convex compact sets. It is thus a convex compact set and:

$$(4.14) \qquad \overline{\mathrm{co}} \bigcup_{n \geqslant k} \Gamma_{t, y_n(t)} \subset \Gamma_{t, y(t)} + \varepsilon B.$$

Substituting this into (4.13):

$$\frac{\mathrm{d}\bar{y}(t)}{\mathrm{d}t} \in \Gamma_{t, \bar{y}(t)} + \varepsilon B.$$

It only remains to let ε tend to zero. Since $\Gamma_{t, \bar{y}(t)}$ is closed, we obtain in the limit for $0 \leqslant t \leqslant T$:

$$(4.15) \qquad \frac{\mathrm{d}\bar{y}(t)}{\mathrm{d}t} \in \Gamma_{t, \bar{y}(t)},$$

$$(4.16) \qquad \frac{\mathrm{d}\bar{y}(t)}{\mathrm{d}t} \in \{ \varphi(t, \bar{y}(t), \omega) \mid (t, \bar{y}(t), \omega) \in E \}.$$

There exists a Borel subset $N \subset [0, T]$, with null measure, such that the restriction of $\mathrm{d}\bar{y}/\mathrm{d}t$ to N is Borel. We can then define a Borel subset of $\complement N \times K$ by:

$$(4.17) \qquad G = \{ (t, \omega) \mid (t, \bar{y}(t), \omega) \in E \text{ and } \varphi(t, \bar{y}(t), \omega) \in \mathrm{d}\bar{y}(t)/\mathrm{d}t \}.$$

By Corollary 1.7 there exists a measurable selection $\bar{\omega}$ of G. In suitably extending it over N, we obtain a measurable mapping $\bar{\omega}$ from $[0, T]$ into A which satisfies

$$(4.18) \qquad \mathrm{d}\bar{y}(t)/\mathrm{d}t = \varphi(t, \bar{y}(t), \bar{\omega}(t))$$

$$(4.19) \qquad (t, \bar{y}(t), \bar{\omega}(t)) \in E$$

and hence $\bar{y}$ is indeed an admissible trajectory. ∎

4.3. Existence of optimal controls

Let us define a function g from $[0,T] \times \mathbf{R}^n \times \mathbf{R}^n$ into $\mathbf{R}$ by:

$$(4.20) \quad g(t, y, \eta) = \min \{ f(t, y, \omega) \mid (t, y, \omega) \in E \text{ and } \varphi(t, y, \omega) = \eta \}.$$

Clearly, $g(t,y,\eta) \in \mathbf{R}_+$ if $\eta \in \Gamma_{t,y}$ and $g(t,y,\eta) = +\infty$ otherwise.

Lemma 4.3. *g is a positive normal integrand of* $[0,T] \times (\mathbf{R}^n \times \mathbf{R}^n)$ *into* $\bar{\mathbf{R}}$.

Proof. Let us take $\varepsilon > 0$. From the Scorza–Dragoni theorem, there exists a compact subset $C_\varepsilon \subset [0,T]$ such that $\text{meas}([0,T] - C_\varepsilon) \leqslant \varepsilon$ and for which the restriction of f to $C_\varepsilon \times \mathbf{R}^n \times K$ is continuous. We now show that the restriction of g to $C_\varepsilon \times \mathbf{R}^n \times \mathbf{R}^n$ is l.s.c., which will prove that g is a normal integrand (Theorem 1.2).

Let (t_n, y_n, η_n) be a sequence of $C_\varepsilon \times \mathbf{R}^n \times \mathbf{R}^n$ which converges to $(\bar{t}, \bar{y}, \bar{\eta})$. Set:

$$(4.21) \qquad \varliminf_{n \to \infty} g(t_n, y_n, \eta_n) = \ell.$$

We wish to show that $g(\bar{t},\bar{y},\bar{\eta}) \leqslant \ell$. If $\ell = +\infty$, this is trivial. If ℓ is finite, we may assume that $g(t_n, y_n, \eta_n)$ is finite for all $n \in \mathbf{N}$ and converges to ℓ. From (4.20), there exists ω_n such that:

$$(4.22) \qquad (t_n, y_n, \omega_n) \in E \quad \text{and} \quad \varphi(t_n, y_n, \omega_n) = \eta_n$$

$$(4.23) \qquad f(t_n, y_n, \omega_n) = g(t_n, y_n, \eta_n).$$

Since K is compact, we can extract from the sequence ω_n, $n \in \mathbf{N}$, a subsequence $\omega_{n'}$ converging to a $\bar{\omega} \in K$. We can then pass to the limit in (4.22) and (4.23):

$$(4.24) \qquad (\bar{t}, \bar{y}, \bar{\omega}) \in E \quad \text{and} \quad \varphi(\bar{t}, \bar{y}, \bar{\omega}) = \bar{\eta}$$

$$(4.25) \qquad f(\bar{t}, \bar{y}, \bar{\omega}) = \ell.$$

Whence necessarily, by (4.20):

$$(4.26) \qquad g(\bar{t}, \bar{y}, \bar{\omega}) \leqslant \ell. \ ∎$$

We can now state a sufficient condition for the existence of optimal controls:

Proposition 4.2. *We assume the previous hypotheses. If for all $(t,y) \in [0,T] \times \mathbf{R}^n$, the function $g(t,y,.)$ is convex from $\mathbf{R}^n$ into $\overline{\mathbf{R}}$, then there exists at least one admissible optimal control.*

Let us explain this condition. To say that $g(t,y,.)$ is convex means that its epigraph is convex, i.e. that for $0 \leqslant t \leqslant T$ and $y \in \mathbf{R}^n$, the set

(4.27) $\{ (\eta, a) \in \mathbf{R}^n \times \mathbf{R} \mid \exists\, \omega \in E_{t,y} : \varphi(t, y, \omega) = \eta \text{ and } a \geqslant f(t, y, \omega) \}$

is convex in $\mathbf{R}^n \times \mathbf{R}$. This implies that its horizontal projection $\Gamma_{t,y}$ is convex, i.e. the set of admissible speeds is convex and compact.

Proof. Let ω_n, $n \in \mathbf{N}$, be a minimizing sequence of admissible controls and y_n, $n \in \mathbf{N}$, their corresponding trajectories. By Lemma 4.2, there exists a constant μ such that for all $n \in \mathbf{N}$,

$$\left\| \frac{\mathrm{d}y_n}{\mathrm{d}t} \right\| \leqslant \mu. \tag{4.28}$$

From Proposition 4.1, possibly by extracting a subsequence, we may assume that there exists an admissible trajectory $\bar{y}$ such that:

$$y_n \to \bar{y} \quad \text{uniformly} \tag{4.29}$$

$$\mathrm{d}y_n/\mathrm{d}t \to \mathrm{d}\bar{y}/\mathrm{d}t \quad \text{for} \quad \sigma(L^\infty, L^1). \tag{4.30}$$

It only remains to apply Theorem 2.1 to the integrand

$$\tilde{g}(t, y, \eta) = g(t, y, \eta) + \chi_{\mu B}(\eta).$$

We obtain:

$$\int_0^T \tilde{g}\Big(t, \bar{y}(t), \frac{\mathrm{d}\bar{y}}{\mathrm{d}t}(t)\Big)\, \mathrm{d}t \leqslant \varliminf_{n\to\infty} \int_0^T \tilde{g}\Big(t, y_n(t), \frac{\mathrm{d}y_n}{\mathrm{d}t}(t)\Big)\, \mathrm{d}t. \tag{4.31}$$

Taking (4.28) and (4.20) into account:

$$\begin{aligned} \int_0^T g\Big(t, \bar{y}(t), \frac{\mathrm{d}\bar{y}}{\mathrm{d}t}(t) &\leqslant \varliminf_{n\to\infty} \int_0^T g\Big(t, y_n(t), \frac{\mathrm{d}y_n}{\mathrm{d}t}(t)\Big)\, \mathrm{d}t \\ &\leqslant \varliminf_{n\to\infty} \int_0^T f(t, y_n(t), \omega_n(t))\, \mathrm{d}t. \end{aligned} \tag{4.32}$$

The right-hand side is equal to $\inf(\mathscr{P})$. By the measurable selection theorem (Cor. 1.7), it is easily shown that there exists a measurable mapping $\bar{\omega}: [0, T] \to K$ such that for $0 \leqslant t \leqslant T$:

$$(t, \bar{y}(t), \bar{\omega}(t)) \in E \quad \text{and} \quad \varphi(t, \bar{y}(t), \bar{\omega}(t)) = \frac{\mathrm{d}\bar{y}}{\mathrm{d}t}(t)$$

$$f(t, \bar{y}(t), \bar{\omega}(t)) = g\left(t, \bar{y}(t), \frac{\mathrm{d}\bar{y}}{\mathrm{d}t}(t)\right).$$

Thus $\bar{\omega}$ is an admissible control and on substituting into (4.32) we obtain:

$$\int_0^T f(t, \bar{y}(t), \bar{\omega}(t)) \, \mathrm{d}t \leqslant \inf(\mathscr{P}) \tag{4.33}$$

and $\bar{\omega}$ is an admissible optimal control. ∎

ERRATUM

The authors are grateful to Michel Valadier for pointing out a mistake in the proof of Lemma VIII.2.1. The lines after formula (2.8) should be changed as follows:

Since $\varphi(\bar{e}, \xi) > \ell(\xi) \ \forall \xi \in B(0, M)$, there exists, for each ξ, some neighborhood $V_\xi \times W_\xi$ of $(\bar{e}, \xi)$ such that:

$$(e', \xi') \in V_\xi \times W_\xi \Longrightarrow \varphi(e', \xi') > \ell(\xi').$$

Let $\xi_1, \ldots, \xi_n$ be such that the W_{ξ_i} cover the compact set $B(0, M)$. We then have:

$$e' \in \cap V_{\xi_i} \text{ and } |\xi'| \leq M \Longrightarrow \varphi(e', \xi') > \ell(\xi').$$

This replaces formula (2.12), at which point the proof resumes without further changes.

CHAPTER IX

Relaxation of Non-convex variational Problems (I)

Orientation

We shall continue the study of non-convex problems in the calculus of variations begun in the preceding chapter. The problems which we shall deal with are of the type:

$$(\mathscr{P}) \qquad \left| \begin{array}{l} \displaystyle\operatorname*{Inf}_{u,p} \int_\Omega f(x, u(x), p(x))\,\mathrm{d}x \\ u = \mathscr{G}\,p, \quad p \in L^1(\Omega)^m. \end{array} \right.$$

We make no hypothesis concerning convexity, and as a result there is in general no solution to these problems. We therefore consider the corresponding relaxed problem:

$$(\mathscr{P}\mathscr{R}) \qquad \left| \begin{array}{l} \displaystyle\operatorname*{Inf}_{u,p} \int_\Omega f^{**}(x, u(x); p(x))\,\mathrm{d}x \\ u = \mathscr{G}\,p, \quad p \in L^1(\Omega)^m. \end{array} \right.$$

This problem was introduced in the last chapter, where we showed that Min $(\mathscr{P}\mathscr{R}) \leqslant$ Inf $(\mathscr{P})$. We shall take this study much further, showing that Min $(\mathscr{P}\mathscr{R}) =$ Inf $(\mathscr{P})$, and deducing that the solutions of problem $(\mathscr{P}\mathscr{R})$ are the cluster points of the minimizing sequences of problem $(\mathscr{P})$. The solutions of the relaxed problem thus appear as "generalized solutions" of the original problem.

To establish these results, we proceed by stages. We start by studying the case where u and p (the state and the control in terms of optimal control) occur separately in the integrand, which is thus of the form $g(x, u(x)) + f(x, p(x))$. The corresponding result is given in Section 3. We arrive at this result by studying the mapping $F: p \mapsto \int_\Omega f(x, p(x))\,\mathrm{d}x$ from $L^1(\Omega)^m$ into $\overline{\mathbf{R}}$: we show in Section 1 that the Γ-regularization F^{**} of F coincides with its l.s.c. regularization $\bar{F}$, and in Section 2 we calculate $F^{**}(p) = \int_\Omega f^{**}(x; p(x))\,\mathrm{d}x$. Finally, in Section 4 we consider the general case by combining the above results with a property of equi-integrability.

1. IDENTITY OF THE Γ-REGULARIZATION AND OF THE L.S.C. REGULARIZATION

1.1. An approximation result

We fix the bounded open subset $\Omega \subset \mathbf{R}^n$, and a family $(\alpha_k)_{1 \leqslant k \leqslant m}$ of real positive numbers with sum 1. For $i \in \mathbf{N}$, we denote by $\mathscr{K}_i$ the set of hypercubes of Ω, whose edges have length 2^{-i} and are parallel to the axes of the co-ordinates and whose vertices have multiples of 2^{-i} as co-ordinates. In other words, $K \in \mathscr{K}_i$ if:

$$K = \prod_{j=1}^{n} [m_j 2^{-i}, (m_j + 1)2^{-i}] \subset \Omega, \quad \text{where the } m_j\text{'s} \in \mathbf{N}.$$

We denote by B_i the union of all $K \in \mathscr{K}_i$. Clearly, $B_i \subset B_{i+1} \subset \Omega$ and $\Omega = \bigcup_{i=1}^{\infty} B_i$. In particular, meas $\Omega = \lim_{i \to \infty}$ meas B_i.

For fixed i, we shall divide B_i into m subsets B_i^k, $1 \leqslant k \leqslant m$, corresponding to the m numbers α_k. To do this, we shall divide each hypercube $K \in \mathscr{K}_i$ into m slices K^k, perpendicular to the first axis of co-ordinates, the thickness of the kth cut being $\alpha_k \cdot 2^{-i}$. In other words:

$$K = \prod_{j=1}^{n} [m_j 2^{-i}, (m_j + 1)2^{-i}] \subset \Omega \tag{1.1}$$

$$K^k = \left[\left(m_1 + \sum_{\ell=1}^{k-1} \alpha_\ell\right)2^{-i}, \left(m_1 + \sum_{\ell=1}^{k} \alpha_\ell\right)2^{-i}\right] \times \prod_{j=1}^{n} [m_j 2^{-i}, (m_j + 1)2^{-i}] \tag{1.2}$$

$$K = \bigcup_{k=1}^{m} K^k \tag{1.3}$$

$$\alpha_k \text{ meas } K = \text{meas } K^k. \tag{1.4}$$

For fixed $i \in \mathbf{N}$ and $k \in \{1, \ldots, m\}$, we denote by B_i^k the union of the K^k, for $K \in \mathscr{K}_i$. Note that the hypercubes K^k, for $K \in \mathscr{K}_i$ and $1 \leqslant k \leqslant m$ are not exactly disjoint. We shall therefore term N_i the union of the faces of these hypercubes, which is thus a set of null measure, and we set $N = \bigcup_{i \in \mathbf{N}} N_i$, which is hence of null measure. Then, for all fixed $i \in \mathbf{N}$, the $B_i^k \cap \complement N$, $1 \leqslant k \leqslant m$, are disjoint.

Definition 1.1. *Let $f=(f_1,\ldots,f_m)\in[L^1(\Omega)]^m$. For all $i\in\mathbf{N}$, we define a measurable mapping $T_i f$ of Ω into $\mathbf{R}$ by:*

$$
(1.5)\qquad \left|\begin{array}{ll} T_i f(x) = f_k(x) & \text{if } x\in B_i^k \cap \complement N \\ T_i f(x) = f_1(x) & \text{if } x\in N \cup \left(\Omega - \bigcup_{k=1}^{m} B_i^k\right) \end{array}\right.
$$

Proposition 1.1. *For all $i\in\mathbf{N}$, the mapping T_i is linear and continuous from $[L^1(\Omega)]^m$ into $L^1(\Omega)$. Moreover,*

$$
(1.6)\qquad \lim_{i\to\infty}\int_\Omega T_i f = \sum_{k=1}^{m}\alpha_k\int_\Omega f_k.
$$

Proof. From equations (1.5) it follows immediately that T_i is linear and that:

$$
(1.7)\qquad \|T_i f\|_1 \leqslant \sum_{k=1}^{m}\|f_k\|_1.
$$

The linear mappings T_i, for $i\in\mathbf{N}$, are thus equi-continuous from $(L^1)^m$ into L^1, and the linear functionals $f\to\int_\Omega T_i f$ are thus equi-continuous over $(L^1)^m$. By Ascoli's theorem, it is sufficient to prove the convergence (1.6) for all f belonging to a dense subset of $(L^1)^m$.

We say that a function is $\mathscr{K}$-tiered if it is almost everywhere equal to a finite linear combination of characteristic functions of hypercubes of $\mathscr{K}_i$, for a sufficiently large value of i. Every uniformly continuous function over Ω is a uniform limit of $\mathscr{K}$-tiered functions. But the continuous functions with compact support are uniformly continuous over Ω and dense in L^1. The $\mathscr{K}$-tiered functions are thus dense in L^1 and the f with $\mathscr{K}$-tiered components are dense in $(L^1)^m$.

Hence let $f=(f_1,\ldots,f_m)$ with $\mathscr{K}$-tiered components. Let us choose i_0 sufficiently large so that the f_k, $1\leqslant k\leqslant m$ are all constant on the hypercubes of $\mathscr{K}_i$, for all $i\geqslant i_0$:

$$
(1.8)\qquad f_k = \sum_{\ell=1}^{r} f_{k_\ell}\mathbf{1}_{K_\ell}, \qquad \text{with } K_\ell\in\mathscr{K}_i.
$$

We thus have

$$
(1.9)\qquad T_i f = \sum_{k=1}^{m}\sum_{\ell=1}^{r} f_k\,\mathbf{1}_{K_\ell^k} \quad \text{a.e.}
$$

$$
(1.10)\qquad \int_\Omega T_i f = \sum_{k=1}^{m}\sum_{\ell=1}^{r} f_{k_\ell}\,\text{meas}\,K_\ell^k.
$$

From (1.4), let:

$$\begin{aligned}\int_\Omega T_i f &= \sum_{k=1}^{\ell} \sum_{\ell=1}^{r} f_{k_\ell} \alpha_k \text{ meas } K_\ell \\ &= \sum_{k=1}^{m} \alpha_k \sum_{\ell=1}^{r} f_{k_\ell} \text{ meas } K_\ell \\ &= \sum_{k=1}^{\ell} \alpha_k \int_\Omega f_k.\end{aligned}$$

This latter equality, which is true for all $i \geqslant i_0$, establishes *a fortiori* the convergence (1.6) for the f with $\mathscr{K}$-tiered components, and thus for all the $f \in (L^1)^m$. ∎

Corollary 1.1. *Let* $1 \leqslant \alpha \leqslant \infty$ *and* $1/\alpha + 1/\alpha' = 1$. *The mapping* T_i *is linear and continuous from* $[L^\alpha(\Omega)]^m$ *into* $L^\alpha(\Omega)$, *and we have*:

$$\sum_{k=1}^{m} \alpha_k f_k = \lim_{i \to \infty} T_i f \quad \textit{for} \quad \sigma(L^\alpha, L^{\alpha'}) \tag{1.11}$$

Proof. We have

$$\| T_i f \|_\alpha \leqslant \sum_{k=1}^{m} \| f_k \|_\alpha, \tag{1.12}$$

whence the continuity, the linearity being obvious. To obtain (1.11), we take any $h \in L^{\alpha'}$, and we apply Proposition 1.1 to $hf = (hf_1, \ldots, hf_m) \in (L^1)^m$. We obtain:

$$\lim_{i \to \infty} \int_\Omega T_i(hf) = \sum_{k=1}^{m} \alpha_k \int_\Omega h f_k. \tag{1.13}$$

Now, from (1.5), it follows that:

$$T_i(hf) = h(T_i f). \tag{1.14}$$

Substituting (1.14) into (1.13), we obtain:

$$\lim_{i \to \infty} \int_\Omega h T_i f = \int_\Omega h \sum_{k=1}^{m} \alpha_k f_k \tag{1.15}$$

which means that $T_i f$ tends to $\sum_{k=1}^{m} \alpha_k f_k$ for $\sigma(L^\alpha, L^{\alpha'})$. ∎

Proposition 1.1 will usually be most useful to us in the following form.

Corollary 1.2. *Let $u_1, \ldots, u_m$ be m mappings from Ω into some arbitrary set $\mathscr{E}$ and let f be a numerical function on $\Omega \times \mathscr{E}$ such that the functions $x \to f(x, u_k(x))$ belong to $L^1(\Omega)$ for all k. Let us define $T_i u: \Omega \to \mathscr{E}$ by:*

$$(1.16) \qquad \begin{cases} T_i u(x) = u_k(x) & \text{if } x \in B_i^k \cap \complement N \\ T_i u(x) = u_1(x) & \text{if } x \in N \cup \left(\Omega - \bigcup_{k=1}^{m} B_i^k\right). \end{cases}$$

Then, when $i \to \infty$:

$$(1.17) \qquad \int_\Omega f(x, T_i u(x))\, dx \to \sum_{k=1}^{m} \alpha_k \int_\Omega f(x, u_k(x))\, dx.$$

Proof. It is sufficient to apply Proposition 1.1 to the m functions:

$$(1.18) \qquad f_k(x) = f(x, u_k(x)).$$

We obtain:

$$(1.19) \qquad \int_\Omega T_i f \to \sum_{k=1}^{m} \alpha_k \int_\Omega f_k.$$

Comparing (1.5) and (1.16), we have:

$$(1.20) \qquad T_i f(x) = f(x, T_i u(x))$$

and substituting (1.20) into (1.19), we obtain (1.17). ■

Remark 1.1. To conclude, let us note that all these results can be extended unchanged to the case where Ω is a bounded measurable subset of $\mathbf{R}^m$, and to the case where Ω is a compact manifold with boundary of dimension n and of class C^0. In the first case, it is sufficient to embed Ω in a bounded open set Ω' of $\mathbf{R}^n$, and to extend by zero on $\Omega' - \Omega$ the function given on Ω. In the second case, we proceed by local charts.

1.2. Application to the Γ-regularization

First of all we recall some results from Chapter I. If V is a separated l.c.s., if F is a mapping of V into $\bar{\mathbf{R}}$, the Γ-regularization of F, denoted by F^{**}, will be the largest convex l.s.c. function everywhere less than F, or again the upper bound of all continuous affine functions everywhere less than F. We have:

$$(1.21) \qquad \text{epi } F^{**} = \overline{\text{co}}\, \text{epi } F.$$

In the same way, we define the l.s.c. regularization of F, denoted by $\bar{F}$: this is the largest l.s.c. function everywhere less than F. We have seen that:

$$\text{epi}\,\bar{F} = \overline{\text{epi}\,F} \tag{1.22}$$

$$\bar{F}(u) = \varliminf_{v \to u} F(v). \tag{1.23}$$

Clearly, $F^{**} \leqslant \bar{F}$. The technicalities of the preceding paragraph will allow us to determine explicitly some (non-convex) functionals for which $F^{**} = \bar{F}$, the space V being endowed with a weak topology.

We shall still denote by Ω a bounded open subset of $\mathbf{R}^n$, and we are given a positive normal integrand f of $\Omega \times \mathbf{R}^m$. We shall take for V the space $L^\alpha(\Omega)^m$, $1 \leqslant \alpha \leqslant \infty$, which we shall endow with the weak topology $\sigma(L^\alpha, L^{\alpha'})$, where $1/\alpha + 1/\alpha' = 1$. We define a function $F \colon V \to \bar{\mathbf{R}}_+$ by:

$$F(u) = \int_\Omega f(x, u(x))\,\mathrm{d}x. \tag{1.24}$$

Proposition 1.2. *When f is a positive normal integrand and $L^\alpha(\Omega)^m$ is endowed with the topology $\sigma(L^\alpha, L^{\alpha'})$, the Γ-regularization of the function F defined by* (1.24) *coincides with its l.s.c. regularization*:

$$F^{**} = \bar{F}.$$

Proof. We already have $F^{**} \leqslant \bar{F}$. We now prove the converse inequality. Let:

$$(\bar{v}, \bar{a}) \in \text{epi}\ F^{**}.$$

By (1.21)

$$(\bar{v}, \bar{a}) \in \overline{\text{co}}\,(\text{epi}\,F). \tag{1.25}$$

For all $\varepsilon > 0$ and every neighbourhood $\mathscr{V}$ of the origin in the topology $\sigma(L^\alpha, L^{\alpha'})$ there exists a family $(u_k, a_k)_{1 \leqslant k \leqslant m}$ of $L^\alpha_m \times \mathbf{R}$ and m numbers $\alpha_k \geqslant 0$ with $\sum_{k=1}^m \alpha_k = 1$ such that:

$$\forall k, \qquad (u_k, a_k) \in \text{epi}\,F \tag{1.26}$$

$$\bar{u} - \sum_{k=1}^m \alpha_k u_k \in \mathscr{V} \tag{1.27}$$

$$\left| \bar{a} - \sum_{k=1}^m \alpha_k a_k \right| \leqslant \varepsilon. \tag{1.28}$$

By (1.26) we have $F(u_k) < +\infty$ for all k. The function $x \to f(x, u_k(x))$ thus belongs to L^1 for any k whatever. From the Corollaries 1.1 and 1.2, we can take i sufficiently large that $T_i u$ satisfies:

$$T_i u - \sum_{k=1}^{m} \alpha_k u_k \in \mathscr{V} \tag{1.29}$$

$$\left| \int_\Omega f(x, T_i u(x))\, dx - \sum_{k=1}^{m} \alpha_k \int_\Omega f(x, u_k(x))\, dx \right| \leqslant \varepsilon. \tag{1.30}$$

By taking (1.27) and (1.29) we then have:

$$T_i u - \bar{u} \in 2\mathscr{V}. \tag{1.31}$$

Then, on taking (1.28) and (1.30)

$$\bar{a} + 2\varepsilon \geqslant \sum_{k=1}^{m} \alpha_k a_k + \varepsilon \geqslant F(T_i u) \tag{1.32}$$

which means that:

$$(T_i u, \bar{a} + 2\varepsilon) \in \overline{\text{epi}\, F}. \tag{1.33}$$

Since ε and V are arbitrary, we deduce from (1.31) and (1.33) that $(\bar{u}, \bar{a}) \in \overline{\text{epi}\, F}$, and hence that

$$\text{epi}\, F^{**} \subset \overline{\text{epi}\, F}, \tag{1.34}$$

and thus $\bar{F} \leqslant F^{**}$. ■

Proposition 1.3. *We assume that F satisfies the hypotheses of Proposition* 1.2 *and for all* $\lambda \in \mathbf{R}$ *we set*

$$S_\lambda = \{ u \in L^\alpha_m \mid F(u) \leqslant \lambda \}. \tag{1.35}$$

Let $G \in \Gamma_0(L^\alpha_m)$ *be a function whose restriction to* S_λ *is continuous with respect to* $\sigma(L^\alpha, L^{\alpha'})$ *for any* λ. *We have*:

$$(G + F)^{**} = G + F^{**} = \overline{(G + F)}.$$

Proof. First of all let us assume that $G \geqslant 0$. We have the inequality

$$G + F^{**} \leqslant (G + F)^{**} \leqslant \overline{(G + F)}. \tag{1.36}$$

Indeed, $G + F^{**}$ is a l.s.c. convex function which is everywhere less than $(G+F)$ and so less than its Γ-regularization. To obtain the converse inequality, we start with equation (1.23):

$$\overline{(G + F)}(u) = \varliminf_{v \to u} (G(v) + F(v)). \tag{1.37}$$

Taking $\lambda > \varliminf_{v\to u}(G(v) + F(v)) \geqslant \varliminf_{v\to u} F(v)$, we clearly have:

$$\varliminf_{v \to u} (G(v) + F(v)) = \varliminf_{\substack{v\in S_\lambda \\ v \to u}} (G(v) + F(v)) \tag{1.38}$$

$$\varliminf_{v \to u} F(v) = \lim_{\substack{v\in S_\lambda \\ v \to u}} F(v). \tag{1.39}$$

Whence on substituting into (1.37) and making use of the continuity of G on S_λ and of Proposition 1.6:

$$\overline{(G + F)}(u) = G(u) + F^{**}(u).$$

The inequalities (1.36) are thus equalities, and the proposition has been proved for the case $G \geqslant 0$. If G is now any function in $\Gamma_0(L^\alpha_m)$, it has an affine continuous minorant ℓ. We thus have, on setting $H = G - \ell \geqslant 0$

$$\begin{aligned}(G + F)^{**} &= (\ell + H + F)^{**}\\ &= \ell + (H + F)^{**}\\ &= \ell + H + F^{**}\\ &= G + F^{**}. \quad \blacksquare\end{aligned}$$

Corollary 1.3. *With the hypotheses of Proposition* 1.3, *we have*:

$$\inf_{u\in L^\alpha} (G(u) + F(u)) = \inf_{u\in L^\alpha} (G(u) + F^{**}(u)). \tag{1.40}$$

It suffices to note that $(G+F)$ and its Γ-regularization have the same lower bound.

Remark 1.2. These results, which depend essentially on Proposition 1.1, remain valid under the more general conditions indicated in Remark 1.1.

2. CALCULATION OF THE Γ-REGULARIZATION

Corollary 1.3 will be of use to us in "making convex" those problems in the calculus of variations which are stated in the form $\inf (G + F)$. To turn this to the best account, we have to make F^{**}, and hence F^*, explicit. Let us resume our notation. Ω is a bounded open subset of $\mathbf{R}^n$, and f a non-negative normal integrand of $\Omega \times \mathbf{R}^m$. We put the space $L^\alpha(\Omega)^m$ $1 \leqslant \alpha \leqslant \infty$, in duality with $L^{\alpha'}(\Omega)^m$, $1/\alpha + 1/\alpha' = 1$, and we define a non-negative function F on L^α_m by:

(2.1)
$$F(u) = \int_\Omega f(x, u(x))\,\mathrm{d}x.$$

Whence its polar F^* on $L^{\alpha'}_m$ is,

(2.2)
$$F^*(u^*) = \sup_{u \in L^\alpha_m}\left[\langle u, u^* \rangle - \int_\Omega f(x, u(x))\,\mathrm{d}x\right]$$
$$= \sup_{u \in L^\alpha_m}\left[\int_\Omega [u(x)u^*(x) - f(x, u(x))]\,\mathrm{d}x\right].$$

2.1. Calculation of the polar F^*

Our principal tool will be the measurable selection theorem (Theorem VIII.1.2).

Proposition 2.1. *Let F be the functional* (2.1) *defined on $L^\alpha(\Omega)^m$, $1 \leqslant \alpha \leqslant \infty$, and let us assume that there exists $u_0 \in L^\infty_m$ such that $F(u_0) < +\infty$. Then for all $u^* \in L^{\alpha'}(\Omega)^m$, we have:*

(2.3)
$$F^*(u^*) = \int_\Omega f^*(x; u^*(x))\,\mathrm{d}x.$$

Proof. We fix $u^* \in L^{\alpha'}_m$ and we introduce the functions:

(2.4)
$$\Phi(x) = \sup_{\xi \in \mathbf{R}^m} \{ \xi u^*(x) - f(x, \xi) \}$$

(2.5)
$$\Phi_n(x) = \max_{|\xi| \leqslant n} \{ \xi u^*(x) - f(x, \xi) \}.$$

Clearly, the sequence Φ_n is increasing, and $\Phi_n(x)$ converges to $\Phi(x)$ for all $x \in \Omega$. Furthermore, for all $n \geqslant \|u_0\|_\infty$, we have:

(2.6)
$$\Phi_n(x) \geqslant u_0(x)u^*(x) - f(x, u_0(x))$$

the function appearing in the right-hand side being integrable over Ω. From Theorem VIII.1.2, for all $n \in \mathbf{N}$, there exists a measurable mapping $\bar{u}_n \colon \Omega \to \mathbf{R}^m$, such that $\|\bar{u}_n\|_\infty \leqslant n$ and:

$$\Phi_n(x) = \bar{u}_n(x)u^*(x) - f(x, \bar{u}_n(x)). \tag{2.7}$$

In particular, Φ_n is measurable for all n, hence Φ is measurable and

$$\int_\Omega \Phi(x)\,\mathrm{d}x = \sup_n \int_\Omega \Phi_n(x)\,\mathrm{d}x \tag{2.8}$$

$$\int_\Omega \Phi(x)\,\mathrm{d}x = \sup_n \left| \int_\Omega [\bar{u}_n(x)u^*(u) - f(x, \bar{u}_n(x))]\,\mathrm{d}x \right|. \tag{2.9}$$

Since $\bar{u}_n \in L^\infty_m \subset L^\alpha_m$ for all n, we obtain:

$$\int_\Omega \Phi(x)\,\mathrm{d}x \leqslant \sup_{u \in L^\alpha_m} \left| \int_\Omega [u(x)u^*(x) - f(x, u(x))]\,\mathrm{d}x \right| \tag{2.10}$$

or, by (2.2):

$$\int_\Omega \Phi(x)\,\mathrm{d}x \leqslant F^*(u^*). \tag{2.11}$$

Conversely, for all $u \in L^\alpha_m$, we obtain from (2.4) the inequalities:

$$u(x)u^*(x) - f(x, u(x)) \leqslant \Phi(x) \tag{2.12}$$

$$\int_\Omega [u(x)u^*(x) - f(x, u(x))]\,\mathrm{d}x \leqslant \int_\Omega \Phi(x)\,\mathrm{d}x. \tag{2.13}$$

Taking the upper bound of the left-hand side in $u \in L^\alpha_m$, we obtain

$$F^*(u^*) \leqslant \int_\Omega \Phi(x)\,\mathrm{d}x. \tag{2.14}$$

Comparing (2.11) and (2.14), we obtain:

$$F^*(u^*) = \int_\Omega \Phi(x)\,\mathrm{d}x. \tag{2.15}$$

But, returning to (2.4), we ascertain that $\Phi(x)$ is none other than $f^*(x; u^*(x))$. Whence the result. ■

Remark 2.1. Proposition 2.1 can easily be extended to the case where the functional F is defined over

$$V = L^{\alpha_1}(\Omega) \times \dots \times L^{\alpha_m}(\Omega), \quad \text{where } 1 \leqslant \alpha_i \leqslant +\infty,$$

this space being in duality with

$$V^* = L^{\alpha'_1}(\Omega) \times \dots \times L^{\alpha'_m}(\Omega), \quad \text{where } 1/\alpha_i + 1/\alpha'_i = 1.$$

Equation (2.3) is valid for all $u^* \in V^*$.

Using a translation, we can also replace the hypothesis $f \geqslant 0$ by:

$$1 \leqslant \alpha_i < \infty, \quad f(x, \xi) \geqslant a(x) - b \sum_{i=1}^{m} |\xi_i|^{\alpha_i},$$

where $a \in L^1(\Omega)$ and $b \geqslant 0$ (if some α_i are $+\infty$, see below).

2.2. Calculation of the Γ-regularization F^{**}

It is sufficient to repeat the process.

Proposition 2.2. *Let us assume that there exist $u_0 \in L^\infty_m$ such that $F(u_0) < \infty$. and $u_0^* \in L^\infty_m$ such that*

$$\int_\Omega f^*(x; u_0^*(x))\, dx < \infty.$$

Then we have:

$$F^{**}(u) = \int_\Omega f^{**}(x; u(x))\, dx. \tag{2.16}$$

Proof. Let us consider the integrand:

$$g(x, \xi) = f(x, \xi - u_0(x)) - f(x, u_0(x)). \tag{2.17}$$

This is a normal integrand. Taking the polars and then the bipolars of both sides:

$$g^*(x; \xi^*) = f^*(x; \xi^*) + u_0(x)\xi^* + f(x, u_0(x)) \tag{2.18}$$

$$g^{**}(x; \xi) = f^{**}(x; \xi - u_0(x)) - f(x, u_0(x)). \tag{2.19}$$

These are normal integrands, by Propositions VIII.1.2 and VIII.1.3, Moreover, we have $g(x, 0) = 0$, and g^* is thus non-negative.

We now set:

$$G(u) = F(u - u_0) - F(u_0). \tag{2.20}$$

Taking the polars of both sides (*cf.* I(4.8) and I(4.9)):

$$G^*(u^*) = F^*(u^*) + \langle u_0, u^* \rangle + F(u_0) \tag{2.21}$$

$$G^{**}(u) = F^{**}(u - u_0) - F(u_0). \tag{2.22}$$

Substituting for $F^*(u^*)$ its value, given by Proposition 2.1:

$$G^*(u^*) = \int_\Omega [f^*(x; u^*(x)) + u_0(x)u^*(x) + f(x, u_0(x))] \, \mathrm{d}x. \tag{2.23}$$

Or again, by (2.18):

$$G^*(u^*) = \int_\Omega g^*(x; u^*(x)) \, \mathrm{d}x. \tag{2.24}$$

By applying Proposition 2.1 to g^*:

$$G^{**}(u) = \int_\Omega g^{**}(x; u(x)) \, \mathrm{d}x. \tag{2.25}$$

Or, by replacing both sides by their value, given respectively by (2.22) and (2.19):

(2.26)
$$F^{**}(u - u_0) - F(u_0) = \int_\Omega [f^{**}(x; u(x) - u_0(x)) - f(x; u_0(x))] \, \mathrm{d}x$$

$$F^{**}(u - u_0) = \int_\Omega f^{**}(x; u(x) - u_0(x)) \, \mathrm{d}x. \tag{2.27}$$

Whence the equation (2.16) follows immediately. ∎

Here is a simple case where there exists $u_0^* \in L_m^\infty$ such that $F^*(u_0^*) < \infty$. Let us assume that there exists a function Φ: $[0, \infty[\to \bar{\mathbf{R}}_+$, which is convex and increasing, such that:

$$\lim_{t \to \infty} \frac{\Phi(t)}{t} = +\infty \tag{2.28}$$

$$\Phi(|\xi|) \leqslant f(x, \xi). \tag{2.29}$$

Making use of (2.28), it is easy to see that the polar $(\Phi \circ |\cdot|)^*$ of the function $\xi \to \Phi(|\xi|)$ is finite everywhere on $(\mathbf{R}^m)^*$. For all $\xi_0^* \in (\mathbf{R}^m)^*$, we deduce from (2.29) that:

$$f^*(x; \xi^*) \leqslant \Phi^*(|\xi_0^*|_*) < \infty \tag{2.30}$$

and it is sufficient to take for u_0^* the constant function equal to ξ_0^*.

Here is another case where the same conclusion holds:

Proposition 2.3. *Assume that* $1 \leqslant \alpha < \infty$ *and that there exists* $u_0 \in L_m^\infty$ *such that* $F(u_0) < \infty$. *We then have*:

$$F^{**}(u) = \int_\Omega f^{**}(x; u(x))\, \mathrm{d}x. \tag{2.31}$$

Proof. We define a function $\tilde{F}$ on L_m^∞ by:

$$\tilde{F}(u) = \int_\Omega f^{**}(x; u(x))\, \mathrm{d}x. \tag{2.32}$$

The integrand f^{**} is normal and non-negative (Prop. VIII.1.3) and the function $\tilde{F}$ is thus strongly l.s.c. on L_m^α (Prop. VIII.1.4). Moreover, since it is convex, it is also weakly l.s.c. on L_m^α, i.e. in the topology $\sigma(L_m^\alpha, L_m^{\alpha'})$.[1] Thus we have:

$$\tilde{F} = (\tilde{F}^*)^*. \tag{2.33}$$

Since $0 \leqslant f^{**} \leqslant f$:

$$0 \leqslant \tilde{F}(u_0) \leqslant F(u_0). \tag{2.34}$$

We can now apply Proposition 2.1 to f^{**}.

$$\begin{aligned} \tilde{F}^*(u^*) &= \int_\Omega f^{***}(x, u(x))\, \mathrm{d}x \\ &= \int_\Omega f^*(x; u(x))\, \mathrm{d}x = F^*(u^*). \end{aligned} \tag{2.35}$$

Hence we have $\tilde{F}^* = F^*$, whence, by applying (2.33), $\tilde{F} = F^{**}$, which is the desired result. ■

[1] It is here that we need the fact that $\alpha \neq +\infty$.

Corollary 2.1. *The following statements are equivalent:*

(a) *F is l.s.c. in the $\sigma(L^\alpha, L^{\alpha'})$-topology,*
(b) *$f(x, .)$ is convex for almost every $x \in \Omega$.*

Proof. In Proposition 2.3 we have just proved that (b) $\Rightarrow$ (a). Conversely, assuming (a), we have by Proposition 1.2:

$$F(u) = \bar{F}(u) = F^{**}(u) \quad \forall u. \tag{2.36}$$

Or again, using Propositions 2.2 or 2.3:

$$\int_\Omega f(x, u(x))\,\mathrm{d}x = \int_\Omega f^{**}(x; u(x))\,\mathrm{d}x \quad \forall u \tag{2.37}$$

which proves, of course, that $f(x, .) = f^{**}(x; .)$ for almost every $x \in \Omega$. ∎

2.3. Recapitulation

We shall gather as a single theorem those results which we shall be using in the following section, namely Propositions 1.3 and 2.2. We recall that, for all $\lambda \in \mathbf{R}$, we have written:

$$S_\lambda = \left[u \in L^\alpha_m \,\middle|\, \int_\Omega f(x, u(x))\,\mathrm{d}x \leqslant \lambda \right] \tag{2.38}$$

and that Φ: $[0, \infty[\to \bar{\mathbf{R}}_+$ is an increasing convex function such that

$$\lim_{t\to\infty} \frac{\Phi(t)}{t} = +\infty.$$

Theorem 2.4. *Let f be a non-negative normal integrand. We assume that there exists $u_0 \in L^\infty_m$ such that $\int_\Omega f(x, \bar{u}_0(x))\mathrm{d}x < +\infty$, and if $\alpha = +\infty$, that $\Phi(|\xi|) \leqslant f(x, \xi)$. Let G be a function in $\Gamma_0(L^\alpha_m)$ whose restriction to S_λ is $\sigma(L^\alpha, L^{\alpha'})$-continuous for every λ. Then we have*

(2.39)

$$\inf_{u\in L^\alpha_m} \left[G(u) + \int_\Omega f(x, u(x))\,\mathrm{d}x \right] = \inf_{u\in L^\alpha_m} \left[G(u) + \int_\Omega f^{**}(x; u(x))\,\mathrm{d}x \right]$$

$$G(u) + \int_\Omega f^{**}(x; u(x))\,\mathrm{d}x = \varliminf_{v\to u} \left[G(u) + \int_\Omega f(x, u(x))\,\mathrm{d}x \right]. \tag{2.40}$$

Remark 2.2. This theorem can obviously be extended to the case considered in Remark 1.1.

3. APPLICATIONS TO THE CALCULUS OF VARIATIONS

3.1. Formulation of the problem

We shall now formulate a typical problem, which covers many of the situations arising in the calculus of variations. As always, we consider an open bounded subset $\Omega \subset \mathbf{R}^n$, and we take $1 < \alpha \leqslant \infty$. Let $\mathscr{G}$ be a continuous mapping from L_m^α to L_ℓ^β, $1 \leqslant \beta \leqslant \infty$, which is *linear* and an (α, β)-compactifier:

(3.1) the restriction of $\mathscr{G}$ to the balls of L_m^α is continuous for the topologies $\sigma(L^\alpha, L^{\alpha'})$ and $\|.\|_\beta$.

From Proposition VIII.2.3, if $1 < \alpha < \infty$, it is sufficient for this that $\mathscr{G}$ be compact. Let g be a Carathéodory function on $\Omega \times \mathbf{R}^\ell$, f a normal integrand on $\Omega \times \mathbf{R}^m$, satisfying respectively:

(3.2) for almost all $x \in \Omega$, $g(x, .)$ is convex.

$(3.3)_1$ If $1 \leqslant \beta < \infty$, there exists $a_1 \in L^1(\Omega)$ and $b_1 \geqslant 0$ such that
$$0 \leqslant g(x, s) \leqslant a_1(x) + b_1 |s|^\beta.$$

$(3.3)_2$ If $\beta = \infty$, for all $k > 0$ there exists $a_1 \in L^1(\Omega)$ such that
$$0 \leqslant g(x, s) \leqslant a_1(x) \quad \text{for} \quad |s| \leqslant k.$$

$(3.4)_1$ If $1 < \alpha < \infty$, there exists $a_2 \in L^1(\Omega)$ and $b_2 \geqslant 0$ such that
$$f(x, \xi) \geqslant a_2(x) + b_2 |\xi|^\alpha.$$

$(3.4)_2$ If $\alpha = \infty$, there exists $a_2 \in L^1(\Omega)$ and a ball B with centre 0 and radius r in $\mathbf{R}^m$ such that:
$$f(x, \xi) \geqslant a_2(x) + \delta(\xi \mid B).$$

(3.5) There exists $p_0 \in L_\infty(\Omega)^m$ such that $\int_\Omega f(x, p_0(x)) < \infty$.

We consider the optimization problem:

$$(\mathscr{P}) \qquad \operatorname*{Inf}_{\substack{u, p \\ p \in L_m^\alpha,\ u = \mathscr{G} p}} \int_\Omega [g(x, u(x)) + f(x, p(x))] \, dx.$$

3.2. The relaxation theorem

As in Chapter VIII, the following will be called the *relaxed problem*:

$$(\mathscr{PR}) \qquad \left| \inf_{\substack{u,p \\ p \in L^\alpha_m,\ u = \mathscr{G}p}} \int_\Omega [g(x, u(x)) + f^{**}(x; p(x))] \,\mathrm{d}x. \right.$$

We now state the fundamental result relating the original problem $(\mathscr{P})$ to its relaxed problem $(\mathscr{PR})$:

Theorem 3.1. *Under the hypotheses* (3.1) *to* (3.5), *the problem* $(\mathscr{PR})$ *has a solution and*

$$\min(\mathscr{PR}) = \inf(\mathscr{P}). \tag{3.6}$$

If $(\bar{u}, \bar{p})$, $\bar{u} = \mathscr{G}\bar{p}$, *is a solution of* $(\mathscr{PR})$, *there exists a minimizing sequence* (u_n, p_n), $u_n = \mathscr{G}p_n$, $n \in \mathbf{N}$, *of* $(\mathscr{P})$ *such that* $u_n \to \bar{u}$ *in* L^β_ℓ *and that* $p_n \to \bar{p}$ *in the topology* $\sigma(L^\alpha, L^{\alpha'})$. *If* (u_n, p_n), $u_n = \mathscr{G}p_n$, $n \in \mathbf{N}$ *is a minimizing sequence of* $(\mathscr{P})$, *there exists a solution* $(\bar{u}, \bar{p})$, $\bar{u} = \mathscr{G}\bar{p}$ *of* $(\mathscr{PR})$ *and a subsequence* $(u_{n_k})_{k \in \mathbf{N}}$ *which converges to* $\bar{u}$ *in* L^β_ℓ, $(p_{n_k})_{k \in \mathbf{N}}$ *which converges to* $\bar{p}$ *for* $\sigma(L^\alpha, L^{\alpha'})$.

For the proof of this theorem, we appeal to the following lemma.

Lemma 3.2. *If* $g \colon \Omega \times \mathbf{R}^l \to \mathbf{R}$ *is a Carathéodory function which satisfies* (3.3), *the mapping*

$$u \mapsto \int_\Omega g(x, u(x)) \,\mathrm{d}x \tag{3.7}$$

from L^β_ℓ *into* $\mathbf{R}$ *is continuous.*

Proof. We already know that this mapping is l.s.c. from applying Proposition VIII.1.4. It is therefore sufficient to show that it is u.s.c. We therefore take any sequence u_n which converges to $\bar{u}$ in L^β; if $\beta = \infty$, we take $k = \sup \|u_n\|_\infty$ in $(3.3)_2$. We define a positive normal integrand h by:

$$\begin{aligned} &(\beta < \infty) \qquad h(x, s) = a_1(x) + b_1 |s|^\beta - g(x, s), \\ &(\beta = \infty) \qquad h(x, s) = a_1(x) - g(x, s). \end{aligned}$$

We can apply Proposition VIII.1.4 to h, obtaining:

$$\varliminf_{n \to \infty} \int_\Omega h(x, u_n(x)) \,\mathrm{d}x \geqslant \int_\Omega h(x, \bar{u}(x)) \,\mathrm{d}x.$$

Furthermore, on substituting its value into h and taking into account the fact that $\int_\Omega |u_n(x)|^\beta \mathrm{d}x \to \int_\Omega |u(x)|^\beta \mathrm{d}x$ (for $\beta < \infty$):

$$\overline{\lim_{n\to\infty}} \int_\Omega g(x, u_n(x))\, \mathrm{d}x \leqslant \int_\Omega g(x, \bar{u}(x))\, \mathrm{d}x. \quad \blacksquare$$

Proof of Theorem 3.1. From Theorem VIII.2.2, the problem $(\mathscr{P}\mathscr{R})$ has a solution. To prove (3.6) we first note that by the preceding lemma the mapping $u \to \int_\Omega g(x, u(x))\,\mathrm{d}x$ is continuous on L^β_ℓ. We define a function G on L^α_m by:

$$\text{(3.8)} \quad \left|\begin{array}{ll} (1 < \alpha < \infty) & G(p) = \displaystyle\int_\Omega g(x, \mathscr{G}p(x))\,\mathrm{d}x \\ (\alpha = \infty) & G(p) = \displaystyle\int_\Omega g(x, \mathscr{G}p(x))\,\mathrm{d}x + \begin{cases} 0 & \text{if } \|p\|_\infty \leqslant r \\ +\infty & \text{otherwise.} \end{cases} \end{array}\right.$$

From (3.2), it is convex. From (3.1), its restriction to the balls of L^α_m is $\sigma(L^\alpha, L^{\alpha'})$-continuous. If $\alpha = \infty$, dom G is contained in a ball, and thus G is $\sigma(L^\infty, L^1)$-continuous and *a fortiori* $G \in \Gamma_0(L^\infty_m)$. If $1 \leqslant \alpha < \infty$, the restriction of G to the balls of L^α_m is $\sigma(L^\alpha, L^{\alpha'})$-continuous and hence continuous in the norm topology. Thus G is convex and continuous on the whole of L^α_m, and so $G \in \Gamma(L^\alpha_m)$.

Replacing $f(x, \xi)$ by $f(x, \xi) - a_2(x)$ if necessary, we can assume that $a_2(x) = 0$. We define a non-negative function F on L^α_m by:

$$\text{(3.9)} \qquad F(p) = \int_\Omega f(x, p(x))\,\mathrm{d}x.$$

For all $\lambda \in \mathbf{R}$, $S_\lambda = \{p \in L^\alpha_m | F(p) \leqslant \lambda\}$ is a bounded set, by virtue of (3.4) and the fact that g is positive. We have already seen that the restriction of G to the S_λ's is $\sigma(L^\alpha, L^{\alpha'})$-continuous. Finally, the existence of a $p_0^* \in L^\infty_m$ such that $\int_\Omega f^*(x; p_0^*(x))\,\mathrm{d}x < +\infty$ results from the inequality (3.4). The hypotheses of Theorem 2.4 thus apply, and we conclude that min $(\mathscr{P}\mathscr{R}) = \inf (\mathscr{P})$.

Let $\bar{u} = \mathscr{G}\bar{p}$ be a solution of $(\mathscr{P}\mathscr{R})$. By Theorem 2.4:

$$\text{(3.10)} \qquad G(\bar{p}) + F^{**}(\bar{p}) = \underline{\lim}_{p\to\bar{p}} \{ G(p) + F(p) \}.$$

Taking λ sufficiently large, we can restrict ourselves to taking the weak lower limit when p varies in S_λ. Now S_λ is bounded in L^α_m and hence is $\sigma(L^\alpha, L^{\alpha'})$-metrizable. There thus exists a sequence $(p_n)_{n\in \mathbf{N}}$ of S_λ, converging to $\bar{p}$ for $\sigma(L^\alpha, L^{\alpha'})$, such that:

$$G(\bar{p}) + F^{**}(\bar{p}) = \varliminf_{p_n \to \bar{p}} \{ G(p_n) + F(p_n) \}. \tag{3.11}$$

As the left-hand side is equal to min $(\mathscr{P}\mathscr{R})$ and hence to inf $(\mathscr{P})$, (3.11) means that the sequence $(p_n)_{n\in\mathbf{N}}$ is minimizing. Finally, from (3.1), $\mathscr{G}p_n$ tends to $\mathscr{G}\bar{p}$ in the L^β_ℓ-norm.

Conversely, let $u_n = \mathscr{G}p_n$, $n \in \mathbf{N}$ be a minimizing sequence of $(\mathscr{P})$. The sequence $(p_n)_{n\in\mathbf{N}}$ is bounded in L^α_m by virtue of the inequality (3.4). We can thus extract a subsequence $p_{n'}$, converging to $\bar{p}$ in $\sigma(L^\alpha, L^{\alpha'})$ and, from (3.1), the sequence $\mathscr{G}p_{n'}$ will converge to $\mathscr{G}\bar{p}$ in the norm. Making use of the fact that F^{**} is l.s.c. or applying Theorem VIII.2.1:

$$G(\bar{p}) + F^{**}(\bar{p}) \leqslant \varliminf_{n' \to \infty} \{ G(p_{n'}) + F(p_{n'}) \}.$$

But the right-hand side is equal to inf $(\mathscr{P})$, since the sequence $(p_{n'})$ is minimizing, and hence equal to min $(\mathscr{P}\mathscr{R})$. Thus $\bar{u} = \mathscr{G}\bar{p}$ is a solution of $(\mathscr{P}\mathscr{R})$.

3.3. A new formulation of the relaxed problem

The relaxed problem $(\mathscr{P}\mathscr{R})$ can be rewritten in an equivalent form. Henceforth we shall denote by E_m, for $m \in \mathbf{N}$, the following simplex of $\mathbf{R}^m$, and by $\tilde{\lambda}$ its elements:

$$E_m = \left[\tilde{\lambda} \in \mathbf{R}^m \,\middle|\, \sum_{i=1}^{m} \lambda_i = 1 \text{ and } \lambda_i \geqslant 0 \quad \forall i \right]. \tag{3.12}$$

Lemma 3.3. *For all $x \in \Omega$ and all $\xi \in \mathbf{R}^m$, we have:*

$$f^{**}(x;\xi) = \min \left[\sum_{i=1}^{m+1} \lambda_i f(x, \xi_i) \,\middle|\, \sum_{i=1}^{m+1} \lambda_i \xi_i = \xi,\ \tilde{\lambda} \in \Lambda_{m+1} \right]. \tag{3.13}$$

Proof. From Carathéodory's Theorem,[1] we have:

$$\text{co epi } f(x,.) = \left[\sum_{i=1}^{m+2} \lambda_i(\xi_i, a_i) \,\middle|\, \tilde{\lambda} \in E_{m+2}, (\xi_i, a_i) \in \text{epi } f(x,.) \right]. \tag{3.14}$$

[1] In a vector space of dimension n, any point on the convex envelope of a set B can be written as a convex combination of at most $n + 1$ points of B; see for example Rockafellar [4].

We shall now show that this is a closed set. Let $(\xi, a) \in \mathbf{R}^{m+1}$ be the limit of a sequence $(\xi^n, a^n) \in \operatorname{co} \operatorname{epi} f(x, .)$, $n \in \mathbf{N}$. We put each (ξ^n, a^n) into the form (3.14):

$$(\xi^n, a^n) = \sum_{i=1}^{m+2} \lambda_i^n(\xi_i^n, a_i^n). \tag{3.15}$$

By extracting subsequences, we may assume that, as $n \to \infty$:

$$\tilde{\lambda}^n \to \bar{\lambda} \in E_{m+2} \tag{3.16}$$

$$\xi_i^n \to \bar{\xi}_i \quad \text{for } i \in I \tag{3.17}$$

$$|\xi_i^n| \to \infty \quad \text{for } i \in J \tag{3.18}$$

with $I \cup J = \{1, \ldots, m+2\}$. Furthermore, by hypothesis:

$$\sum_{i=1}^{m+2} \lambda_i^n \xi_i^n \to \xi \tag{3.19}$$

$$\sum_{i=1}^{m+2} \lambda_i^n f(x, \xi_i^n) \leqslant \sum_{i=1}^{m+2} \lambda_i^n a_i^n \to a. \tag{3.20}$$

Let us now bring in the inequalities (3.4). If $\alpha = \infty$, it follows from (3.18) that $J = \varnothing$. If $1 < \alpha < \infty$, we deduce from (3.20) that $\lambda_i^n |\xi_i^n|^\alpha$ is bounded above by a constant when $n \to \infty$, and thus that:

$$\forall i \in J, \qquad \lambda_i^n\, |\xi_i^n| \to 0. \tag{3.21}$$

Hence, by substituting into (3.19), we have

$$\sum_{i \in I} \lambda_i^n \xi_i^n \to \xi \tag{3.22}$$

$$\sum_{i \in I} \lambda_i^n \to 1. \tag{3.23}$$

Or, taking (3.16) and (3.17) into account

$$\sum_{i \in I} \bar{\lambda}_i \bar{\xi}_i = \xi \ \text{ and } \ \sum_{i \in I} \bar{\lambda}_i = 1. \tag{3.24}$$

To obtain non-negative terms, we subtract $a_2(x)$ from both sides of (3.20). We then have:

$$a - a_2(x) \geqslant \varliminf_{n\to\infty}\left\{\sum_{i=1}^{m+2} \lambda_i^n(f(x, \xi_i^n) - a_2(x))\right\} \tag{3.25}$$

$$a - a_2(x) \geqslant \varliminf_{n\to\infty}\left\{\sum_{i\in I} \lambda_i^n(f(x, \xi_i^n) - a_2(x))\right\}. \tag{3.26}$$

Taking (3.23) into account,

$$a - a_2(x) \geqslant \varliminf_{n\to\infty}\left\{\sum_{i\in I} \lambda_i^n f(x, \xi_i^n)\right\} - a_2(x). \tag{3.27}$$

Hence again:

$$a \geqslant \sum_{i\in I} \lambda_i^n \varliminf_{n\to\infty} f(x, \xi_i^n) \tag{3.28}$$

Or finally, $f(x, .)$ being l.s.c., and by virtue of (3.16),

$$a \geqslant \sum_{i\in I} \bar{\lambda}_i f(x, \bar{\xi}_i). \tag{3.29}$$

Let us set $b = a - \sum_{i\in I} \bar{\lambda}_i f(x, \bar{\xi}_i)$. We take $i_0 \in I$ such that $\bar{\lambda}_{i_0} \neq 0$, which is possible by (3.23).

$$\bar{a}_i = f(x, \bar{\xi}_i) \quad \text{if} \quad i \in I,\ i \neq i_0$$
$$\bar{a}_{i_0} = f(x, \bar{\xi}_{i_0}) + b/\bar{\lambda}_{i_0}.$$

We then have $(\bar{\xi}_i, \bar{a}_i) \in \operatorname{epi} f(x, .)$ for all $i \in I$, $\xi = \sum_{i\in I} \bar{\lambda}_i \bar{\xi}_i$, $a \in \sum_{i\in I} \bar{\lambda}_i \bar{a}_i$, which means that $(\xi, a) \in \operatorname{co}(\operatorname{epi} f(x, .))$. We have thus proved that this is a closed set.

The lemma can easily be deduced from this. Indeed, we know by equation (1.21) that:

$$\begin{aligned} \operatorname{epi} f^{**}(x; .) &= \overline{\operatorname{co}} \operatorname{epi} f(x, .) \\ &= \operatorname{co} \operatorname{epi} f(x, .). \end{aligned} \tag{3.30}$$

In particular, for all $\xi \in \mathbf{R}^m$,

$$(\xi, f^{**}(x; \xi)) \in \operatorname{co} \operatorname{epi} f(x, \xi) \tag{3.31}$$

$$\forall \varepsilon > 0, \qquad (\xi, f^{**}(x, \xi) - \varepsilon) \notin \operatorname{co} \operatorname{epi} f(x, \xi). \tag{3.32}$$

Whence immediately, making use of Carathéodory's Theorem again, we have:

$$(3.33)\qquad f^{**}(x;\xi) = \min\left[\sum_{i=1}^{m+2} \lambda_i f(x,\xi_i) \,\middle|\, \sum_{i=1}^{m+2} \lambda_i \xi_i = \xi,\ \tilde{\lambda} \in E_{m+2}\right].$$

In particular, we can find $\bar{\tilde{\lambda}} \in E_{m+2}$ and $\bar{\xi}_1, \ldots, \bar{\xi}_{m+2}$ such that:

$$(3.34)\qquad f^{**}(x;\xi) = \sum_{i=1}^{m+2} \bar{\lambda}_i f(x;\bar{\xi}_i), \quad \text{with} \quad \sum_{i=1}^{m+2} \bar{\lambda}_i \bar{\xi}_i = \xi.$$

If, in the equation (3.34), all the λ_i were non-null, and all the $(\bar{\xi}_i, f(x, \bar{\xi}_i))$ affinely independent in $\mathbf{R}^{m+1}$, then $(\xi, f^{**}(x;\xi))$ would be an internal point of co epi $f(x, .)$,(1) thus contradicting (3.32). We may therefore assume that $(\xi, f^{**}(x,\xi))$ is a convex combination of at most $(m+1)$ points of co epi $f(x, .)$. Taking $\lambda \in E_{m+1}$ in (3.33) we obtain equation (3.13). ∎

We shall now allow x and ξ to vary in (3.13).

Proposition 3.1. *For any measurable mapping p of Ω in $\mathbf{R}^m$, there exists a measurable mapping $\tilde{\ell}$ of Ω into E_{m+1} and $(m+1)$ measurable mappings q_i of Ω in $\mathbf{R}^m$, such that for almost all $x \in \Omega$:*

$$(3.35)\qquad f^{**}(x;p(x)) = \sum_{i=1}^{m+1} \ell_i(x) f(x, q_i(x))$$

$$(3.36)\qquad \sum_{i=1}^{m+1} \ell_i(x) q_i(x) = p(x).$$

Proof. The integrands f and f^{**} are normal and the mapping p is measurable. By modifying them on a set $N \subset \Omega$ of null measure, we may therefore assume f and f^{**} to be Borel on $\Omega \times \mathbf{R}^m$ and p to be Borel on Ω.

Let us consider in $\Omega \times E_{m+1} \times (\mathbf{R}^m)^{m+1}$ the subset:

$$C = \left[(x, \tilde{\lambda}, \xi_i, \ldots, \xi_{m+1}) \,\middle|\, \sum_{i=1}^{m+1} \lambda_i \xi_i = p(x),\ \sum_{i=1}^{m+1} \lambda_i f(x,\xi_i) = f^{**}(x, p(x))\right].$$

It is clearly Borel and its section C_x is closed for all $x \in \Omega$. The indicator function of C is thus a normal integrand. By a homeomorphism of $\mathbf{R}^m$

(1) Indeed, this point $(\xi, f^{**}(x,\xi))$ will then be interior to the $(m+2)$-simplex generated by the $(\bar{\xi}_i, f^{**}(x, \bar{\xi}_i))$, $1 \leqslant i \leqslant m+2$.

on the interior of its unit ball B_m, we can arrive at the case where $C \subset \Omega \times E_{m+1} \times (B_m)^{m+1}$. We can then apply Corollary VIII.1.1: there exists a measurable mapping u from Ω into $E_{m+1} \times (\mathbf{R}^m)^{m+1}$ such that $u(x) \in C_x$ for all x in Ω, i.e., $(x, \tilde{u}(x)) \in C$ for all $x \in \Omega$. The various components of $\tilde{u}$ give us $\tilde{\ell}$ and the q_i, $1 \leqslant i \leqslant m+1$. ■

We are thus led to reformulate the relaxed problem $(\mathscr{PR}')$

$$(3.37)\quad \left|\begin{array}{l} \operatorname{Inf} \displaystyle\int_\Omega \left[g(x, u(x)) + \sum_{i=1}^{m+1} \ell_i(x) f(x, q_i(x))\right] \mathrm{d}x \\ \ell_i : \Omega \to \mathbf{R}, \quad q_i : \Omega \to \mathbf{R}^m \text{ measurable, } 1 \leqslant i \leqslant m+1 \\ \displaystyle\sum_{i=1}^{m+1} \ell_i(x) = 1 \text{ and } \ell_i(x) \geqslant 0 \text{ a.e., } 1 \leqslant i \leqslant m+1 \\ p = \displaystyle\sum_{i=1}^{m+1} \ell_i q_i \in L_m^\alpha \text{ and } u = \mathscr{G}p. \end{array}\right.$$

Proposition 3.2. *The problems* $(\mathscr{PR})$ *and* $(\mathscr{PR}')$ *have the same value and* min $(\mathscr{PR}) =$ min $(\mathscr{PR}')$. *Furthermore, the mapping*:

$$(3.38)\quad (\bar{\ell}_1, \ldots, \bar{\ell}_{m+1}, \bar{q}_1, \ldots, \bar{q}_{m+1}) \to \bar{p} = \sum_{i=1}^{m+1} \bar{\ell}_i \bar{q}_i$$

sends the set of solutions of the problem $(\mathscr{PR}')$ *onto the set of solutions of the problem* $(\mathscr{PR})$.

Proof. If $\tilde{\ell}: \Omega \to E_{m+1}$ and $q_i: \Omega \to \mathbf{R}^m$, $1 \leqslant i \leqslant m+1$ are measurable mappings and if $p = \sum_{i=1}^{m+1} \ell_i q_i$, we deduce from Lemma 3.3 that, for all $x \in \Omega$

$$(3.39)\quad \sum_{i=1}^{m+1} \ell_i(x) f(x, q_i(x) \geqslant f^{**}(x; p(x)).$$

We thus have $\inf(\mathscr{PR}') \geqslant \min(\mathscr{PR})$. But, if $\bar{p}$ is a solution of $(\mathscr{PR})$, from Proposition 3.1, there exist measurable mappings $\bar{\ell}: \Omega \to E_{m+1}$ and $q_i: \Omega \to \mathbf{R}^m$, $1 \leqslant i \leqslant m+1$, such that, for almost all $x \in \Omega$, we have:

$$(3.40)\quad \sum_{i=1}^{m+1} \bar{\ell}_i(x) f(x, \bar{q}_i(x)) = f^{**}(x; \bar{p}(x))$$

$$(3.41)\quad \sum_{i=1}^{m+1} \bar{\ell}_i(x) \bar{q}_i(x) = \bar{p}(x).$$

Setting $\bar{u} = \mathscr{G}\bar{p}$, we thus have:

$$\begin{aligned}\min(\mathscr{P}\mathscr{R}) &= \int_\Omega [g(x, \bar{u}(x)) + f^{**}(x; \bar{p}(x))]\,\mathrm{d}x \\ &= \int_\Omega [g(x, u(x)) + \sum_{i=1}^{m+1} \bar{\ell}_i(x) f(x, \bar{q}_i(x))]\,\mathrm{d}x. \\ &\geqslant \inf(\mathscr{P}\mathscr{R}'). \end{aligned} \tag{3.42}$$

Comparing (3.39) and (3.42), we deduce that $\inf(\mathscr{P}\mathscr{R}') = \min(\mathscr{P}\mathscr{R})$ and that $(\ell_1, \ldots, \ell_{mn}, \bar{q}_1, \ldots, \bar{q}_{m+1})$ is thus a solution of $(\mathscr{P}\mathscr{R}')$. Therefore $\min(\mathscr{P}\mathscr{R}) = \min(\mathscr{P}\mathscr{R}')$, and the mapping (3.38) is surjective. Conversely, if $(\ell_1, \ldots, \ell_{mn}, \bar{q}_1, \ldots, \bar{q}_{m+1})$ is a solution of $(\mathscr{P}\mathscr{R}')$, if $\bar{p} = \sum \ell_i \bar{q}_i$ and $\bar{u} = \mathscr{G}\bar{p}$, we deduce from (3.39) that

$$\begin{aligned}\min(\mathscr{P}\mathscr{R}') &= \int_\Omega \left[g(x, \bar{u}(x)) + \sum_{i=1}^{m+1} \bar{\ell}_i(x) f(x, \bar{q}_i(x)) \right] \mathrm{d}x \\ &\geqslant \int_\Omega [g(x, \bar{u}(x)) + f^{**}(x, \bar{p}(x))]\,\mathrm{d}x.\end{aligned}$$

Since $\min(\mathscr{P}\mathscr{R}') = \min(\mathscr{P}\mathscr{R})$, p is a solution of $(\mathscr{P}\mathscr{R})$, and (3.38) thus maps the solutions of $(\mathscr{P}\mathscr{R}')$ into solutions of $(\mathscr{P}\mathscr{R})$. ■

Problem $(\mathscr{P}\mathscr{R}')$ has an interesting probabilistic interpretation. For all $x \in \Omega$, instead of choosing a deterministic control $p(x) \in \mathbf{R}^m$, we simply choose a set of $m + 1$ possible controls $q_i(x)$, and a probability distribution $\ell_i(x)$ among them. We can think of it as a unique control switching very quickly, and assuming the value $q_i(x)$ with frequency $\ell_i(x)$ around x. The system, by linearity, will react as to the control $\sum_{i=1}^{m+1} \ell_i(x) q_i(x)$, and will assume the corresponding state. It only remains to take the mathematical expectation of the criterion, i.e., to integrate it.

4. RELAXATION. THE GENERAL CASE

We shall now deal with problems in the calculus of variations where the criterion cannot be decomposed into a part which is a convex function of $u(x)$ and a part which depends only on $p(x)$, as in the previous section. The proof will rely on equi-integrability properties and a direct application of Proposition 1.2.

4.1. Formulation of the problem

Henceforth, Φ: $[0, +\infty[\to \mathbf{R}_+$ will be a non-negative, increasing, convex and l.s.c. function such that

$$\lim_{t \to \infty} \frac{\Phi(t)}{t} = +\infty \tag{4.1}$$

and g will be a normal integrand of $\Omega \times \mathbf{R}^m$ which satisfies:

$$\Phi(|\xi|) \leqslant g(x, \xi). \tag{4.2}$$

β will be a given exponent with $1 \leqslant \beta \leqslant \infty$, and f will be a normal integrand of $\Omega \times (\mathbf{R}^l \times \mathbf{R}^m)$ such that

$(4.3)_1$ If $1 \leqslant \beta < \infty$, there exist a_1 and $a_2 \in L^1(\Omega)$, $b \geqslant 0$ and $c \geqslant 1$ such that:
$$g(x, \xi) + a_2(x) \leqslant f(x, s, \xi) \leqslant cg(x, \xi) + b\,|s|^\beta + a_1(x).$$

$(4.3)_2$ If $\beta = \infty$, there exists $a_2 \in L^1(\Omega)$ and, for all $k > 0$, there exist $c \geqslant 1$ and $a_1 \in L^1(\Omega)$ such that
$$g(x, \xi) + a_2(x) \leqslant f(x, s, \xi) \leqslant cg(x, \xi) + a_1(x) \quad \text{for} \quad |s| \leqslant k.$$

(4.4) For almost all $x \in \Omega$, the restriction of $f(x, ., .)$ to $\mathbf{R}^l \times \operatorname{dom} g(x, .)$ is continuous.

We note that if f does not depend on s and satisfies $f(x, \xi) \geqslant \Phi(|\xi|)$, hypotheses (4.2) and (4.3) are satisfied with $g(x, \xi) = f(x, \xi)$, $b = a_1 = a_2 = 0$. In the general case, $f(x, s, \xi)$ is finite if and only if $g(x, \xi)$ is finite. If we therefore set $C = \{(x, \xi) \mid g(x, \xi) < +\infty\}$, hypothesis (4.4) means that g is the sum of a Carathéodory function and the indicator function of the subset $C \times \mathbf{R}^l \subset \Omega \times \mathbf{R}^l \times \mathbf{R}^m$. *Note also that, in contradiction to Section* 3, *we make no assumption concerning the convexity of* $f(x, ., \xi)$ *in* s.

Finally, we take a mapping $\mathcal{G}$ from L^Φ_m into L^β_m, which is a (Φ, β)-compactifier:

(4.5) if $(p_n)_{n \in \mathbf{N}}$ converges weakly to $\bar{p}$ in L^1_m and if $\sup_{n \in \mathbf{N}} \int \Phi \circ |p_n| < \infty$, then $(\mathcal{G}p_n)_{n \in \mathbf{N}}$ converges to $\mathcal{G}\bar{p}$ in L^β.

Hypothesis (4.5) is satisfied if $\mathcal{G}$ is a continuous linear mapping from L^1_m into L^β_l $(1 < \beta \leqslant \infty)$, or if $\mathcal{G}$ is a compact continuous linear mapping from L^α_m $(1 < \alpha < \infty)$ into L^β_l $(1 \leqslant \beta \leqslant \infty)$ (Propositions VIII.2.3 and VIII.2.4).

We consider the optimization problem:

$$(\mathscr{P}) \qquad \left| \begin{array}{l} \operatorname*{Inf}_{u,p} \displaystyle\int_\Omega f(x, u(x), p(x))\,\mathrm{d}x \\ p \in L^{\Phi}_m, \quad u = \mathscr{G}p. \end{array} \right.$$

4.2. The relaxation theorem

We have associated with the problem $(\mathscr{P})$ the relaxed problem $(\mathscr{P}\mathscr{R})$:

$$(\mathscr{P}\mathscr{R}) \qquad \left| \begin{array}{l} \operatorname*{Inf}_{u,p} \displaystyle\int_\Omega f^{**}(x, u(x); p(x))\,\mathrm{d}x \\ u = \mathscr{G}p, \quad p \in L^{\Phi}_m. \end{array} \right.$$

From the results of Chapter VIII, f^{**} is a normal integrand (Prop. 2.1) and the problem $(\mathscr{P}\mathscr{R})$ has a solution (Prop. 2.5). The following theorem is the analogue of Theorem 3.1:

Theorem 4.1. *Under the hypotheses* (4.1) *to* (4.5), *the problem* $(\mathscr{P}\mathscr{R})$ *has a solution and*

$$\operatorname{Min}(\mathscr{P}\mathscr{R}) = \operatorname{Inf}(\mathscr{P}). \tag{4.6}$$

If $(\bar{u}, \bar{p})$, $\bar{u} = \mathscr{G}\bar{p}$, *is a solution of* $(\mathscr{P}\mathscr{R})$, *there is a minimizing sequence* (u_n, p_n), $u_n = \mathscr{G}p_n$ *of* $(\mathscr{P})$ *such that* $u_n \to \bar{u}$ *in* L^{β}_l *and that* $p_n \to \bar{p}$ *in the* $\sigma(L^1 L^\infty)$ *topology. If* (u_n, p_n), $u_n = \mathscr{G}p_n$ *is a minimizing sequence of* $(\mathscr{P})$, *there exists a solution* $(\bar{u}, \bar{p})$, $\bar{u} = \mathscr{G}\bar{p}$ *of* $(\mathscr{P}\mathscr{R})$ *and a subsequence* $(u_{n'}, p_{n'})$ *such that* $u_{n'} \to \bar{u}$ *in* L^{β}_l *and that* $p_{n'} \to \bar{p}$ *in the* $\sigma(L^1, L^\infty)$ *topology.*

Theorem 4.1 can be extended to the case where Ω is no longer an open subset of $\mathbf{R}^n$ as will be pointed out in Remark 4.3. We shall prove Theorem 4.1 in Section 4.4, after giving in Section 4.3 some preliminary results.

Remark 4.1. When the integrand $f(x, s, \xi)$ is of the form $g(x, s) + f(x, \xi)$ where $g(x, .)$ is convex for all $x \in \Omega$, we find that, as in Section 3, $\operatorname{Inf}(\mathscr{P}) = \operatorname{Min}(\mathscr{P}\mathscr{R})$. But Theorem 4.1 does not repeat all the results of Section 3, since we had established there a considerably more definite result:

$$G + F^{**} = \overline{(G + F)}.$$

Remark 4.2. From Theorem 4.1, the two following statements are equivalent to each other:

(i) the sequence (u_n, p_n), $u_n = \mathscr{G}p_n$, is minimizing for $(\mathscr{P})$,

(ii) the sequence (u_n, p_n), $u_n = \mathscr{G}p_n$, is minimizing for $(\mathscr{P}\mathscr{R})$, and, when $n \to \infty$:

$$\int_\Omega [f(x, u_n(x), p_n(x)) - f^{**}(x, u_n(x); p_n(x))] \, dx \to 0 .$$

We note that the integrand $f - f^{**}$ is non-negative, and therefore, under conditions (i) or (ii), there exists a subsequence $(u_{n'}, p_{n'})$ of the original sequence such that, when $n' \to \infty$:

$$f(x, u_n(x), p_n(x)) - f^{**}(x, u_n(x); p_n(x)) \to 0 \qquad \text{a.e.}$$

Remark 4.3. We obtain analogous results if we consider the operators $\mathscr{G}$ of $L^\Phi_m(\tilde{\Omega}, d\mu)$ into $L^\beta_l(\tilde{\Omega})$, where $\tilde{\Omega}$ is a set provided with a finite positive measure $d\mu$ ($\tilde{\Omega} = \Omega$, $d\mu = dx$, the Lesbesgue measure in the foregoing).

This allows us in particular to study problems with boundary controls, to be encountered in Section 5. ■

4.3. Some preliminary lemmas

Lemma 4.1. *Let $(\bar{u}, \bar{p})$, $\bar{u} = \mathscr{G}\bar{p}$, a solution of $(\mathscr{P}\mathscr{R})$. There exists a measurable mapping $\bar{\ell}: \Omega \to E_{m+1}$ and $(m + 1)$ measurable mappings $q_i: \Omega \to \mathbf{R}^m$ such that:*

$$\sum_{i=1}^{m+1} \bar{\ell}_i(x)\bar{q}_i(x) = \bar{p}(x) \qquad \text{a.e.} \tag{4.7}$$

$$\sum_{i=1}^{m+1} \bar{\ell}_i(x) f(x, \bar{u}(x), \bar{q}_i(x)) = f^{**}(x, \bar{u}(x); \bar{p}(x)) \qquad \text{a.e.} \tag{4.8}$$

This lemma can be proved in the same way as Proposition 3.4, by making use of the inequalities (4.3). The same conclusion holds whenever $(\bar{u}, \bar{p})$ is a point such that $\bar{u} = \mathscr{G}\bar{p}$ and such that the right-hand side in (4.8) is finite almost everywhere.

Lemma 4.2. *There exists a function ψ: $[0, +\infty[\to \overline{\mathbf{R}}_+$, which is increasing and convex such that* $\lim_{t\to\infty}(\psi(t)/t) = +\infty$ *and that:*

$$\int_\Omega \sum_{i=1}^{m+1} \bar{\ell}_i(x)\psi \circ g(x, \bar{q}_i(x)) \, dx < +\infty. \tag{4.9}$$

Proof. From (4.3):

$$\int_\Omega \sum_{i=1}^{m+1} \bar{\ell}_i(x) g(x, \bar{q}_i(x))\,\mathrm{d}x \leqslant \int_\Omega \sum_{i=1}^{m+1} \bar{\ell}_i(x) f(x, \bar{u}(x), \bar{q}_i(x))\,\mathrm{d}x - \int_\Omega a_2(x)\,\mathrm{d}x. \tag{4.10}$$

Or by substituting (4.8) into it:

(4.11)

$$\int_\Omega \sum_{i=1}^{m+1} \bar{\ell}_i(x) g(x, \bar{q}_i(x))\,\mathrm{d}x \leqslant \int_\Omega f^{**}(x, \bar{u}(x); \bar{p}(x))\,\mathrm{d}x - \int_\Omega a_2(x)\,\mathrm{d}x$$

$$\int_\Omega \sum_{i=1}^{m+1} \bar{\ell}_i(x) g(x, \bar{q}_i(x))\,\mathrm{d}x < +\infty. \tag{4.12}$$

For $n \in \mathbf{N}$, we set:

$$C_n^i = \{ x \in \Omega \mid n \leqslant g(x, \bar{q}_i(x)) < n + 1 \} \tag{4.13}$$

$$\alpha_n = \sum_{i=1}^{m+1} \int_{C_n^i} \bar{\ell}_i(x) g(x, \bar{q}_i(x))\,\mathrm{d}x. \tag{4.14}$$

From (4.12), we have $\sum_{n \in \mathbf{N}} \alpha_n < +\infty$. We know[1] that in that case we can find an increasing sequence k_n of integers tending to $+\infty$ and such that $\sum_{n \in \mathbf{N}} k_n \alpha_n < +\infty$. We define $\psi : [0, +\infty[\to \mathbf{R}_+$ by induction:

$$\psi(t) = k_n (t - n) + \psi(n) \quad \text{if } n \leqslant t < n + 1 \text{ and } \psi(0) = 1. \tag{4.15}$$

We check that:

$$\lim_{t \to \infty} \frac{\psi(t)}{t} = \lim_{n \to \infty} k_n = +\infty \tag{4.16}$$

$$\begin{aligned} \int_\Omega \sum_{i=1}^{m+1} \bar{\ell}_i(x) \psi \circ g(x, \bar{q}_i(x))\,\mathrm{d}x &= \sum_{n \in \mathbf{N}} \sum_{i=1}^{m+1} \int_{C_n^i} \bar{\ell}_i(x) \psi \circ g(x, \bar{q}_i(x))\,\mathrm{d}x \\ &\leqslant \sum_{n \in \mathbf{N}} \sum_{i=1}^{m+1} \int_{C_n^i} k_n \bar{\ell}_i(x) g(x, \bar{q}_i(x))\,\mathrm{d}x \\ &= \sum_{n \in \mathbf{N}} k_n \alpha_n < +\infty. \quad \blacksquare \end{aligned} \tag{4.17}$$

[1] See the appendix at the end of this chapter.

Lemma 4.3. *There exists a sequence $(p_n)_{n\in\mathbf{N}}$ of L^1_m, converging to $\bar{p}$ in the $\sigma(L^1, L^\infty)$ topology, and such that:*

$$(4.18)\qquad \int_\Omega f(x, \bar{u}(x), p_n(x))\,\mathrm{d}x \to \int_\Omega f^{**}(x, \bar{u}(x); \bar{p}(x))\,\mathrm{d}x$$

$$(4.19)\qquad \int_\Omega \psi\circ g(x, p_n(x)) \to \int_\Omega \sum_{i=1}^{m+1} \bar{\ell}_i(x)\psi\circ g(x, \bar{q}_i(x))\,\mathrm{d}x.$$

Proof. For all $k\in\mathbf{N}$ and all $I\subset\{1,\ldots,m+1\}$, we introduce:

$$(4.20)\quad \Omega_{k,I} = \{\, x\in\Omega \mid \psi\circ g(x, \bar{q}_i(x)) \leqslant k \ \ \forall i\in I,\quad \bar{\ell}_i(x) = 0\ \ \forall i\notin I \,\}.$$

From (4.9), Ω is the countable union of measurable sets $\Omega_{k,I}$. Let $(\Omega_j)_{j\in\mathbf{N}}$ be a denumeration of the $\Omega_{k,I}$. If we write $\Omega'_j = \Omega_j - \bigcup_{i<j}\Omega_i$ then the Ω'_j, $j\in\mathbf{N}$, form a partition of Ω into measurable sets. Clearly it is sufficient to prove Lemma 4.4 for the Ω'_j; to obtain a p which is an ε-approximation of $\bar{p}$ on Ω, we construct a family p_j of $\varepsilon 2^{-1-j}$-approximations on Ω'_j, $j\in\mathbf{N}$, and we piece them together by defining $p(x) = p_j(x)$ on Ω'_j.

We can thus arrive at the case where Ω is measurable and where the functions $x\to\psi\circ g(x, \bar{q}_i(x))$ belong to $L^\infty(\Omega)$. From (4.2), the $\bar{q}_i$ belong to $L^\infty(\Omega)$ and from (4.3) the functions $x\to f(x, \bar{u}(x), \bar{q}_i(x))$ belong to $L^1(\Omega)$. But we know that the tiered functions[1] are dense in $L^\infty(\Omega)$. For all $\varepsilon>0$ and every family $h_k\in L^\infty(\Omega)$, $1\leqslant k\leqslant N$, there thus exists a measurable mapping $\tilde{\ell}$ from Ω into E_{m+1} which only takes a finite number of values, such that:

$$(4.21)\quad \left|\int_\Omega \sum_{i=1}^{m+1} (\ell_i(x) - \bar{\ell}_i(x))\bar{q}_i(x)h_k(x)\,\mathrm{d}x\right| \leqslant \varepsilon/2, \qquad 1\leqslant k\leqslant N$$

$$(4.22)\qquad \left|\int_\Omega \sum_{i=1}^{m+1} (\ell_i(x) - \bar{\ell}_i(x)) f(x, \bar{u}(x), \bar{q}_i(x))\,\mathrm{d}x\right| \leqslant \varepsilon/2$$

$$(4.23)\qquad \left|\int_\Omega \sum_{i=1}^{m+1} (\ell_i(x) - \bar{\ell}_i(x))\psi\circ g(x, \bar{q}_i(x))\,\mathrm{d}x\right| \leqslant \varepsilon/2.$$

We have a partition of Ω into measurable disjoint subsets on which ℓ is constant. On each of them, we can apply Proposition 1.1. By piecing them

[1] I.e. those measurable functions which only take a finite number of values.

together we obtain a partition of Ω into $(m+1)$ measurable subsets B_i such that:

(4.24)
$$\left|\int_\Omega \left(\sum_{i=1}^{m+1} \ell_i(x)\bar{q}_i(x) - \sum_{i=1}^{m+1} \mathbf{1}_{B_i}\bar{q}_i(x)\right)h_k(x)\,\mathrm{d}x\right| \leqslant \varepsilon/2, \quad 1 \leqslant k \leqslant N.$$

$$\left|\int_\Omega \left(\sum_{i=1}^{m+1} \ell_i(x) - \sum_{i=1}^{m+1} \mathbf{1}_{B_i}\right) f(x, \bar{u}(x), \bar{q}_i(x))\,\mathrm{d}x\right| \leqslant \varepsilon/2 \tag{4.25}$$

$$\left|\int_\Omega \left(\sum_{i=1}^{m+1} \ell_i(x) - \sum_{i=1}^{m+1} \mathbf{1}_{B_i}\right)\psi \circ g(x, q_i(x))\,\mathrm{d}x\right| \leqslant \varepsilon/2. \tag{4.26}$$

We define $p_\varepsilon \in L_m^\Phi$ by:

$$p_\varepsilon(x) = q_i(x) \quad \text{on} \quad B_i. \tag{4.27}$$

By adding (4.21) to (4.24), (4.22) to (4.25) and (4.23) to (4.26), we obtain:

$$\left|\int_\Omega \left(p_\varepsilon(x) - \sum_{i=1}^{m+1} \bar{\ell}_i(x)\bar{q}_i(x)\right)h_k(x)\,\mathrm{d}x\right| \leqslant \varepsilon, \qquad 1 \leqslant k \leqslant N, \tag{4.28}$$

$$\left|\int_\Omega f(x, \bar{u}(x), p_\varepsilon(x))\,\mathrm{d}x - \int_\Omega \sum_{i=1}^{m+1} \bar{\ell}_i(x) f(x, \bar{u}(x), \bar{q}_i(x))\,\mathrm{d}x\right| \leqslant \varepsilon, \tag{4.29}$$

$$\left|\int_\Omega \psi \circ g(x, p_\varepsilon(x))\,\mathrm{d}x - \int_\Omega \sum_{i=1}^{m+1} \bar{\ell}_i(x)\psi \circ g(x, \bar{q}_i(x))\,\mathrm{d}x\right| \leqslant \varepsilon. \tag{4.30}$$

Let us choose in $L^\infty(\Omega)$ a sequence $(h_k)_{k\in\mathbf{N}}$ which is dense in the $\sigma(L^\infty, L^1)$ topology. To obtain the required sequence p_n, it is sufficient to take $\varepsilon = 1/n$ in equations (4.28), (4.29) and (4.30), and the first n terms of the sequence $(h_k)_{k\in\mathbf{N}}$ in (4.28). We note a fact which will be useful to us later; namely that in any point $x \in \Omega$ the sequence $p_n(x)$ only takes a finite number of values (the $q_i(x)$, $1 \leqslant i \leqslant m+1$). ∎

4.4. Proof of the relaxation theorem

As before, let $\bar{u} = \mathscr{G}\bar{p}$ be a solution of $(\mathscr{P}\mathscr{R})$, and $(p_n)_{n\in\mathbf{N}}$ a sequence converging to $\bar{p}$ for $\sigma(L^1, L^\infty)$ and such that:

$$\int_\Omega f(x, \bar{u}(x), p_n(x))\,\mathrm{d}x \to \int_\Omega f^{**}(x, \bar{u}(x); \bar{p}(x))\,\mathrm{d}x \tag{4.31}$$

$$\sup_{n\in\mathbb{N}} \psi \circ g(x, p_n(x))\,\mathrm{d}x < +\infty \tag{4.32}$$

$$\forall x \in \Omega, \quad p_n(x) \text{ takes at most } (m+1) \text{ distinct values.}^{(1)} \tag{4.33}$$

The existence of such a sequence has just been established, as (4.32) is a consequence of (4.19). We set $u_n = \mathscr{G}p_n$. From estimate (4.32) and hypotheses (4.2) and (4.5), u_n converges to $\bar{u}$ in L^β. Extracting a subsequence if necessary, we may assume that u_n converges to $\bar{u}$ almost everywhere. Let us set:

$$h_n(x) = f(x, u_n(x), p_n(x)) - f(x, \bar{u}(x), p_n(x)). \tag{4.34}$$

For almost all $x \in \Omega$, $u_n(x)$ tends to $\bar{u}(x)$, and $p_n(x)$ only takes a finite number of values in dom $g(x, .)$. From hypothesis (4.4) we at once deduce that $h_n(x) \to 0$. Hence the sequence h_n converges to zero almost everywhere.

From the inequalities (4.3), we have, if $\beta < \infty$:

$$|h_n(x)| \leqslant |a_1(x) - a_2(x)| + b(|\bar{u}(x)|^\beta + |u_n(x)|^\beta) + (c-1)g(x, p_n(x)) \tag{4.35}$$

and if $\beta = \infty$, taking $k = \sup_n \|u_n\|_\infty$:

$$|h_n(x)| \leqslant |h_n(x)| \leqslant |a_1(x) - a_2(x)| + (c-1)g(x, p_n(x)) \tag{4.36}$$

From Theorem VIII.1.3, the inequality (4.32) means that the family of mappings $x \to g(x, p_n(x))$, for $n \in \mathbb{N}$, is equi-integrable. If $\beta < \infty$, $|u_n|^\beta$ converges to $|\bar{u}|^\beta$ in L^1, and the family of $|u_n|^\beta$, $n \in \mathbb{N}$, is equi-integrable. The right-hand sides of (4.35) and (4.36) are thus equi-integrable, and hence the family of h_n, $n \in \mathbb{N}$ is equi-integrable. Still using Theorem VIII.1.3, we can extract a subsequence $(h_{nk})_{k\in\mathbb{N}}$ converging weakly to $\bar{h}$ in L^1. But we already know that the sequence $(h_n)_{n\in\mathbb{N}}$ converges to zero almost everywhere. Hence of necessity $\bar{h} = 0$, and the sequence $(h_n)_{n\in\mathbb{N}}$ as a whole weakly converges to zero in L^1. In particular:

$$\int_\Omega h_n(x)\,\mathrm{d}x \to 0 \tag{4.37}$$

$$\int_\Omega f(x, u_n(x), p_n(x))\,\mathrm{d}x - \int_\Omega f(x, \bar{u}(x), p_n(x))\,\mathrm{d}x \to 0 \tag{4.38}$$

Bringing together (4.31) and (4.47), we have the desired result:

$$\int_\Omega f(x, u_n(x), p_n(x))\,\mathrm{d}x \to \int_\Omega f^{**}(x, \bar{u}(x), \bar{p}(x))\,\mathrm{d}x. \tag{4.39}$$

(1) More exactly, $p_n(x) \in \{q_1(x), \ldots, q_{m+1}(x)\}$.

Whence it immediately follows that inf $(\mathscr{P}) \leqslant$ min $(\mathscr{P}\mathscr{R})$. Since the inverse inequality is trivial, we have the required equality, and hence the theorem. ■

4.5. A new formulation of the relaxed problem

As in the preceding paragraph, we can clearly give the relaxed problem the following equivalent form:

$$\left|\begin{array}{l} \operatorname{Inf} \displaystyle\sum_{i=1}^{m+1} \ell_i(x) f(x, u(x), q_i(x))\, \mathrm{d}x \\ \ell_i : \Omega \to \mathbf{R},\ q_i : \Omega \to \mathbf{R}^m \text{ measurable}, \quad 1 \leqslant i \leqslant m+1 \\ \displaystyle\sum_{i=1}^{m+1} \ell_i(x) = 1 \text{ and } \ell_i(x) \geqslant 0 \qquad \text{a.e.,} \qquad 1 \leqslant i \leqslant m+1 \\ p = \displaystyle\sum_{i=1}^{m+1} \ell_i q_i \in L^{\Phi}_m \text{ and } u = \mathscr{G}p. \end{array}\right.$$

Problems $(\mathscr{P}\mathscr{R})$ and $(\mathscr{P}\mathscr{R}')$ have the same value: min $(\mathscr{P}\mathscr{R})$ = min $(\mathscr{P}\mathscr{R}')$. Furthermore, the mapping

$$(\bar{\ell}_1, \ldots, \bar{\ell}_{m+1}, \bar{q}_1, \ldots, \bar{q}_{m+1}) \to \bar{p} = \sum_{i=1}^{m+1} \bar{\ell}_i \bar{q}_i$$

sends the set of solutions of problem $(\mathscr{P}\mathscr{R}')$ onto the set of solutions of problem $(\mathscr{P}\mathscr{R})$.

5. EXAMPLES

Example 1

Consider the problem:

$$(\mathscr{P}) \qquad \operatorname*{Inf}_{u \in H^1_0(\Omega) \cap H^2(\Omega)} \int_\Omega (1 - |\Delta u|)^2 \, \mathrm{d}x$$

where Ω is a bounded open subset of $\mathbf{R}^n$, and Δ the Laplace operator. It is associated with a Green operator $\mathscr{G}$ which has been shown in Chapter VIII, Section 3, to be a (2, 2)-compactifier, and the function $f(x, s, \xi) = (1 - |\xi|)^2$ clearly satisfies assumptions (4.1)–(4.4) with $\beta = 2$. Denoting by $(|\xi| - 1)_+$ the non-negative part of $(|\xi| - 1)$:

$$(|\xi| - 1)_+ = \begin{cases} 0 & \text{if} \quad (|\xi| - 1) \leqslant 0 \\ (|\xi| - 1) & \text{if} \quad (|\xi| - 1) > 0 \end{cases}$$

we can state the relaxed problem:

$$(\mathscr{P}\mathscr{R}) \qquad \operatorname*{Min}_{u \in H_0^1(\Omega)} \int_\Omega (|\Delta u| - 1)_+^2 \, dx,$$

and apply Theorem 4.1. Note first that $\bar{u} = 0$ is the only optimal solution of $(\mathscr{P}\mathscr{R})$. Hence:

$$\text{(5.1)} \qquad \operatorname{Inf}(\mathscr{P}) = \operatorname{Min}(\mathscr{P}\mathscr{R}) = 0$$

$$\text{(5.2)} \qquad \begin{cases} \text{if } u_n \text{ is a minimizing sequence of } (\mathscr{P}), \text{ then} \\ u_n \to 0 \text{ strongly and } \Delta u_n \to 0 \text{ weakly in } L^2(\Omega). \end{cases}$$

We deduce, first, that $(\mathscr{P})$ has no solution; indeed, $\bar{u} = 0$ is the only candidate, and substituting it into $(\mathscr{P})$, we get the value meas (Ω), which is positive, and hence greater than inf $(\mathscr{P})$ by (5.1). Moreover, if u_n is a minimizing sequence of $(\mathscr{P})$, then:

$$\int_\Omega (1 - |\Delta u_n|)^2 \, dx \to \inf(\mathscr{P}) = 0.$$

There exists a subsequence $u_{n'}$ such that:

$$\text{(5.3)} \qquad |\Delta u_{n'}| \to 1 \quad \text{a.e. when} \quad n' \to \infty.$$

Compare (5.2) and (5.3): $\Delta u_{n'} \to 0$ weakly, and yet $\Delta u_{n'} \to 1$ almost everywhere. This means that the functions $\Delta u_{n'}$ switch very quickly between values close to $+1$ and -1, taking care to be half the time close to $+1$ and half the time close to -1. This is a characteristic feature of minimizing sequences of controls.

Remark 5.1. If we consider the problem:

$$(\mathscr{P}) \qquad \operatorname*{Inf}_{u \in H_0^1(\Omega)} \int_\Omega (f(u) + (1 - |\Delta u|)^2 \, dx$$

with f a non-negative continuous function growing no faster than u^2, we obtain similar results. The relaxed problem is:

$$(\mathscr{P}\mathscr{R}) \qquad \operatorname*{Min}_{u \in H_0^1(\Omega)} \int_\Omega (f(u) + (|\Delta u| - 1)_+^2) \, dx$$

and by Remark 4.2, from every minimizing subsequence u_n we can extract a subsequence $u_{n'}$ such that:

$$(|\Delta u_{n'}(x)| - 1)^2 - (|\Delta u_{n'}(x)| - 1)_+ \to 0 \quad \text{a.e.},$$

i.e., almost certainly, the open interval $(-1, +1)$ holds no cluster point of the sequence $\Delta u_{n'}(x)$.

Example 2

Denote by $\partial\Omega$ the boundary of Ω, and define $u = \mathscr{G}p$ as the unique solution of the inhomogeneous Dirichlet problem:

$$(5.4) \qquad \begin{cases} -\Delta u = p_0 & \text{in} \quad \Omega \\ u = p_1 & \text{on} \quad \partial\Omega, \end{cases}$$

where $p = (p_0, p_1)$. From Lions and Magenes [1], the mapping $\mathscr{G}$ is well-defined, linear, and continuous from $L^2(\Omega) \times L^2(\partial\Omega)$ into $H^{1/2}(\Omega)$, and thus compact into $L^2(\Omega)$, if Ω is regular enough.

Consider, for instance, the problem:

$$(\mathscr{P}) \qquad \begin{cases} \operatorname{Inf} \displaystyle\int_\Omega (f(u) + (p_0 - 1)^2)\,\mathrm{d}x + \int_{\partial\Omega} (p_1 - 1)^2\,\mathrm{d}\sigma \\ -\Delta u = p_0 \quad \text{in} \quad \Omega, \qquad u = p_1 \quad \text{on} \quad \partial\Omega. \end{cases}$$

The results of Section 4 are applicable (Remark 4.3). The relaxed problem is:

$$(\mathscr{P}\mathscr{R}) \qquad \begin{cases} \operatorname{Min} \displaystyle\int_\Omega (f(u) + (p_0 - 1)_+^2)\,\mathrm{d}x + \int_{\Omega} (p_1 - 1)_+^2\,\mathrm{d}\sigma \\ -\Delta u = p_0 \quad \text{in} \quad \Omega, \qquad u = p_1 \quad \text{on} \quad \partial\Omega, \end{cases}$$

with the usual results, provided f is a continuous non-negative function growing no faster than u^2.

Example 3

As in Section 1.2 of Chapter VIII, we denote by C a Borel subset of $\Omega \times \mathbf{R}^m$, such that the section C_x is closed and non-empty for almost every $x \in \Omega$. Let f denote its indicator function:

$$\begin{aligned} f(x, a) &= 0 && \text{if} \quad a \in C_x \\ f(x, a) &= +\infty && \text{if} \quad a \notin C. \end{aligned}$$

We denote by $\mathscr{S}$ the set of measurable selections of C (which is non-empty by Corollary VIII.1.1), and for $1 \leqslant \alpha < \infty$, we consider the mapping F on $L_m^\alpha(\Omega)$ defined by:

$$F(u) = \int_\Omega f(x, u(x))\,\mathrm{d}x, \qquad \forall u \in L_m^\alpha.$$

In other words:

$$\begin{aligned} F(u) &= 0 && \text{if} \quad u \in L_m^\alpha \cap \mathscr{S} \\ F(u) &= +\infty && \text{if} \quad u \in L_m^\alpha \setminus \mathscr{S}. \end{aligned}$$

Note that F is the indicator function of $\mathscr{S} \cap L^\alpha_m$ in L^α_m; in particular, F is weakly l.s.c. if and only if $\mathscr{S} \cap L^\alpha_m$ is weakly closed.

Let us now assume that $\mathscr{S} \cap L^\infty_m$ is non-empty, i.e., that there is at least one measurable selection of C which is essentially bounded. Let us then apply Corollary 2.1: $\mathscr{S} \cap L^\alpha_m$ is weakly closed in L^α_m if and only if the sections C_x are convex in $\mathbf{R}^m$, for almost every $x \in \Omega$. If they are not, we apply Proposition 2.3: the weak closure of $\mathscr{G} \cap L_m$ in L_m is the set $\mathscr{G} \cap L_m$, where $\mathscr{G}$ is the set of measurable selections of the closed convex hull $\overline{\text{co}}\, C_x$ of the sections C_x:[1]

$$\mathscr{G} \cap L^\alpha_m = \{u \in L^\alpha_m \,|\, u(x) \in \overline{\text{co}}\, C_x \quad \text{a.e.}\}.$$

APPENDIX

We give here, for the reader's convenience, a well-known result which we used in the proof of Lemma 4.2.

Lemma. *Let $\sum_{n=0}^\infty \alpha_n$ be a convergent series with positive terms. There exists an increasing sequence $(k_n)_{n\in\mathbf{N}}$ of integers such that:*

$$k_n \to \infty \quad \textit{when} \quad n \to \infty$$

the series $\sum_{n=0}^\infty k_n \alpha_n$ converges.

Proof. We set $S = \sum_{n=0}^\infty \alpha_n$ and $p_0 = 0$. We define $p_1 \in \mathbf{N}$ as the smallest integer > 0 such that $\sum_{n=p_1}^\infty \alpha_n \leqslant S/2$. By induction, for all $r \geqslant 1$, we define $p_r \in \mathbf{N}$ as the smallest integer $> p_{r-1}$ such that $\sum_{n=p_r}^\infty \alpha_n \leqslant S/2^r$. We then have:

$$2S = \sum_{r=1}^\infty \frac{S}{2^r} \geqslant \sum_{r=1}^\infty \sum_{n=p_r}^\infty \alpha_n.$$

For all $n \in \mathbf{N}$, there exists a unique number r such that $p_{r-1} \leqslant n < p_r$. We then take, by definition, $k_n = r$. Clearly, $k_n \to \infty$, and we have:

$$\sum_{n=0}^\infty k_n \alpha_n = \sum_{r=0}^\infty \sum_{n=p_r}^\infty \alpha_n \leqslant 2S < +\infty.$$

[1] Note that $\overline{\text{co}}\, C_x$ is not a section of $\overline{\text{co}}\, C$.

CHAPTER X

Relaxation of Non-convex Variational Problems (II)

Orientation

Here we continue the study of variational problems without solutions, which we began in the previous chapter. This chapter will be concerned with the fundamental problem of the calculus of variations for dimensions greater than one:

$$(\mathscr{P}) \qquad \left| \begin{array}{l} \operatorname{Inf} \displaystyle\int_\Omega f(x, u(x), \operatorname{grad} u(x))\, \mathrm{d}x \\ u \in W^{1,\alpha}(\Omega), \qquad u - u_0 \in W_0^{1,\alpha}(\Omega). \end{array} \right.$$

This problem does not fit into the framework of the preceding chapter. Indeed, the equation $\operatorname{grad} u = p$ does not have a solution for all functions p in $L^1(\Omega)$, which means that we cannot associate with it a Green operator $\mathscr{G}: p \to u$ defined over the whole of $L^1(\Omega)$. We must therefore develop special techniques, which are unfortunately more complicated than in the previous chapter, although the general theme remains the same: we are trying to relax the problem. Formally, the results obtained are practically identical: the greater than one:

$$(\mathscr{P}\mathscr{R}) \qquad \left| \begin{array}{l} \operatorname{Inf} \displaystyle\int_\Omega f^{**}(x, u(x), \operatorname{grad} u(x))\, \mathrm{d}x \\ u \in W^{1,\alpha}(\Omega), \qquad u - u_0 \in W_0^{1,\alpha}(\Omega). \end{array} \right.$$

and we shall show in passing that the Γ-regularization of certain functionals of this type on $W_0^{1,\alpha}(\Omega)$ coincides with their l.s.c. regularization.

Finally, we have assembled at the end of this chapter all those results which concern the Euler equations: we shall show that, even if a variational problem has no solution, the associated variational equations have at least approximate solutions. A few examples will help the understanding.

1. AN APPROXIMATION RESULT

1.1. Extension of Lipschitzian functions

The following is an essential tool:

MacShane's lemma. *Let X be a metric space, E a subspace of X, k a positive real number. Then any k-Lipschitz mapping from E into* $\mathbf{R}$ *can be extended by a k-Lipschitz mapping from X into* $\mathbf{R}$.

Proof. For all $x \in X$, we define:

$$\tilde{u}(x) = \sup_{e \in E} \{ u(e) - kd(x, e) \} \in \mathbf{R} \cup \{ +\infty \}. \tag{1.1}$$

If $\bar{e} \in E$, we have, from the Lipschitz condition over E, $u(e) - kd(\bar{e}, e) \leqslant u(\bar{e})$ for all e in E, and so $\tilde{u}(\bar{e}) = u(\bar{e})$. Thus $\tilde{u}$ is an extension of u.

Let x and y belong to X, with for example $\tilde{u}(x) \leqslant \tilde{u}(y)$. Then:

(1.2)

$$\begin{aligned} 0 \leqslant \tilde{u}(y) - \tilde{u}(x) &= \sup_{e \in E} \{ u(e) - kd(y, e) \} - \sup_{e \in E} \{ u(e) - kd(x, e) \} \\ &\leqslant \sup_{e \in E} \{ u(e) - kd(y, e) - u(e) + kd(x, e) \} \\ &\leqslant k \sup_{e \in E} \{ d(x, e) - d(y, e) \}. \end{aligned}$$

Finally, from the triangle inequality:

$$0 \leqslant \tilde{u}(y) - \tilde{u}(x) \leqslant kd(y, x). \tag{1.3}$$

In particular, if we take x in E, we have $\tilde{u}(y) \leqslant u(x) + kd(y, x) < +\infty$ for all y in X and thus $\tilde{u}$ is indeed a mapping from X into $\mathbf{R}$. The inequality (1.3) then means that $\tilde{u}$ is Lipschitz with constant k. ■

1.2. Convex combination of piecewise affine functions

Henceforth, Ω will be a bounded open subset of $\mathbf{R}^n$, with boundary $\partial\Omega$. We shall call a function $u: \Omega \to \mathbf{R}$ *affine* if it is the restriction to Ω of an affine function over $\mathbf{R}^n$. In particular, if u is affine, $\operatorname{grad} u$ is constant, the converse being true if Ω is connected. We then have $u(x) = \langle x, \operatorname{grad} u \rangle + \text{constant}$. We say that a function $u: \Omega \to \mathbf{R}$ is *piecewise affine* if it is continuous and if there exists a partition of Ω into a negligible set and a finite number of open sets on which u is affine.

We shall also fix a family $(\alpha_k)_{1 \leqslant k \leqslant p}$ of real positive numbers with sum 1. The approximation method developed in the previous chapter easily gives us, for the one-dimensional case ($n = 1$):

Proposition 1.1. *Let Ω be a bounded open interval in* $\mathbf{R}$. *If u_k, $1 \leqslant k \leqslant p$, are affine functions from Ω into* $\mathbf{R}$, *for all given $\varepsilon > 0$ there exists a piecewise affine function $u: \Omega \to \mathbf{R}$ and m disjoint open sets Ω_k, $1 \leqslant k \leqslant p$, such that:*

$$\text{(1.4)} \qquad \operatorname{meas} \Omega_k = \alpha_k \operatorname{meas} \Omega \quad \text{for} \quad 1 \leqslant k \leqslant p,$$

$$\text{(1.5)} \qquad \forall x \in \Omega_k, \qquad u'(x) = u'_k,$$

$$\text{(1.6)} \qquad \forall x \in \Omega, \qquad \left| u(x) - \sum_{k=1}^{p} \alpha_k u_k(x) \right| \leqslant \varepsilon$$

$$\text{(1.7)} \qquad \forall x \in \partial\Omega, \qquad u(x) = \sum_{k=1}^{p} \alpha_k u_k(x).$$

Proof. We have $\Omega =]a, b[$. We take $i \in \mathbf{N}$ and divide Ω into 2^i equal sub-intervals and each of these into p sub-intervals with lengths proportional respectively to $\alpha_1, \alpha_2 \ldots, \alpha_p$. We denote by ${}_i\Omega_k$ the union of the 2^i open sub-intervals with length $\alpha_k(b-a)/2^i$, and by ${}_iu$ the unique piecewise affine function such that:

$$\text{(1.8)} \qquad {}_iu(a) = \sum_{k=1}^{p} \alpha_k u_k(a)$$

$$\text{(1.9)} \qquad {}_iu' = u'_k \quad \text{on} \quad {}_i\Omega_k, \qquad \text{for} \quad 1 \leqslant k \leqslant p.$$

Clearly, (1.4), (1.5) and (1.6) are satisfied by ${}_iu$ and the ${}_i\Omega_k$ for all $i \in \mathbf{N}$. Moreover, the sequence $({}_iu)_{i \in \mathbf{N}}$ converges uniformly to $\sum_{k=1}^{p} \alpha_k u_k$ and we can thus choose i sufficiently large so that (1.7) is also satisfied. ∎

It is obvious that this result cannot be extended to higher dimensions as it stands. In fact, if we split up Ω into p open subsets Ω_k, the function $\sum_{k=1}^{p} \mathbf{1}_{\Omega_k} \operatorname{grad} u_k$ is by no means necessarily a gradient if $n > 1$. It is thus necessary to appeal to MacShane's lemma to obtain an approximation u which will be no longer piecewise affine but locally Lipschitz instead. We recall that saying that u is locally Lipschitz with constant k is equivalent to saying that all the derivatives[1] $\partial u / \partial x_j$, $1 \leqslant j \leqslant n$, are in L^∞, with $|\operatorname{grad} u(x)| \leqslant k$ almost everywhere.

We shall use the notation of Chapter IX, Section 1.1: for $i \in \mathbf{N}$, $\mathscr{K}_i$ will denote the set of hypercubes of Ω of the form:

$$K = \prod_{j=1}^{p} [m_j 2^{-i}, (m_j + 1) 2^{-i}],$$

[1] In the sense of distributions.

where the $m_j \in \mathbf{N}$ and A_i will denote their union so that

$$\text{meas}\, \Omega = \lim_{i \to \infty} \text{meas}\, A_i.$$

Theorem 1.2. *Let u_k, $1 \leqslant k \leqslant p$ be piecewise affine functions from Ω into* $\mathbf{R}$. *For all $\varepsilon > 0$, there exists a locally Lipschitz function u, on Ω and p disjoint open spaces Ω_k, $1 \leqslant k \leqslant p$, such that:*

$$|\text{meas}\, \Omega_k - \alpha_k \,\text{meas}\, \Omega| \leqslant \alpha_k \varepsilon \quad \textit{for} \quad 1 \leqslant k \leqslant p \tag{1.10}$$

$$\text{grad}\, u(x) = \text{grad}\, u_k(x) \quad \textit{a.e. on } \Omega_k, \qquad 1 \leqslant k \leqslant p \tag{1.11}$$

$$|\text{grad}\, u(x)| \leqslant \max_{1 \leqslant k \leqslant p} \{ |\text{grad}\, u_k(x)| \} \quad \textit{a.e. on } \Omega \tag{1.12}$$

$$\forall x \in \Omega, \qquad \left| u(x) - \sum_{k=1}^{p} \alpha_k u_k(x) \right| \leqslant \varepsilon \tag{1.13}$$

$$\forall x \in \partial\Omega, \qquad u(x) = \sum_{k=1}^{p} \alpha_k u_k(x). \tag{1.14}$$

Proof. We shall prove the theorem in the particular case where the u_k are affine on Ω. We can deduce the general case by making a partition[1] of Ω into open spaces on which the u_k are affine, by constructing u over each of these open spaces and by collecting the pieces together by virtue of (1.14). Moreover we shall assume that the constants grad u_k are not all identical; for if they were, the property would be trivial on taking $u = \sum_{k=1}^{p} \alpha_k u_k$.

We then reason by induction simultaneously in n and p. The result is true in the one-dimensional case: it suffices to apply Proposition 1.1 to each of the connected components of Ω. Let us assume that the result has been proved up to and including dimension $n - 1$. In the nth dimension the result is trivially true for $p = 1$. Let us assume it has been established for $p - 1$. If we now prove it for p, it will be true for all p (induction on p) and hence for all n (induction on n).

We set

$$\bar{u}_k = u_k - \sum_{k=1}^{p} \alpha_k u_k.$$

By changing coordinates in $\mathbf{R}^n$, we arrive at the case where $\bar{u}_1$ only depends on the first coordinate:

$$\bar{u}_1(x) = g_1 x_1 + c_1. \tag{1.15}$$

[1] Up to a negligible set.

We set

$$\rho = \max_{1 \leqslant k \leqslant p} \{ |\text{grad } u_k| \} - \left| \sum_{k=1}^{p} \alpha_k \text{ grad } u_k \right| \tag{1.16}$$

$$\rho > 0. \tag{1.17}$$

Indeed, all boundary points of an Euclidean ball of $\mathbf{R}^n$ are extremal. We cannot therefore have $\rho = 0$ unless all the grad u_k are equal, contrary to hypothesis.

Let us choose $i \in \mathbf{N}$ sufficiently large so that:

$$\text{meas}\,(\Omega - A_i) \leqslant \varepsilon/2 \tag{1.18}$$

and let us denote by N the number of hypercubes of $\mathscr{K}_i$. We take $K \in \mathscr{K}_i$, $K = \Pi_{j=1}^n [a_j, b_j]$; for all $\eta \in]0, 2^{-(i+1)}[$ we denote by $\tilde{K}_\eta$ the hypercube $\Pi_{j=2}^n]a_j + \eta, b_j - \eta[$. Let us choose η sufficiently small so that:

$$\text{meas}\,(K - [a_1, b_1] \times \tilde{K}_\eta) \leqslant \varepsilon/4N \tag{1.19}$$

$$\eta \max_{1 \leqslant k \leqslant p} \{ |\text{grad } u_k| \} \leqslant \varepsilon/2. \tag{1.20}$$

We then subdivide $[a_1, b_1]$ into intervals of length $(b_1 - a_1)2^{-m}$, $m \in \mathbf{N}$, and each of these into two sub-intervals which are open and have lengths $\alpha_1(b_1 - a_1)2^{-m}$ and $(1 - \alpha_1)(b_1 - a_1)2^{-m}$ respectively. We denote by I_m the union of the former and by J_m the union of the latter so that

$$[a_1, b_1] = I_m \cup J_m \tag{1.21}$$

$$\text{length of } I_m = \alpha_1(b_1 - a_1) \tag{1.22}$$

$$\text{length of } J_m = (1 - \alpha_1)(b_1 - a_1). \tag{1.23}$$

We then define a piecewise affine mapping v_m from $[a_1, b_1]$ into $\mathbf{R}$ as follows:

$$v_m(a_1) = 0 \tag{1.24}$$

$$v'_m(x_1) = g_1 \quad \text{on } I_m \tag{1.25}$$

$$v'_m(x_1) = \frac{-\alpha_1}{1 - \alpha_1} g_1 \quad \text{on } J_m. \tag{1.26}$$

We thus have $v_m(b_1) = 0$, and the v_m converge uniformly to zero on $[a_1, b_1]$. We can therefore choose $m \in \mathbf{N}$ sufficiently large so that:

$$\forall x_1 \in [a_1, b_1], \qquad |v_m(x_1)| \leqslant \min \{ \varepsilon/2, \rho\eta/2 \}. \tag{1.27}$$

Having fixed m in this way, we define

(1.28) $$\Omega_1 = I_m \times \tilde{K}_\eta$$

(1.29) $$\forall x \in \Omega_1, \qquad u(x) = v_m(x_1) + \sum_{k=1}^{p} \alpha_k u_k(x).$$

From (1.19), (1.22) and (1.28) we deduce that:

(1.30) $$(\text{meas}\, \Omega_1 - \alpha_1 \, \text{meas}\, K) \leqslant \alpha_1 \frac{\varepsilon}{4N}.$$

From (1.27) and (1.29) we deduce that:

(1.31) $$\forall x \in \Omega_1, \qquad \left| u(x) - \sum_{k=1}^{p} \alpha_k u_k(x) \right| \leqslant \min \{ \varepsilon/2, \rho\eta/2 \}.$$

By differentiating (1.29) and substituting (1.15) into it:

(1.32) $$\forall x \in \Omega_1, \qquad \text{grad}\, u(x) = \text{grad}\, u_1.$$

By definition, we have $\sum_{k=1}^{p} \alpha_k \bar{u}_k = 0$. Over each connected component of $J_m \times \tilde{K}_\eta$, we have from (1.26):

(1.33) $$\text{grad}\, v_m = \sum_{k=2}^{p} \frac{\alpha_k}{1 - \alpha_1} \text{grad}\, \bar{u}_k$$

(1.34) $$\text{grad}\left(v_m + \sum_{k=1}^{p} \alpha_k u_k \right) = \sum_{k=2}^{p} \frac{\alpha_k}{1 - \alpha_1} \text{grad}\, u_k.$$

We then apply the induction hypothesis for each of the connected components of $J_m \times \tilde{K}_\eta$ to the $(p-1)$ scalars $\alpha_k/(1-\alpha_1)$ which are positive with sum 1, to the $(p-1)$ functions u_k, $2 \leqslant k \leqslant p$ and to $\min_{2 \leqslant k \leqslant p}\{\alpha_k \varepsilon/4N, \rho\eta/2\}$. There thus exists a mapping $u : J_m \times \tilde{K}_\eta \to R$ which is locally Lipschitz and $(p-1)$ open disjoint spaces Ω_k, $2 < k < p$ such that:

(1.35)
$$\left| \text{meas}\, \Omega_k - \frac{\alpha_k}{1 - \alpha_1} \text{meas}\, (J_m \times \tilde{K}_n) \right| \leqslant \alpha_k \frac{\varepsilon}{4N}, \qquad 2 \leqslant k \leqslant p$$

(1.36) $$\text{grad}\, u(x) = \text{grad}\, u_k(x) \quad \text{a.e. on } \Omega_k, \qquad 2 \leqslant k \leqslant p$$

(1.37) $$|\text{grad}\, u(x)| \leqslant \max_{2 \leqslant k \leqslant p} \{ |\text{grad}\, u_k| \} \qquad \text{a.e. on } J_m \times \tilde{K}_n$$

$$\forall x \in J_m \times \tilde{K}_\eta, \qquad \left| u(x) - v_m(x) - \sum_{k=1}^{p} \alpha_k u_k(x) \right| \leqslant \min\{\varepsilon/2, \rho\eta/2\} \tag{1.38}$$

$$\forall x \in \partial(J_m \times \tilde{K}_\eta), \qquad u(x) = v_m(x) + \sum_{k=1}^{m} \alpha_k u_k(x). \tag{1.39}$$

We transform (1.35) by (1.19) and (1.23):

$$|\text{meas}\, \Omega_k - \alpha_k \,\text{meas}\, K| \leqslant \alpha_k \frac{\varepsilon}{2N}, \qquad 2 \leqslant k \leqslant p. \tag{1.40}$$

We have thus defined u over $I_m \times \tilde{K}_\eta$ and over $J_m \times \tilde{K}_\eta$. By comparing (1.29) and (1.39), we ascertain that these two definitions can be pieced together to define a single continuous function u on $[a_1, b_1] \times \tilde{K}_\eta$. Since v_m has been constructed in such a way as to be null at a_1 and at b_1, we thus obtain:

$$u(x) = \sum_{k=1}^{p} \alpha_k u_k(x) \quad \text{if} \quad x \in \{a_1, b_1\} \times \tilde{K}_\eta. \tag{1.41}$$

Let us make an extension of u over the whole of ∂K by:

$$u(x) = \sum_{k=1}^{p} \alpha_k u_k(x) \quad \text{if} \quad x \in \partial K. \tag{1.42}$$

The function thus defined over $\partial K \cup [a_1, b_1] \times \tilde{K}_\eta$ is continuous, by virtue of (1.41). It is even Lipschitz with constant $\max_{1 \leqslant k \leqslant p}\{|\text{grad}\, u_k|\}$. This is easily checked over ∂K (by virtue of (1.42)) and over $\tilde{K}_\eta$ (by virtue of (1.32) and (1.37)) separately. Let us now take a point x of ∂K, with $x_1 \neq a_1$ and $x_1 \neq b_1$, and a point y of $[a_1, b_1] \times \tilde{K}_\eta$. Let us denote by z the point where the segment joining x and y cuts $[a_1, b_1] \times \partial \tilde{K}_\eta$. We have:

$$\eta \leqslant |z - x|. \tag{1.43}$$

Making use of (1.42) and of the triangle inequality, we can write:

$$|u(x) - u(y)| \leqslant \left| \sum_{k=1}^{p} \alpha_k u_k(x) - \sum_{k=1}^{p} \alpha_k u_k(z) \right| + \left| \sum_{k=1}^{p} \alpha_k u_k(z) - u(z) \right| + |u(z) - u(y)|. \tag{1.44}$$

The second term is bounded above by $\rho\eta$ (by virtue of (1.31) if $z_1 \in I_m$, by virtue of (1.27) and (1.38) if $z_1 \in J_m$). The third term can be bounded above using the Lipschitz condition on $[a_1, b_1] \times \tilde{K}_\eta$. This gives us:

$$|u(x) - u(y)| \leqslant \left| \sum_{k=1}^{p} \alpha_k \operatorname{grad} u_k \right| |z - x| + \rho\eta + \max_{1 \leqslant k \leqslant p} \{ |\operatorname{grad} u_k| \} |y - z|. \tag{1.45}$$

Making use of (1.43):

$$|u(x) - u(y)| \leqslant \left(\left| \sum_{k=1}^{p} \alpha_k \operatorname{grad} u_k \right| + \rho \right) |z - x| + \max_{1 \leqslant k \leqslant p} \{ |\operatorname{grad} u_k| \} |y - z|. \tag{1.46}$$

Hence, from the definition of ρ (1.16):

$$|u(x) - u(y)| \leqslant \max_{1 \leqslant k \leqslant p} \{ |\operatorname{grad} u_k| \} (|z - x| + |y - z|). \tag{1.47}$$

As z belongs to the line segment joining x and y:

$$|u(x) - u(y)| \leqslant \max_{1 \leqslant k \leqslant p} \{ |\operatorname{grad} u_k| \} (|x - y|). \tag{1.48}$$

From MacShane's lemma, we can extend u as a Lipschitz function with constant $\max_{1 \leqslant k \leqslant p}\{|\operatorname{grad} u_k|\}$ over the whole of K.

Let us recall the properties of this extension, still denoted by $\bar{u}$:

$$|\operatorname{meas} \Omega_k - \alpha_k \operatorname{meas} K| \leqslant \alpha_k \frac{\varepsilon}{2N}, \qquad 1 \leqslant k \leqslant p \tag{1.49}$$

$$\operatorname{grad} u(x) = \operatorname{grad} u_k \quad \text{a.e. on } \Omega_k, \qquad 1 \leqslant k \leqslant p \tag{1.50}$$

$$|\operatorname{grad} u(x)| \leqslant \max_{1 \leqslant k \leqslant p} \{ |\operatorname{grad} u_k| \} \quad \text{a.e. on } K \tag{1.51}$$

$$\forall x \in K, \qquad \left| u(x) - \sum_{k=1}^{p} \alpha_k u_k(x) \right| \leqslant \varepsilon \tag{1.52}$$

$$\forall x \in \partial K, \qquad u(x) = \sum_{k=1}^{p} \alpha_k u_k(x). \tag{1.53}$$

Property (1.49) arises from (1.30) and (1.40), (1.50) from (1.32) and (1.36), (1.53) from (1.42), and (1.51) is none other than the Lipschitz condition over

K. Property (1.52) arises from (1.31) if $x \in \Omega_1$, from (1.27) and (1.38) if $x \in J_m \times \tilde{K}_\eta$. There remains the case when $x \in K - [a_1, b_1] \times \tilde{K}_\eta$. It is then sufficient to choose a point $y \in \partial K$ such that $|y - x| \leqslant \eta$, and, taking into account (1.42), to write

(1.54)

$$\left| u(x) - \sum_{k=1}^{p} \alpha_k u_k(x) \right| \leqslant |u(x) - u(y)| + \left| \sum_{k=1}^{p} \alpha_k u_k(y) - \sum_{k=1}^{p} \alpha_k u_k(x) \right|$$

Using the Lipschitz condition:

(1.55)
$$\begin{aligned} \left| u(x) - \sum_{k=1}^{p} \alpha_k u_k(x) \right| &\leqslant 2 \max_{1 \leqslant k \leqslant p} \{ |\mathrm{grad}\, u_k| \} \, |x - y| \\ &\leqslant 2 \max_{1 \leqslant k \leqslant p} \{ |\mathrm{grad}\, u_k| \} \, \eta \\ &\leqslant \varepsilon \ \text{(from (1.20))}. \end{aligned}$$

We have thus defined u and the Ω_k over all the hypercubes K of $\mathscr{K}_i$. We piece u and the Ω_k together, and extend u by $\sum_{k=1}^{m} \alpha_k u_k$ out of A_i. The function thus obtained is continuous by virtue of (1.53), and satisfies (1.14). It satisfies (1.11) due to (1.50), (1.12) due to (1.51), and (1.12) due to (1.52). It only remains to show that it satisfies (1.10); by virtue of (1.18) and (1.49) it is sufficient to write:

(1.56)
$$\begin{aligned} |\mathrm{meas}\, \Omega_k - \alpha_k \,\mathrm{meas}\, \Omega| &\leqslant |\mathrm{meas}\, \Omega_k - \alpha_k \,\mathrm{meas}\, A_i| + \alpha_k |\mathrm{meas}\, A_i - \mathrm{meas}\, \Omega| \\ &\leqslant \alpha_k \frac{\varepsilon}{2} + \alpha_k \frac{\varepsilon}{2} \leqslant \alpha_k \varepsilon. \ \blacksquare \end{aligned}$$

1.3. Convex combination of integrals

We shall now apply the foregoing result to the study of convex combinations of integrals:

$$\sum_{k=1}^{p} \alpha_k \int_\Omega f(x, \mathrm{grad}\, u_k(x))\, dx.$$

In other words, with reference to Chapter IX, Theorem 1.2 is analogous to Proposition 1.1, and we seek the analogy of Corollary 1.2. As before, we settle on a family $(\alpha_k)_{1 \leqslant k \leqslant p}$ of real positive numbers with sum 1.

Proposition 1.3. *Let u_k, $1 \leqslant k \leqslant p$, be piecewise affine functions from Ω into $\mathbf{R}$. We take a finite family $\mathscr{F}$ of normal integrands of $\Omega \times \mathbf{R}^n$ and a positive function $c \in L^1(\Omega)$ which satisfy:*

(1.57)

$$c(x) \geqslant \sup \{ |f(x, \xi)| \,|\, f \in \mathscr{F}, |\xi| \leqslant \max_{1 \leqslant k \leqslant p} \{ \|\text{grad}\, u_k\|_\infty \}\} \quad \text{a.e.}$$

For all $\varepsilon > 0$, there exists a locally Lipschitz function u: $\Omega \to \mathbf{R}$ such that:

(1.58)

$$\forall f \in \mathscr{F} \qquad \left| \int_\Omega f(x, \text{grad}\, u(x))\, dx - \sum_{k=1}^{p} \alpha_k \int_\Omega f(x, \text{grad}\, u_k(x))\, dx \right| \leqslant \varepsilon$$

(1.59) $$|\text{grad}\, u(x)| \leqslant \max_{1 \leqslant k \leqslant p} \{ |\text{grad}\, u_k(x)| \} \quad \textit{a.e. on } \Omega$$

(1.60) $$\forall x \in \Omega, \qquad \left| u(x) - \sum_{k=1}^{p} \alpha_k u_k(x) \right| \leqslant \varepsilon$$

(1.61) $$\forall x \in \partial\Omega, \qquad u(x) = \sum_{k=1}^{p} \alpha_k u_k(x).$$

Proof. Here also it is sufficient to derive the proof for the case where the functions u_k are affine over Ω. The general case may be deduced therefrom by dividing Ω into open spaces over which the u_k are affine, by constructing u over each of these open spaces, and by collecting together the pieces by virtue of (1.61). Let us therefore take $\varepsilon > 0$.

We know that the tiered functions are dense in $L^1(\Omega)$. There thus exists a partition of Ω into a finite number of open subsets $\mathcal{O}_i$, $1 \leqslant i \leqslant N$ and a negligible set and there exist functions $\tilde{f}_k$, constant over each of the $\mathcal{O}_i$ such that:

(1.62)

$$\forall f \in \mathscr{F}, \qquad \int_\Omega |f(x, \text{grad}\, u_k) - \tilde{f}_k(x)|\, dx \leqslant \varepsilon \quad \text{for} \quad 1 \leqslant k \leqslant p.$$

Let us now choose $\delta > 0$ which satisfies the following two properties:

(1.63) $$\delta \leqslant \frac{\varepsilon}{N(1 + \max_{1 < k < p} \{ \|\tilde{f}_k\|_\infty \})}$$

$$(1.64) \qquad B \subset \Omega \text{ measurable and } \operatorname{meas} B \leqslant \delta \Rightarrow \int_B c \leqslant \varepsilon/N.$$

We apply Theorem 1.2 to each of the open subsets $\mathcal{O}_i$: there exists a locally Lipschitz function u from $\mathcal{O}_i$ into $\mathbf{R}$ and there exist p open disjoints spaces Ω_k^i, $1 \leqslant k \leqslant p$, such that:

$$(1.65) \qquad |\operatorname{meas} \Omega_k^i - \alpha_k \operatorname{meas} \mathcal{O}_i| \leqslant \alpha_k \delta \quad \text{for} \quad 1 \leqslant k \leqslant p$$

$$(1.66) \qquad \operatorname{grad} u(x) = \operatorname{grad} u_k \quad \text{a.e. on } \Omega_k^i, \qquad 1 \leqslant k \leqslant p$$

$$(1.67) \qquad |\operatorname{grad} u(x)| \leqslant \max_{1 \leqslant k \leqslant p} \{|\operatorname{grad} u_k|\} \quad \text{a.e. on } \mathcal{O}_i$$

$$(1.68) \qquad \forall x \in \mathcal{O}_i, \qquad \left| u(x) - \sum_{k=1}^{p} \alpha_k u_k(x) \right| \leqslant \delta$$

$$(1.69) \qquad \forall x \in \partial\mathcal{O}_i, \qquad u(x) = \sum_{k=1}^{p} \alpha_k u_k(x).$$

We have thus constructed u over all the $\mathcal{O}_i$ and we set $u = \sum_{k=1}^{p} \alpha_k u_k$ over the negligible set $\Omega - \bigcup_{i=1}^{N} \mathcal{O}_i$. The function thus defined is continuous from (1.69) and satisfies (1.59) (and is thus locally Lipschitz), (1.60) and (1.61). It remains to verify (1.58). From (1.66) for all $f \in \mathscr{F}$:

$$(1.70) \qquad \int_{\mathcal{O}_i} f(x, \operatorname{grad} u(x))\, dx - \sum_{k=1}^{p} \int_{\Omega_k^i} f(x, \operatorname{grad} u_k)\, dx$$
$$= \int_{\mathcal{O}_i - \bigcup_{k=1}^{p} \Omega_k^i} f(x, \operatorname{grad} u(x))\, dx.$$

Let us examine the right-hand side. From (1.67) and (1.57), $|f(x, \operatorname{grad} u(x))| \leqslant c(x)$. From (1.65), $\operatorname{meas}(\mathcal{O}_i - \bigcup_{k=1}^{p} \Omega_k^i) \leqslant \delta$. From (1.64), the norm of the right-hand side is less than ε/N, whence:

$$(1.71) \qquad \left| \int_{\mathcal{O}_i} f(x, \operatorname{grad} u(x))\, dx - \sum_{k=1}^{p} \int_{\Omega_k^i} f(x, \operatorname{grad} u_k)\, dx \right| \leqslant \varepsilon/N.$$

Now, by virtue of (1.63) and (1.65), by taking account of the fact that f_k is constant over $\mathcal{O}_i$, we have:

$$(1.72) \qquad \sum_{k=1}^{p} \left| \int_{\Omega_k^i} \tilde{f}_k(x)\, dx - \alpha_k \int_{\mathcal{O}_i} \tilde{f}_k(x)\, dx \right| \leqslant \varepsilon/N.$$

We write $\Omega_k = \bigcup_{i=1}^N \Omega_k^i$. We have the fundamental inequality:

$$\left|\int_\Omega f(x, \operatorname{grad} u(x))\,dx - \sum_{k=1}^p \alpha_k \int_\Omega f(x, \operatorname{grad} u_k)\,dx\right| \tag{1.73}$$

$$\leqslant \left|\int_\Omega f(x, \operatorname{grad} u(x))\,dx - \sum_{k=1}^p \int_{\Omega_k} f(x, \operatorname{grad} u_k)\,dx\right|$$

$$+ \sum_{k=1}^p \int_{\Omega_k} \left| f(x, \operatorname{grad} u_k) - \tilde{f}_k(x)\right| dx + \sum_{k=1}^p \left|\int_{\Omega_k} \tilde{f}_k(x)\,dx - \alpha_k \int_\Omega \tilde{f}_k(x)\,dx\right|$$

$$+ \sum_{k=1}^p \alpha_k \int_\Omega |\tilde{f}_k(x) - f(x, \operatorname{grad} u_k)|\,dx.$$

On the right-hand side, the first and third terms are less than ε on summing the inequalities (1.71) and (1.72) from $i = 1$ to N. The second and fourth terms are similarly less than ε from (1.62). Whence we have the desired inequality:

$$\left|\int_\Omega f(x, \operatorname{grad} u(x))\,dx - \sum_{k=1}^p \alpha_k \int_\Omega f(x, \operatorname{grad} u_k)\,dx\right| \leqslant 4\varepsilon \quad \blacksquare \tag{1.74}$$

2. IDENTITY OF THE Γ-REGULARIZATION AND OF THE L.S.C. REGULARIZATION

In Chapter IX, Section 2, we studied functionals over $L^\alpha(\Omega)$ of the form

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

and we showed that their Γ-regularization coincided with their l.s.c. regularization. The aim of this section is to do the same for functionals of the type

$$F(u) = \int_\Omega f(x, \operatorname{grad} u(x))\,dx \qquad \text{over } W^{1,\alpha}(\Omega).$$

We shall achieve this by approximating the functions of $W^{1,\alpha}(\Omega)$ by piecewise affine functions and by applying Proposition 1.3. We must therefore establish at the outset properties of density and continuity in $W^{1,\alpha}(\Omega)$; for this it will be necessary to distinguish between the cases $1 \leqslant \alpha \leqslant \infty$ and $\alpha = \infty$.

2.1. Approximation of an indefinitely differentiable function by a piecewise affine function

The following theorem is a standard result in numerical analysis (finite elements method).

Proposition 2.1. *Let u be a function which is indefinitely differentiable on a bounded open space Ω of $\mathbf{R}^n$ and continuous on $\bar{\Omega}$. There exists a sequence $(u_i)_{i\in\mathbf{N}}$ of piecewise affine functions on Ω such that, when $i \to \infty$,*

(2.1) $\quad$ $\operatorname{grad} u_i \to \operatorname{grad} u$ *uniformly over every compact subset of* Ω

$$\|\operatorname{grad} u_i\|_\infty \leqslant \|\operatorname{grad} u\|_\infty \tag{2.2}$$

$$u_i \to u \text{ uniformly.} \tag{2.3}$$

If in addition the support of u is compact in Ω, we can impose on the u_i, $i \in \mathbf{N}$ the condition that they are zero outside a fixed compact $K \subset \Omega$. In particular

$$\forall x \in \partial\Omega, \qquad u_i(x) = 0. \tag{2.4}$$

The proof consists of covering $\bar{\Omega}$ with a net of n-simplexes with diameter $\leqslant 2^{-i}$, of taking for u_i the piecewise affine function which coincides with u at the vertices of the net and of applying Taylor's formula. From this we deduce in particular, that the piecewise affine functions which are null on the boundary are dense in $W_0^{1,\alpha}(\Omega)$, $1 \leqslant \alpha < \infty$.

2.2. Lipschitzian open subsets of $\mathbf{R}^n$

We denote by B the unit ball of $\mathbf{R}^n$. Henceforth we shall assume that the following regularity hypothesis holds on Ω:

Definition 2.2 *We shall term a bounded open space Ω of $\mathbf{R}^n$ Lipschitz if, for any point x of $\partial\Omega$, there exists a neighbourhood $\mathcal{O}$ of x such that $\Omega \cap \mathcal{O}$ is of the form*

$$\Omega \cap \mathcal{O} = [y \in \mathbf{R}^n \mid y^n \leqslant \theta(y^1, ..., y^{n-1})] \cap \mathcal{O}, \tag{2.5}$$

where θ is a Lipschitz function of $\mathbf{R}^{n-1}$ in $\mathbf{R}$, and the y_n^i are a system of Cartesian coordinates in $\mathbf{R}^n$.

We immediately deduce from this that $\partial\Omega$ locally allows a representation of the form $y^n = \theta(y^1, \ldots, y^{n-1})$, where θ is a Lipschitz function. The class of Lipschitz bounded open subsets is large. Thus, if $\bar{\Omega}$ is a compact polyhedron, Ω

is Lipschitz. If Ω is a regular open space in the usual sense, i.e. if $\bar{\Omega}$ is a C^∞ submanifold of $\mathbf{R}^n$ with boundary $\partial\Omega$, then Ω is Lipschitz.

We shall subsequently make use of the fact that every Lipschitz bounded open subset Ω is locally star-shaped, i.e. Ω is a finite union of spaces $\mathcal{O}_i$, each one of which is star-shaped with respect to one of its points ω_i (i.e. $h(\mathcal{O}_i) \subset \mathcal{O}_i$ for every homothety h with centre ω_i and ratio $\leqslant 1$).

We prove this briefly. From the compactness of $\bar{\Omega}$, it is sufficient to show that each point x of $\bar{\Omega}$ possesses a neighbourhood $\mathcal{O}$ such that $\Omega \cap \mathcal{O}$ is star-shaped with respect to one of its points. Clearly, this is so if $x \in \Omega$: it is sufficient to take for $\mathcal{O}$ a ball with centre x which is contained in Ω. If $x \in \partial\Omega$, this is locally reduced to the case where

$$\Omega = \{\, y \in \mathbf{R}^n \mid y^n \leqslant \theta(\tilde{y}) \,\},$$

where $\tilde{y} = (y^1, \ldots y^{n-1})$ and where $\theta : \mathbf{R}^{n-1} \to \mathbf{R}$ is Lipschitz. Furthermore, we can assume that $\tilde{x} = 0$ and that $x^n = \theta(\tilde{x}) > 0$. But it is then easily shown that $\tilde{x}$ possesses a neighbourhood $\mathscr{V}$ in $\mathbf{R}^{n-1}$ such that the cylinder $(\mathscr{V} \times \mathbf{R}) \cap \Omega$ is star-shaped with respect to the origin. It suffices to see that $\theta(\lambda\tilde{y}) \geqslant \lambda\theta(\tilde{y})$ for all y sufficiently close to zero and all $\lambda \in [0, 1]$. Let us therefore choose $|\tilde{y}| < \theta(0)/2k$, where k is the Lipschitz constant for θ. We then have for $0 \leqslant \lambda \leqslant 1$:

$$\begin{aligned} 0 &< \theta(0) - 2k\lambda|\tilde{y}| \\ &< \theta(\lambda\tilde{y}) - k\lambda|\tilde{y}| \\ &< \theta(\lambda\tilde{y}) - \lambda\theta(\tilde{y}). \end{aligned} \tag{2.6}$$

We call every Lipschitz mapping of an interval of $\mathbf{R}$ into $\bar{\Omega}$ a *path* in $\bar{\Omega}$. For x and y in $\bar{\Omega}$ we denote by $\delta(x, y)$ the infimum of 1 and of the lengths of the paths in Ω joining x to y. It is a distance over $\bar{\Omega}$. But we also have another distance over $\bar{\Omega}$: the Euclidean distance $d(x,y)$. We show that they are equivalent, i.e. that for all $\alpha > 0$ there exists $\beta > 0$ such that $\delta(x,y) \leqslant \beta$ implies $d(x,y) \leqslant \alpha$ and that $d(x,y) \leqslant \beta$ implies that $\delta(x,y) \leqslant \alpha$. This will subsequently be useful when we wish to avoid going back to local mappings in certain proofs.

Proposition 2.3. *If Ω is a bounded open Lipschitz space, the distances δ and d are equivalent over $\bar{\Omega}$.*

Proof. Let us denote by $\bar{\Omega}_\delta$ (or $\bar{\Omega}_d$) the set $\bar{\Omega}$ endowed with the metric δ (or d). We must show that the identity mapping from Ω_δ into Ω_d is uniformly continuous and conversely.

Clearly, $d \leqslant \delta$, and thus the identity mapping from $\bar{\Omega}_\delta$ into $\bar{\Omega}_d$ is uniformly continuous. It remains for us to show that the inverse mapping is continuous from $\bar{\Omega}_d$ into $\bar{\Omega}_\delta$: since $\bar{\Omega}_d$ is compact, it will then be uniformly continuous.

Let there be in $\bar{\Omega}$ a sequence $(x_i)_{i\in\mathbf{N}}$ and a point x such that $d(x,x_i) \to 0$. If $x \in \Omega$, there is an $\varepsilon > 0$ such that the Euclidean ball with centre x and radius ε is contained within Ω; we then have $\delta(x,y) = d(x,y)$ for all the points y in this ball, and so, after a certain stage, $\delta(x,x_i) = d(x,x_i) \to 0$.

If $x \in \partial\Omega$, in accordance with Definition 2.2, we arrive locally at the situation where

$$\bar{\Omega} = \{\, y \in \mathbf{R}^n \mid y^n \leqslant \theta(y^1, \dots, y^{n-1}) \,\}, \tag{2.7}$$

where $\theta : \mathbf{R}^{n-1} \to \mathbf{R}$ is a Lipschitz function with constant k, and where

$$x^1 = \dots = x^n = 0.$$

For each point y in Ω, we denote by z the point in Ω with co-ordinates

$$z = (y^1, \dots, y^{n-1}, -k[(y^1)^2 + \dots + (y^{n-1})^2]^{1/2}). \tag{2.8}$$

We have $z^n < y^n$, and the segment $[y,z]$ is this contained in Ω, as well as the segment $[x,z]$. Hence we have:

$$\begin{aligned} &\delta(x,y) \leqslant d(x,z) + d(z,y) \\ &\delta(x,y) \leqslant (1+k^2)^{1/2} d(x,y) + 2kd(x,y) \\ &\delta(x,y) \leqslant [(1+k^2)^{1/2} + 2k] d(x,y). \end{aligned} \tag{2.9}$$

After a certain stage, x_i enters a neighbourhood of x which is subject to the representation (2.7) and hence, by (2.9), $\delta(x,x_i) \to 0$. The result is established for all cases.

Here is an example of its application:

Corollary 2.4. *If Ω is a bounded open Lipschitz space, the following are equivalent to each other*:

(a) $$\|\mathrm{grad}\, u\|_\infty \leqslant k$$

(b) *u is δ-Lipschitz with constant k on $\bar{\Omega}$.*

In particular, $W^{1,\infty}(\Omega) \subset \mathscr{C}(\bar{\Omega})$.

Proof. We have $\delta(x,y) = d(x,y)$ whenever x and y belong to a ball contained within Ω. If then u is δ-Lipschitz with constant k over Ω, it is d-Lipschitz with constant k over all the balls contained within Ω, and hence $\|\mathrm{grad}\, u\|_\infty < k$.

Conversely, (a) means that u is locally d-Lipschitz with constant k. Then if c is a path in Ω joining x to y, we obtain:

$$|u(x) - u(y)| < k(x,y). \text{ Length } (c). \tag{2.10}$$

By definition of $\delta(x,y)$ as the infimum of the lengths of the paths joining x to y, we obtain (b):

$$|u(x) - u(y)| \leqslant k\delta(x, y). \tag{2.11}$$

In particular, u can be extended as a δ-Lipschitz function on $\bar{\Omega}$ with the same constant k; from Proposition 2.3, this extension will be d-continuous on $\bar{\Omega}$.

2.3. Approximation of a function of $W_0^{1,\alpha}$, $1 \leqslant \alpha < \infty$ by a piecewise affine function

Henceforth, Ω will be a Lipschitz bounded open subset of $\mathbf{R}^n$, and φ will be a non-negative convex mapping of $\mathbf{R}^n$ into $\overline{\mathbf{R}}_+$. We now introduce the following concept:

Definition 2.5. *We denote by $W_0^{1,\varphi}(\Omega)$ the set of functions $u \in W_0^{1,1}(\Omega)$ such that*:

$$\int_\Omega \varphi \circ \operatorname{grad} u < +\infty. \tag{2.12}$$

If for example we take $\varphi(\xi) = |\xi|^\alpha$, $1 \leqslant \alpha < \infty$, we obtain $W_0^{1,\varphi}(\Omega) = W_0^{1,\alpha}(\Omega)$, minus the topology; we thus recover the usual Sobolev spaces. But we can construct very different spaces: anisotropic (take for φ a positive quadratic form) or contained in $W_0^{1,\infty}(\Omega)$ (take $\varphi(\xi) = +\infty$ if $|\xi| \geqslant 1$, $\varphi(\xi) = |\xi|/(1 - |\xi|)$ if $|\xi| < 1$). The following proposition generalizes the fact that $\mathscr{D}(\Omega)$ is dense in $W_0^{1,\alpha}(\Omega)$ (in a weak form).

Proposition 2.6. *If $u \in W_0^{1,\varphi}(\Omega)$, there exists a sequence $(u_i)_{i\in\mathbf{N}}$ of functions of $\mathscr{D}(\Omega)$ such that, when $i \to \infty$*:

$$u_i \to u \quad \text{in} \quad L^1(\Omega) \tag{2.13}$$

$$\operatorname{grad} u_i \to \operatorname{grad} u \quad \text{in} \quad L_n^1(\Omega) \tag{2.14}$$

$$\int_\Omega \varphi \circ \operatorname{grad} u_i \to \int_\Omega \varphi \circ \operatorname{grad} u. \tag{2.15}$$

Proof. First we shall show that u can be assumed to have compact support in Ω, and then we shall prove the proposition for this case. To simplify the reasoning, we shall assume that $\varphi(0) = 0$. Since Ω is locally star-shaped we can write $\Omega = \bigcup_{j=1}^{N} \mathcal{O}_j$ where the $\mathcal{O}_j$ are open subsets which are star-shaped with respect to ω_j, $1 \leqslant j \leqslant N$. Then let $u \in W_0^{1,\varphi}(\Omega)$; the functions:

$$u^+(x) = \sup \{ u(x), 0 \},$$
$$u^-(x) = \sup \{ -u(x), 0 \},$$

also belong to $W_0^{1,\varphi}(\Omega)$. Extending u, u^+ and u^- by zero outside Ω, we obtain functions of $W_0^{1,\varphi}(\mathbf{R}^n)$ which will again be denoted by u, u^+ and u^-. Naturally, we have:

$$u = u^+ - u^-.$$

Let ρ_k, $k \in \mathbf{N}$, be an increasing sequence of positive numbers tending to 1, $0 < \rho_k < 1$. We set:

$$u_{k,j}^+(x) = \rho_k u^+\left(\omega_j + \frac{1}{\rho_k}(x - \omega_j)\right), \qquad x \in \mathbf{R}^n,\ 1 \leqslant j \leqslant N;$$

$$u_{k,j}^-(x) = \rho_k u^-\left(\omega_j + \frac{1}{\rho_k}(x - \omega_i)\right), \qquad x \in \mathbf{R}^n,\ 1 \leqslant j \leqslant N.$$

Clearly, $u_{k,j}^+ \to u^+$ in $W_0^{1,1}(\mathbf{R}^n)$ when $k \to \infty$, and furthermore:

$$\int_{\mathbf{R}^n} \varphi \circ \operatorname{grad} u_{k,j}^+(x)\,\mathrm{d}x = \int_{\mathbf{R}^n} \varphi \circ \operatorname{grad} u^+\left(\omega_j + \frac{1}{\rho_k}(x - \omega_j)\right)\mathrm{d}x$$

$$= \rho_k^n \int_{\mathbf{R}^n} \varphi \circ \operatorname{grad} u^+(y)\,\mathrm{d}y$$

by the rule for changing the variable in a multiple integral. Finally, when $k \to \infty$:

$$\int_{\mathbf{R}^n} \varphi \circ \operatorname{grad} u_{k,j}^+(x)\,\mathrm{d}x \to \int_{\mathbf{R}^n} \varphi \circ \operatorname{grad} u^+(x)\,\mathrm{d}x.$$

We have analogous results for $\bar{u}_{k,j}$. Lastly, we define:

$$\text{(2.16)} \qquad \left|\begin{array}{ll} u_k^+(x) = \inf\limits_{1 \leqslant j \leqslant N} \{ u^+(x), u_{k,j}^+(x) \}, & x \in \mathbf{R}^n \\ u_k^-(x) = \inf\limits_{1 \leqslant j \leqslant N} \{ u^-(x), u_{k,j}^-(x) \}, & x \in \mathbf{R}^n. \end{array}\right.$$

The functions u_k^+ and u_k^- converge to u^+ and u^- respectively in $W_0^{1,1}(\mathbf{R}^n)$ for $k \to \infty$. Furthermore, from (2.16):

$$
\left|\begin{aligned}
&\int_{\mathbf{R}^n} \varphi \circ \operatorname{grad} u_k^+(x)\,dx \to \int_{\mathbf{R}^n} \varphi \circ \operatorname{grad} u^+(x)\,dx \\
&\int_{\mathbf{R}^n} \varphi \circ \operatorname{grad} u_k^-(x)\,dx \to \int_{\mathbf{R}^n} \varphi \circ \operatorname{grad} u^-(x)\,dx.
\end{aligned}\right. \tag{2.17}
$$

We set:

$$
\begin{aligned}
B_k &= \{\, x \mid u_k^+(x) \neq 0 \,\} \subset \{\, x \mid u^+(x) \neq 0 \,\} \\
C_k &= \{\, x \mid u_k^-(x) \neq 0 \,\} \subset \{\, x \mid u^-(x) \neq 0 \,\}.
\end{aligned}
$$

We have $B_k \cap C_k = \varnothing$. The function $u_k = u_k^+ - u_k^-$ converges to u in $W_0^{1,1}(\mathbf{R}^n)$ when $k \to \infty$, and satisfies:

$$
\begin{aligned}
\int_{\mathbf{R}^n} \varphi \circ \operatorname{grad} u_k(x)\,dx &= \int_{B_k} \varphi \circ \operatorname{grad} u_k(x)\,dx + \int_{C_k} \varphi \circ \operatorname{grad} u_k(x)\,dx \\
&= \int_{\mathbf{R}^n} \varphi \circ \operatorname{grad} u_k^+(x)\,dx + \int_{\mathbf{R}^n} \varphi \circ \operatorname{grad} u_k^-(x)\,dx \\
&\to \int_{\mathbf{R}^n} \varphi \circ \operatorname{grad} u^+(x)\,dx + \int_{\mathbf{R}^n} \varphi \circ \operatorname{grad} u^-(x)\,dx \\
&= \int_{\mathbf{R}^n} \varphi \circ \operatorname{grad} u(x)\,dx.
\end{aligned}
$$

Finally, the support of u_k, by construction, is contained in

$$
\bigcup_{j=1}^{N} (\omega_j + \rho_k(\bar{\mathcal{O}}_j - \omega_j)),
$$

which is a compact set contained within Ω. We have thus been able to restrict ourselves to the case where $u \in W_0^{1,\varphi}(\Omega)$ has compact support in Ω, which we shall henceforth assume.

For $i \in \mathbf{N}$, let $\rho_i \in \mathscr{D}(\mathbf{R}^n)$ be non-negative with support contained within the ball $\{\xi \mid |\xi| \leqslant 2^{-i}\}$, and such that $\int_{\mathbf{R}^n} \rho = 1$. Let us set $u_i = u \star \rho_i$.[1] Since u has compact support, $u_i \in \mathscr{D}(\Omega)$ for i sufficiently large. We know that $u \star \rho_i \to u$ in

[1] Convolution product.

L^1 and that $\operatorname{grad}(u \bigstar \rho_i) = (\operatorname{grad} u) \bigstar \rho_i \to \operatorname{grad} u$ in L^1_n, whence (2.13) and (2.14). It remains for us to verify (2.15).

Since $\operatorname{grad} u = (\operatorname{grad} u) \bigstar \rho_1$, we have:

$$\varphi \circ \operatorname{grad} u_i(x) = \varphi\left(\int_{\mathbf{R}^n} \rho_i(x - y) \operatorname{grad} u(y) \, \mathrm{d}y\right).$$

But $\rho_i(x-y)\mathrm{d}y$ is a positive Radon measure on $\mathbf{R}^n$ of unit mass, and φ is convex and continuous. By virtue of a lemma whose proof we leave to the end (Lemma 2.7):

$$\varphi\left(\int_{\mathbf{R}^n} \rho_i(x - y) \operatorname{grad} u(y) \, \mathrm{d}y\right) \leqslant \int_{\mathbf{R}^n} \rho_i(x - y)\varphi \circ \operatorname{grad} u(y) \, \mathrm{d}y$$

To sum up:

$$0 \leqslant \varphi \circ \operatorname{grad} u_i \leqslant \rho_i \bigstar (\varphi \circ \operatorname{grad} u). \tag{2.18}$$

We know that, when $i \to \infty$, the sequence $(\rho_i \bigstar (\varphi \circ \operatorname{grad} u))_{i \in \mathbf{N}}$ converges to $\varphi \circ \operatorname{grad} u$ in $L^1(\Omega)$. In particular, it is equi-integrable in $L^1(\Omega)$ (Theorem VIII.1.3, (a) $\Rightarrow$ (c)); for all $\varepsilon > 0$ there exists $\delta > 0$ such that:

$$A \text{ measurable, meas } A \leqslant \delta \Rightarrow \int_A \rho_i \bigstar (\varphi \circ \operatorname{grad} u) \leqslant \varepsilon \qquad \forall i \in \mathbf{N}. \tag{2.19}$$

But from (2.29) and the estimate (2.18) we deduce that:

$$A \text{ measurable, meas } A \leqslant \delta \Rightarrow \int_A \varphi \circ \operatorname{grad} u_i \leqslant \varepsilon \qquad \forall i \in \mathbf{N}. \tag{2.20}$$

The sequence $(\varphi \circ \operatorname{grad} u_i)_{i \in \mathbf{N}}$ is thus equi-integrable. Moreover, each extracted subsequence possesses a subsequence which converges almost everywhere to $\varphi \circ \operatorname{grad} u$, since $\operatorname{grad} u_i \to \operatorname{grad} u$ in L^1 and φ is continuous. Therefore we can apply Corollary VIII.1.3, which gives us:

$$\varphi \circ \operatorname{grad} u_i \to \varphi \circ \operatorname{grad} u \quad \text{in} \quad L^1(\Omega) \tag{2.21}$$

whence in particular (2.15).

The proof has used the following lemma:

Lemma 2.7. *Let φ be a function in $\Gamma_0(\mathbf{R}^n)$, μ a positive Radon measure on $\mathbf{R}^n$ of unit mass and p a μ-integrable mapping from $\mathbf{R}^n$ into $\mathbf{R}^n$. We then have:*

$$\varphi(\textstyle\int p \,\mathrm{d}\mu) \leqslant \int \varphi \circ p \,\mathrm{d}\mu. \tag{2.22}$$

Proof. We can write φ as the pointwise supremum of a denumerable family of continuous affine functions:

$$\forall \xi \in \mathbf{R}^n, \qquad \varphi(\xi) = \sup_{i\in\mathbf{N}} \{ \langle \xi, \xi_i' \rangle + a_i \}. \tag{2.23}$$

For this, it is sufficient that the family $(\xi_i', -a_i)_{i\in\mathbf{N}}$ is dense in $\operatorname{epi}\varphi^*$. For all $i \in \mathbf{N}$, we have:

$$\langle \textstyle\int p(x) \,\mathrm{d}\mu(x), \xi_i' \rangle + a_i = \int (\langle p(x), \xi_i' \rangle + a_i) \,\mathrm{d}\mu(x). \tag{2.24}$$

Taking the supremum of each side for $i \in \mathbf{N}$ and applying (2.23), we obtain:

$$\varphi(\textstyle\int p \,\mathrm{d}\mu) = \sup_{i\in\mathbf{N}} \int (\langle p(x), \xi_i' \rangle + a_i) \,\mathrm{d}\mu(x). \tag{2.25}$$

But the integrable functions $\langle p, \xi_i' \rangle + a_i$ are less than $\varphi \circ p$ by virtue of (2.23). If $\int \varphi \circ p = +\infty$, the inequality (2.22) is trivial. Otherwise, $\varphi \circ p$ is integrable, and we can apply Fatou's lemma:

$$\sup_{i\in\mathbf{N}} \int (\langle p(x), \xi_i' \rangle + a_i) \,\mathrm{d}\mu(x) \leqslant \int \sup_{i\in\mathbf{N}} \{ \langle p(x), \xi_i' \rangle + a_i \} \,\mathrm{d}\mu(x). \tag{2.26}$$

Using (2.23) once again, we get:

$$\sup_{i\in\mathbf{N}} \int (\langle p(x), \xi_i' \rangle + a_i) \,\mathrm{d}\mu(x) \leqslant \int \varphi \circ p(x) \,\mathrm{d}\mu(x). \tag{2.27}$$

We obtain (2.22) on combining (2.25) and (2.27). ∎

We can thus approximate the functions of $W_0^{1,\varphi}(\Omega)$ by indefinitely differentiable functions with compact support (Prop. 2.6). But these in their turn can be approximated by piecewise affine functions (Prop. 2.1). This amounts to:

Proposition 2.8. *If $u \in W_0^{1,\varphi}(\Omega)$, there exists a sequence $(u_i)_{i\in\mathbf{N}}$ of piecewise affine functions over Ω, null on $\partial\Omega$, such that:*

$$u_i \to u \quad \textit{in} \quad L^1(\Omega) \tag{2.28}$$

$$\operatorname{grad} u_i \to \operatorname{grad} u \quad \textit{in} \quad L^1(\Omega) \tag{2.29}$$

$$\int_\Omega \varphi \circ \operatorname{grad} u_i \to \int_\Omega \varphi \circ \operatorname{grad} u. \tag{2.30}$$

Proof. Taking $\varepsilon > 0$, from Proposition 2.6, there exists $v \in \mathscr{D}(\Omega)$ such that:

$$\|u - v\|_1 \leqslant \varepsilon/2 \tag{2.31}$$

$$\|\operatorname{grad} u - \operatorname{grad} v\|_1 \leqslant \varepsilon/2 \tag{2.32}$$

$$|\textstyle\int \varphi \circ \operatorname{grad} u - \varphi \circ \operatorname{grad} v| \leqslant \varepsilon/2. \tag{2.33}$$

From Proposition 2,1, there exists a sequence $(v_i)_{i\in\mathbf{N}}$ of piecewise affine functions, null outside a compact $K \subset \Omega$, converging uniformly to v, such that:

$$\operatorname{grad} v_i \to \operatorname{grad} v \quad \text{uniformly.} \tag{2.34}$$

Since φ is continuous, $\varphi \circ \operatorname{grad} v_i$ converges uniformly to $\varphi \circ \operatorname{grad} v$. Taking $w = v_i$ for i sufficiently large, we conclude that:

$$\|v - w\|_1 \leqslant \varepsilon/2 \tag{2.35}$$

$$\|\operatorname{grad} v - \operatorname{grad} w\|_1 \leqslant \varepsilon/2 \tag{2.36}$$

$$|\textstyle\int \varphi \circ \operatorname{grad} v - \int \varphi \circ \operatorname{grad} w| = \varepsilon/2. \tag{2.37}$$

It only remains to gather together (2.31) and (2.35), (2.32) and (2.36), and (2.33) and (2.37). We then take $u_i = w$ for $\varepsilon = 2^{-i}$. ■

2.4. Approximation of a function of $W^{1,\infty}$ by a piecewise affine function

We recall that $W^{1,\infty}(\Omega)$ is the set of locally Lipschitz functions over Ω, i.e. of δ-Lipschitz functions over Ω, and that they can be extended by continuity to $\bar{\Omega}$ (Prop. 2.4). Let $u \in W^{1,\infty}(\Omega)$, not necessarily null on the boundary. We shall show that we can approximate u everywhere, except on a small neighbourhood of the boundary, by piecewise affine functions which satisfy the same boundary condition.

Proposition 2.9. *Let Ω be a Lipschitz bounded open subset of $\mathbf{R}^n$, and $u \in W^{1,\infty}(\Omega)$. There exists a sequence $(u_i, \Omega_i)_{i\in\mathbf{N}}$, where $u_i \in W^{1,\infty}(\Omega)$ and Ω_i is an open subspace of Ω, such that as $i \to \infty$,*

$$\Omega_i \subset \Omega_{i+1} \quad \textit{and} \quad \operatorname{meas}(\Omega - \Omega_i) \to 0 \tag{2.38}$$

$$u_i \textit{ is piecewise affine over } \Omega_i \tag{2.39}$$

$$\forall x \in \partial\Omega, \qquad u_i(x) = u(x) \tag{2.40}$$

$$u_i \to u \quad \textit{uniformly} \tag{2.41}$$

$$\operatorname{grad} u_i \to \operatorname{grad} u \qquad \textit{a.e. in } \Omega \tag{2.42}$$

$$\|\operatorname{grad} u_i\|_\infty \leqslant \|\operatorname{grad} u\|_\infty + c(i), \quad \textit{where } c(i) \to 0 \tag{2.43}$$

Proof. We set $\Omega_i = \{x \in \Omega | \delta(x, \partial\Omega) > 1/i\}$. We regularize u by convolution with a non-negative function $\rho \in \mathscr{D}(\mathbf{R}^n)$ such that $\int_{\mathbf{R}^n} \rho = 1$. If the support of ρ is sufficiently small, we obtain, on writing $k = \|\operatorname{grad} u\|_\infty$:

$$\forall x \in \Omega_{2i}, \qquad |\rho \star u(x) - u(x)| \leqslant 1/i^2 \tag{2.44}$$

$$\|\operatorname{grad} \rho \star u - \operatorname{grad} u\|_1 \leqslant 1/i \tag{2.45}$$

$$\|\operatorname{grad} \rho \star u\|_\infty \leqslant k. \tag{2.46}$$

We then apply Proposition 2.1 to $\rho \star u$ over the open subspace Ω_{2i}. There exists a piecewise affine function w_i over Ω_{2i}, such that:

$$\forall x \in \Omega_i, \qquad |\operatorname{grad} w_i(x) - \operatorname{grad} \rho \star u(x)| \leqslant 1/i \tag{2.47}$$

$$\forall x \in \Omega_i, \qquad |\operatorname{grad} w_i(x)| \leqslant k \tag{2.48}$$

$$\forall x \in \Omega_i, \qquad |w_i(x) - \rho \star u(x)| \leqslant 1/i^2. \tag{2.49}$$

Equation (2.47) results from (2.1) and from the fact that $\bar{\Omega}_i$ is compact in Ω_{2i}, equation (2.48) from (2.2) and (2.46), and equation (2.49) from (2.3).

We define the function u_i over $\partial\Omega \cup \Omega_i$ by:

$$u_i(x) = u(x) \quad \text{if} \quad x \in \partial\Omega \tag{2.50}$$

$$u_i(x) = w_i(x) \quad \text{if} \quad x \in \Omega_i. \tag{2.51}$$

By hypothesis, u is δ-Lipschitz with constant k; the same applies to w_i from (2.48). Thus from (2.50) and (2.51), we have:

$$|u_i(x) - u_i(y)| \leqslant k\delta(x, y) \quad \text{if} \quad x \text{ and } y \in \partial\Omega \tag{2.52}$$

$$|u_i(x) - u_i(y)| \leqslant k\delta(x, y) \quad \text{if} \quad x \text{ and } y \in \Omega_i. \tag{2.53}$$

If now $x \in \partial\Omega$ and $y \in \Omega_i$, we have, from (2.50), (2.51) and the triangle inequality:

$$|u_i(x) - u_i(y)| \leqslant |u(x) - u(y)| + |u(y) - w_i(y)|. \tag{2.54}$$

On the right-hand side, the first term is bounded above by $k\delta(x,y)$ since it is δ-Lipschitz with constant k. The second term is bounded above by $2/i^2$, by virtue of (2.44) and (2.49). Taking account of the fact that $y \in \Omega_i$, and hence that $\delta(x,y) > 1/i$, we obtain:

$$|u_i(x) - u_i(y)| \leqslant (k + 2/i)\delta(x, y). \tag{2.55}$$

Gathering together (2.52), (2.53) and (2.55), we deduce that u_i is δ-Lipschitz with constant $k + 2/i$. It can therefore be extended, from MacShane's lemma, as a function u_i over $\bar{\Omega}$, which is δ-Lipschitz with constant $(k + 2/i)$.

By construction, the functions $(u_i)_{i\in\mathbf{N}}$ satisfy (2.39), (2.40) and (2.43). Equation (2.42) results from (2.45), possibly by extracting a subsequence, and (2.47). Finally, the u_i are δ-Lipschitz with constant $(k + 2/i)$, and are thus equi-continuous; they converge simply over Ω ((2.44) and (2.49)), and therefore uniformly by Ascoli's theorem. Whence we arrive at (2.41). ∎

2.5. Functionals over $W_0^{1,\alpha}(\Omega)$, $1 < \alpha < \infty$

We recall the notation of Section 2.3: Ω is a Lipschitz bounded open subset of $\mathbf{R}^n$ and φ is a convex mapping from $\mathbf{R}^n$ into $\mathbf{R}_+$, and hence is continuous and non-negative. We shall study functionals of the type $\int f(x, \operatorname{grad} u(x))\,\mathrm{d}x$. We begin with a continuity theorem:

Proposition 2.10. *Let f be a Carathéodory function from $\Omega \times \mathbf{R}^n$ into $\mathbf{R}$, which satisfies the estimate*:

$$|f(x, \xi)| \leqslant a(x) + c\varphi(\xi), \qquad \textit{where } \alpha \in L^1(\Omega) \textit{ and } c \in \mathbf{R}_+. \tag{2.56}$$

For all sequences $(p_i)_{i\in\mathbf{N}}$ of measurable functions from Ω into $\mathbf{R}^n$ which satisfies, as $i \to \infty$:

$$p_i(x) \to \bar{p}(x) \qquad \textit{a.e. in } \Omega \tag{2.57}$$

$$\int_\Omega \varphi \circ p_i \to \int_\Omega \varphi \circ \bar{p} \tag{2.58}$$

we have:

$$\int_\Omega f(x, p_i(x))\,\mathrm{d}x \to \int_\Omega f(x, \bar{p}(x))\,\mathrm{d}x. \tag{2.59}$$

Proof. We consider the Carathéodory function

$$g(x, \xi) = f(x, \xi) + c\varphi(\xi) + a(x) \tag{2.60}$$

which is non-negative by (2.56).

When $i \to \infty$, $g(x,p_i(x))$ converges almost everywhere to $g(x,\bar{p}(x))$. Applying Fatou's lemma, we have:

$$\underline{\lim} \left\{ \int g(x, p_i(x)) \, dx \right\} \geqslant \int g(x, \bar{p}(x)) \, dx \tag{2.61}$$

$$\underline{\lim} \left\{ \int f(x, p_i(x) \, dx + c \int \varphi \circ p_i + \int a \right\} \geqslant \int f(x, \bar{p}(x)) \, dx + c \int \varphi \circ \bar{p} + \int a. \tag{2.62}$$

On subtracting the constant $\int a$ from both sides and on making use of (2.58), we obtain:

$$\underline{\lim} \int f(x, p_i(x) \, dx \geqslant \int f(x, \bar{p}(x)) \, dx. \tag{2.63}$$

But $-f$ is also a Carathéodory function, and satisfies (2.56). Applying (2.63) to it, we obtain:

$$\overline{\lim} \int f(x, p_i(x) \, dx \leqslant \int f(x, p(\bar{x})) \, dx \tag{2.64}$$

And we arrive at (2.59) by comparing (2.63) and (2.64). ∎

We deduce from this result a generalization of Theorem 1.2.

Proposition 2.11. *Let u_k, $1 \leqslant k \leqslant p$, be functions in $W_0^{1,\Phi}(\Omega)$ and α_k, $1 \leqslant k \leqslant p$, be positive scalars with sum unity. Let $\mathcal{F}$ be a finite family of Carathéodory functions which satisfy*

$$|f(x, \xi)| \leqslant a(x) + c\varphi(\xi), \qquad \text{where} \quad a \in L^1(\Omega) \quad \text{and} \quad c \geqslant 0. \tag{2.65}$$

For all $\varepsilon > 0$, there exists a function $u \in W^{1,\infty}(\Omega)$ such that:

$$\forall x \in \partial\Omega, \qquad u(x) = 0 \tag{2.66}$$

$$\left\| u - \sum_{k=1}^{p} \alpha_k u_k \right\|_1 \leqslant \varepsilon \tag{2.67}$$

$\forall f \in \mathcal{F}$,

$$\left| \int_\Omega f(x, \operatorname{grad} u(x)) \, dx - \sum_{k=1}^{p} \alpha_k \int_\Omega f(x, \operatorname{grad} u_k(x)) \, dx \right| \leqslant \varepsilon.$$

Proof. From Propositions 2.8 and 2.10, there exists for each u_k, $1 \leqslant k \leqslant p$, a piecewise affine function v_k such that:

$$\forall x \in \partial\Omega, \qquad v_k(x) = 0 \tag{2.68}$$

(2.69) $$\|u_k - v_k\|_1 \leqslant \varepsilon/2$$

(2.70) $\forall f \in \mathscr{F}$,

$$\left| \int_\Omega f(x, \operatorname{grad} u_k(x))\, dx - \int_\Omega f(x, \operatorname{grad} v_k(x))\, dx \right| \leqslant \varepsilon/2.$$

We now apply Proposition 1.3 to the family v_k, $1 \leqslant k \leqslant p$. First, we must verify hypothesis (1.57) which follows immediately from (2.65) and the continuity of φ:

(2.71) $$\max \{ |f(x, \xi)| \mid f \in \mathscr{F}, |\xi| \leqslant \rho \} \leqslant a(x) + \max_{|\xi| \leqslant \rho} \varphi(\xi).$$

There thus exists a function $u \in W^{1,\infty}(\Omega)$ such that:

(2.72) $$\forall x \in \partial\Omega, \qquad u(x) = \sum_{k=1}^{p} \alpha_k v_k(x)$$

(2.73) $$\left\| u - \sum_{k=1}^{p} \alpha_k v_k \right\|_\infty \leqslant \frac{\varepsilon}{2 \operatorname{meas} \Omega}$$

(2.74) $\forall f \in \mathscr{F}$,

$$\left| \int_\Omega f(x, \operatorname{grad} u(x))\, dx - \sum_{k=1}^{p} \alpha_k \int_\Omega f(x, \operatorname{grad} v_k(x))\, dx \right| \leqslant \varepsilon/2.$$

Whence we obtain the result: (2.66) results from (2.68) and (2.72), (2.67) from (2.69) and (2.73) and (2.68) from (2.70) and (2.74). ■

We have thus obtained a generalization of Proposition 1.3. We may ask ourselves which Carathéodory functions satisfy (2.56) or (2.65). If, for example, $\varphi(\xi) \geqslant |\xi|^\alpha$, all functions of the type $f(x,\xi) = \langle h(x), \xi \rangle$, with $h \in L_n^{\alpha'}(\Omega)$, $1/\alpha + 1/\alpha' = 1$, satisfy them by virtue of the inequality:

(2.75) $$\begin{cases} |\langle h(x), \xi \rangle| \leqslant \dfrac{1}{\alpha'} |h(x)|^{\alpha'} + \dfrac{1}{\alpha} |\xi|^\alpha & \text{if } \alpha > 1 \\ |\langle h(x), \xi \rangle| \leqslant \|h\|_\infty |\xi| & \text{if } \alpha = 1. \end{cases}$$

Whence we immediately have the following corollary which at first seems more precise than Proposition 2.11.

Corollary 2.12. *We assume that $\varphi(\xi) \geqslant |\xi|^\alpha$, where $\alpha \geqslant 1$. Let u_k, $1 \leqslant k \leqslant p$ be functions of $W_0^{1,\alpha}(\Omega)$, and α_k, $1 \leqslant k \leqslant p$, be positive real numbers with sum* *unity. Let $\mathscr{F}$ be a finite family of Carathéodory functions satisfying* (2.65). *For all $\varepsilon < 0$ and every neighbourhood V of zero for $\sigma(L_n^{\alpha}, L_n^{\alpha'})$, there exists a function $u \in W^{1,\infty}(\Omega)$ satisfying* (2.66), (2.67), (2.68) *and*:

$$\operatorname{grad} u - \operatorname{grad}\left(\sum_{k=1}^{p} \alpha_k u_k\right) \in V. \tag{2.76}$$

Proof. Let us define by means of a finite family $(h_i)_{i\in I}$ of $L_n^{\alpha'}$ a weak neighbourhood contained in V:

$$V \supset \left\{ q \in L_n^{\alpha} \,\middle|\, \left| \int h_i q - \sum_{k=1}^{p} \alpha_k \int h_i \operatorname{grad} u_k \right| \leqslant \varepsilon,\ i \in I \right\}. \tag{2.77}$$

From the foregoing remarks, for all $i \in I$ the integrand $h_i(x)\xi$ satisfies (2.65) for a suitable choice of a and c. We denote by $\mathscr{H}$ the family of $h_i(x)\xi$, for $i \in I$, and it only remains for us to apply Proposition 2.11 to $\mathscr{F} \cup \mathscr{H}$. ■

2.6. Functionals on $W^{1,\infty}(\Omega)$

Our aim is to obtain similar results for $W^{1,\infty}$, with a constraint of the type $\|\operatorname{grad} u\|_\infty \leqslant c$. We shall make a simplifying hypothesis, which is a kind of compatibility condition between the Lipschitz bounded open subset Ω, the boundary value and the constraint.

Proposition 2.13 *We take a function $u_0 \in W^{1,\infty}(\Omega)$ and a number c such that*

$$\|\operatorname{grad} u_0\|_\infty < c. \tag{2.78}$$

Let u_k, $1 \leqslant k \leqslant p$ be functions of $W^{1,\infty}(\Omega)$ satisfying the boundary condition $u_k = u_0$ over $\partial\Omega$, and the constraint $\|\operatorname{grad} u_k\| \leqslant c$ a.e. in Ω, and let $\alpha_k, 1 \leqslant k \leqslant p$ be positive numbers with sum unity. Let $\mathscr{F}$ be a finite family of Carathéodory functions satisfying

$$\max_{|\xi| \leqslant c} |f(x, \xi)| \leqslant a(x), \quad \text{where } a \in L^1(\Omega). \tag{2.79}$$

For all $\varepsilon > 0$ there exists a function u $W^{1,\infty}(\Omega)$ such that:

$$\forall x \in \partial\Omega, \qquad u(x) = u_0(x) \tag{2.80}$$

$$\left\| u - \sum_{k=1}^{p} \alpha_i u_i \right\|_\infty \leqslant \varepsilon \tag{2.81}$$

(2.82) $$\|\operatorname{grad} u\|_\infty \leqslant c$$

(2.83) $\forall f \in \mathscr{F}$,

$$\left| \int_\Omega f(x, \operatorname{grad} u(x))\, dx - \sum_{k=1}^{p} \alpha_k \int_\Omega f(x, \operatorname{grad} u_k(x))\, dx \right| \leqslant \varepsilon.$$

Proof. For λ tending to 0 in $]0,1[$, we have:

(2.84) $$\|\operatorname{grad}(\lambda u_0 + (1-\lambda)u_k)\|_\infty < c$$

(2.86) $$\|\operatorname{grad}(\lambda u_0 + (1-\lambda)u_k) - \operatorname{grad} u_k\|_\infty \to 0.$$

For all $\eta < 0$, we can thus choose functions v_k, $1 \leqslant k \leqslant p$, such that:

(2.87) $$\|u_k - v_k\|_\infty \leqslant \varepsilon/3$$

(2.88) $$\|\operatorname{grad} v_k\|_\infty < c$$

(2.89) $$\|\operatorname{grad} v_k - \operatorname{grad} u_k\|_\infty \leqslant \frac{\eta}{2 \operatorname{meas} \Omega}$$

(2.90) $$\forall x \in \partial\Omega, \qquad v_k(x) = u_0(x).$$

We therefore choose $\eta > 0$ sufficiently small so that, for all $f \in \mathscr{F}$ and $1 \leqslant k \leqslant p$:

(2.91) $$A \text{ measurable, meas } A \leqslant \eta \Rightarrow \int_\Omega a(x)\, dx \leqslant \varepsilon/4$$

(2.92)
$$\left.\begin{array}{l} \|p\|_\infty \leqslant c, \\ \|p - \operatorname{grad} u_k\|_1 \leqslant \eta \end{array}\right. \Rightarrow \int_\Omega |f(x, p(x)) - f(x, \operatorname{grad} u_k(x))|\, dx \leqslant \varepsilon/4.$$

We now apply Proposition 2.9 to the family v_k, $1 \leqslant k \leqslant p$. There exists an open space $\Delta \subset \Omega$ and functions $w_k \in W^{1,\infty}(\Omega)$, $1 \leqslant k \leqslant p$, such that:

(2.93) $$\operatorname{meas}(\Omega - \Delta) \leqslant \eta,$$

(2.94) the w_k are piecewise affine over Δ,

(2.95) $$\forall x \in \partial\Omega, \qquad w_k(x) = v_k(x) \quad \text{for} \quad 1 \leqslant k \leqslant p,$$

(2.96) $$\|w_k - v_k\|_\infty \leqslant \varepsilon/3 \quad \text{for} \quad 1 \leqslant k \leqslant p$$

(2.97) $$\|\operatorname{grad} w_k - \operatorname{grad} v_k\|_1 \leqslant \eta/2$$

(2.98) $$\|\text{grad}\, w_k\|_\infty \leqslant c.$$

Inequality (2.97) arises from (2.42), (2.43) and from the application of Lebesgue's theorem, (2.98) results from (2.43) and (2.88) (this is why we had to make hypothesis (2.78))

Let us now apply Proposition 1.3 to the family of w_k, $1 \leqslant k \leqslant p$, over the open subspace Δ. Hypothesis (1.57) is indeed satisfied, by virtue of (2.79). There therefore exists a function w which is locally Lipschitz on Δ such that:

(2.99) $$\forall f \in \mathscr{F},$$
$$\left| \int_\Delta f(x, \text{grad}\, w(x))\, \mathrm{d}x - \sum_{k=1}^{p} \alpha_k \int_\Delta f(x, \text{grad}\, w_k(x))\, \mathrm{d}x \right| \leqslant \varepsilon/4$$

(2.100) $$|\text{grad}\, w(x)| \leqslant c \quad \text{a.e. on } \Delta$$

(2.101) $$\left| w(x) - \sum_{k=1}^{p} \alpha_k w_k(x) \right| \leqslant \varepsilon/3 \qquad \forall x \in \Delta$$

(2.102) $$\forall x \in \partial\Delta, \qquad w(x) = \sum_{k=1}^{p} \alpha_k w_k(x).$$

We define a function u on Ω by:

(2.103) $$u(x) = w(x) \qquad \text{if } x \in \Delta$$

(2.104) $$u(x) = \sum_{k=1}^{p} \alpha_k w_k(x) \qquad \text{if } x \notin \Delta.$$

This is a continuous function by virtue of (2.102) and it is δ-Lipschitz with constant c, by virtue of (2.98) and (2.100), whence (2.82) Equation (2.80) results from (2.90), and estimate (2.81) results from (2.87), (2.96) and (2.101). It remains for us to establish (2.83). For all $f \in \mathscr{F}$, by virtue of (2.79) and (2.91), we have:

(2.105) $$\left| \int_{\Omega - \Delta} f(x, \text{grad}\, u(x))\, \mathrm{d}x \right| \leqslant \varepsilon/4$$

(2.106) $$\sum_{k=1}^{p} \alpha_k \left| \int_{\Omega - \Delta} f(x, \text{grad}\, u_k(x))\, \mathrm{d}x \right| \leqslant \varepsilon/4.$$

We now apply (2.92), by virtue of (2.89) and (2.97):

(2.107) $$\sum_{k=1}^{p} \alpha_k \int_\Delta |f(x, \text{grad}\, u_k(x)) - f(x, \text{grad}\, w_k(x))|\, \mathrm{d}x \leqslant \varepsilon/4.$$

We obtain (2.83) by adding (2.105), (2.99) (taking into account the fact that $w = u$ on Δ), (2.107) and (2.106). Whence the result. ∎

Here we must point out that all the functions f which are continuous on $\bar{\Omega} \times \mathbf{R}^n$ satisfy (2.79): it is sufficient to denote by m the maximum of $|f(x, \xi)|$ over the compact set $\bar{\Omega} \times \{\xi \mid |\xi| \leqslant c\}$, and to write:

(2.108) $$\max_{|\xi| \leqslant c} |f(x, \xi)| \leqslant m.$$

If $h \in L^1_n(\Omega)$, the function $f(x, \xi) = \langle h(x), \xi \rangle$ satisfies (2.79) trivially, whence we immediately have the following corollary, which at first seemes more precise than Proposition 2.13.

Corollary 2.14. *We use the hypotheses and notations of Proposition* 2.13. *For all* $\varepsilon > 0$ *and every neighbourhood* V *of zero for* $\sigma(L^\infty, L^1)$, *there exists a function* $u \in W^{1,\infty}(\Omega)$ *satisfying* (2.80), (2.81), (2.82), (2.83) *and*:

(2.109) $$\operatorname{grad} u - \operatorname{grad}\left(\sum_{k=1}^{p} \alpha_k u_k\right) \in V.$$

2.7. Γ-regularization

We still assume that Ω is a Lipschitz bounded open subspace of $\mathbf{R}^n$. By virtue of the above results, we shall now give certain situations where the Γ-regularization of the functional $u \rightsquigarrow \int_\Omega f(x, \operatorname{grad} u(x))$ coincides with its l.s.c. regularization, thus generalizing Proposition IX.1.2. Essentially we shall give two cases: firstly $W_0^{1,\alpha}$, $1 \leqslant \alpha \leqslant \infty$ and secondly $W^{1,\infty}$, under a common statement.

Let φ be a convex function of $\mathbf{R}^n$ into $\bar{\mathbf{R}}$, of one of the following types:

(H_1) $$\forall \xi \in \mathbf{R}^n, \qquad |\xi| \leqslant \varphi(\xi) < \infty$$

(H_∞) $$\varphi(\xi) = \delta(\xi \mid kB), \text{ where } k > 0 \text{ and } B = \{\xi \mid |\xi| \leqslant 1\}.$$

In the case of (H_1), φ certainly is non-negative and continuous. In the case of (H_∞), φ is the indicator function of a ball in $\mathbf{R}^n$. We take a Carathéodory function of $\Omega \times \mathbf{R}^n$ into $\mathbf{R}$, satisfying

(2.110) $$|f(x, \xi)| \leqslant a(x) + c\varphi(\xi), \text{ where } a \in L^1(\Omega) \text{ and } c \geqslant 0.$$

Finally we take a boundary condition in $W^{1,1}(\Omega)$

(H_1) $$u \in W_0^{1,1}(\Omega)$$

(H_∞) $u = u_0$ on $\partial\Omega$, where $u_0 \in W^{1,\infty}(\Omega)$ satisfies $\|\operatorname{grad} u_0\|_\infty < k$.

Let us define a mapping F of $W^{1,1}(\Omega)$ into $\mathbf{R}$ by:

$$
(2.111)\quad \left|\begin{array}{ll} F(u) = +\infty & \text{if} \displaystyle\int_\Omega \varphi \circ \operatorname{grad} u = +\infty \\ F(u) = +\infty & \text{if } u \text{ does not satisfy the boundary conditions,} \\ F(u) = \displaystyle\int_\Omega f(x, \operatorname{grad} u(x))\,\mathrm{d}x & \text{otherwise.} \end{array}\right.
$$

Finally, we put $W^{1,1}$ into duality with $L^\infty \times L^\infty_n$ by the usual bilinear form:

$$
(2.112)\qquad \langle u, (h, p) \rangle = \int_\Omega hu + \sum_{i=1}^n \int_\Omega p_i \frac{\partial u}{\partial x_i}.
$$

Proposition 2.15. *Hypotheses* (H_1) *or* (H_∞). *For the topology* $\sigma(W^{1,1}, L^\infty \times L^\infty_n)$ *the* Γ*-regularization of* F *coincides with its l.s.c. regularization*:

$$
(2.113)\qquad F^{**} = \bar{F}.
$$

Proof. It is necessary to show that $\operatorname{epi} F^{**} = \operatorname{epi} \bar{F}$, or again that:

$$
(2.114)\qquad \overline{\mathrm{co}}\ \operatorname{epi} F = \overline{\operatorname{epi} F}.
$$

We therefore take a point of co epi F, i.e. $(\sum_{k=1}^p \alpha_k u_k, \sum_{k=1}^p \alpha_k a_k)$, where the α_k are positive real numbers with sum unity and where:

$$
(2.115)\qquad F(u_k) < a_k < +\infty \quad \text{for} \quad 1 \leqslant k \leqslant p.
$$

We take $\varepsilon > 0$ and a weak neighbourhood V of the origin for the given duality. From Corollary 2.12 (for hypothesis (H_1)) or Corollary 2.14 (in the case of (H^∞)), there exists $u \in W^{1,1}(\Omega)$ such that:

$$
(2.116)\qquad u - \sum_{k=1}^p \alpha_k u_k \in V
$$

$$
(2.117)\qquad \left| F(u) - \sum_{k=1}^p \alpha_k F(u_k) \right| \leqslant \varepsilon.
$$

Substituting (2.115) into (2.117), we obtain:

$$
(2.118)\qquad \sum_{k=1}^p \alpha_k a_k + \varepsilon \geqslant F(u).
$$

But (2.116) and (2.118) indicate that:

$$\text{co epi } F \subset \overline{\text{epi } F} \tag{2.119}$$

Whence (2.114) and the result. ■

Corollaries 2.12 and 2.14 even enable us to state this result more precisely in the special cases $\varphi(\xi) = |\xi|^\alpha$, $1 \leqslant \alpha < \infty$, and (H^∞).

Corollary 2.16. *Let f be a Carathéodory function of $\Omega \times \mathbf{R}^n$* into $\mathbf{R}$, *which satisfies*

$$|f(x, \xi)| \leqslant a(x) + c\,|\xi|^\alpha \tag{2.120}$$

where $a \in L^1(\Omega)$, $c \geqslant 0$ and $1 < \alpha < \infty$. We then define a mapping F of $W_0^{1,\alpha}(\Omega)$ into $\mathbf{R}$ by:

$$F(u) = \int_\Omega f(x, \operatorname{grad} u(x))\,\mathrm{d}x. \tag{2.121}$$

For the duality $(W_0^{1,\alpha}, W^{-1,\alpha'})$, we have:

$$F^{**} = \bar{F}. \tag{2.122}$$

Corollary 2.17. *Let f be a Carathéodory function of $\Omega \times kB$ into $\mathbf{R}$, which satisfies*

$$\max_{|\xi| \leqslant k} |f(x, \xi)| \leqslant a(x) \tag{2.123}$$

where $a \in L^1(\Omega)$. We take $u_0 \in W^{1,\infty}(\Omega)$, such that

$$\|\operatorname{grad} u_0\|_\infty < k. \tag{2.124}$$

We then define a mapping F of $W^{1,\infty}(\Omega)$ into $\mathbf{R}\{+\infty\}$ by

$$\left|\begin{array}{ll} F(u) = +\infty & \textit{if } \; u \neq u_0 \textit{ on } \partial\Omega \; \textit{ or if } \; \|\operatorname{grad} u\|_\infty > k \\ F(u) = \displaystyle\int_\Omega f(x, \operatorname{grad} u(x))\,\mathrm{d}x & \textit{otherwise.} \end{array}\right. \tag{2.125}$$

For the duality $(W^{1,\infty}, L^1 \times L^1_n)$, we have:

$$F^{**} = \bar{F}. \tag{2.126}$$

3. APPLICATIONS TO THE CALCULUS OF VARIATIONS

We shall now apply the above results to the study of the integral $\int_\Omega f^{**}(x; \operatorname{grad} u(x))\,\mathrm{d}x$, which will enable us to calculate the Γ-regularization of the functional F and will provide us with a relaxation theorem for variational problems of the following type:

$$(\mathscr{P}) \qquad \left| \begin{array}{l} \operatorname{Inf} \displaystyle\int_\Omega f(x, u(x), \operatorname{grad} u(x))\,\mathrm{d}x \\ u = u_0 \quad \text{on } \partial\Omega. \end{array} \right.$$

3.1. A crucial lemma

Let Ω be a Lipschitz bounded open subspace of $\mathbf{R}^n$. We shall state our lemma under one of the following sets of hypotheses:

Hypotheses (H_1). We take a continuous convex mapping φ from $\mathbf{R}^n$ into $\mathbf{R}_+$, satisfying $\varphi(\xi) \geqslant |\xi|$, and a Carathéodory function f from $\Omega \times R^n$ into $\mathbf{R}$, satisfying

$$(3.1) \qquad 0 \leqslant f(x, \xi) \leqslant a(x) + c\varphi(\xi), \quad \text{where } a \in L^1(\Omega) \text{ and } c \geqslant 0.$$

We denote by $\mathscr{W} = W_0^{1,\varphi}(\Omega)$ the set of functions u of $W_0^{1,1}(\Omega)$ ("null on the boundary") such that $\int \varphi \circ \operatorname{grad} u < +\infty$. This is a convex set of $W^{1,1}(\Omega)$. We define a positive function F on $W^{1,1}$ by:

$$(3.2) \qquad \left| \begin{array}{ll} F(u) = +\infty & \text{if } u \notin \mathscr{W} \\ F(u) = \displaystyle\int_\Omega f(x, \operatorname{grad} u(x))\,\mathrm{d}x & \text{if } u \in \mathscr{W}. \end{array} \right.$$

Hypotheses (H_∞). We take a number $k > 0$ and a Carathéodory function f from $\Omega \times kB$ into $\mathbf{R}$, satisfying:

$$(3.3) \qquad 0 \leqslant f(x, \xi) \leqslant a(x) \qquad \forall \xi \in kB, \quad \text{where } a \in L^1(\Omega).$$

In addition we take $u_0 \in W^{1,\infty}(\Omega)$ such that:

$$(3.4) \qquad \|\operatorname{grad} u_0\|_\infty < k.$$

We denote by $\mathscr{W}$ the set of functions u of $W^{1,1}(\Omega)$ such that $u = u_0$ on $\partial\Omega$ and such that $\|\operatorname{grad} u\|_\infty \leqslant k$ a.e. in Ω. This is a closed convex subset of $W^{1,1}(\Omega)$.

We define a non-negative function F on $W^{1,1}$ by

$$(3.5)\quad \begin{cases} F(u) = +\infty & \text{if } u \notin \mathscr{W} \\ F(u) = \displaystyle\int_\Omega f(x, \operatorname{grad} u(x))\, dx & \text{if } u \in \mathscr{W}. \end{cases}$$

Lemma 3.1. *Hypotheses* (H_1) *or* (H_∞). *For $\rho > 0$, we denote by f_ρ the truncated function*:

$$(3.6)\quad \begin{cases} f_\rho(x, \xi) = +\infty & \text{if } |\xi| > \rho \\ f_\rho(x, \xi) = f(x, \xi) & \text{if } |\xi| \leqslant \rho. \end{cases}$$

Let us make ρ tend to $+\infty$. For all $x \in \Omega$ fixed, $f_\rho^{**}(x;.)$ is continuous on its effective domain and converges to $f^{**}(x;.)$ uniformly over all compact subsets. For all $\xi \in \mathbf{R}^n$, $f^{**}(.;\xi)$ converges to $f^{**}(.;\xi)$ in $L^1(\Omega)$.

Proof. In the case of (H_∞), $f_\rho^{**} = f_k^{**} = f^{**}$ if $\rho \geqslant k$. We can thus assume that $\rho < k$. We fix $x \in \Omega$. We know (from Lemma IX.3.3.) that for all $\xi \in \rho B$:

$$(3.7)\quad f_\rho^{**}(x;\xi) = \min\left[\sum_{i=1}^{n+1} \lambda_i f(x, \xi_i) \,\middle|\, \lambda \in E_{n+1}, \xi_i \in \rho B, \sum_{i=1}^{n+1} \lambda_i \xi_i = \xi\right].$$

If in particular $|\xi| = \rho$, ξ is extremal in ρB, and cannot be the centre of gravity of separate points. Thus, for all $\xi \in \partial(\rho B)$, $f(x;\xi) = f_\rho^{**}(x;\xi)$. We already know that $f^{**}(x,.)$ is finite and l.s.c. on ρB, and thus continuous in the interior. To check the continuity on the boundary, we take $\xi \in \partial(\rho B)$ and a sequence $\zeta_n \in \rho B$ converging to ξ.

Since f_ρ^{**} is l.s.c.:

$$\underline{\lim}\, f_\rho^{**}(x;\zeta_n) \geqslant f_\rho^{**}(x;\xi) = f(x,\xi)$$

and since f_ρ^{**} is less than the continuous function f on ρB:

$$\overline{\lim}\, f_\rho^{**}(x;\zeta_n) \leqslant \lim f(x,\zeta_n) = f(x,\xi).$$

The continuity of $f_\rho^{**}(x;.)$ is thus established.

If $\sigma > \rho$, it is obvious that $f_\rho^{**} \geqslant f^{**} \geqslant f^{**}$. The pointwise infimum $\inf f_\rho^{**}$ is therefore a convex function, which is continuous over the domain of f^{**}, everywhere less than f and greater than f^{**}. This can only be f^{**}, which is thus the decreasing pointwise limit of the f_ρ^{**} as $\rho \to \infty$. Hence we deduce the first part of the lemma by virtue of Dini's theorem and the second part from Lebesgue's theorem (by the inequality (3.1)). ■

Proposition 3.2. *Under hypotheses* (H_1) *or* (H_∞), *for each piecewise affine function* $u \in \mathscr{W}$, *for all* $\varepsilon > 0$ *and every neighbourhood* V *of* 0 *for* $\sigma(L^1_n, L^\infty_n)$, *there exists a function* $v \in \mathscr{W}$ *such that:*

(3.8) $$\operatorname{grad} v - \operatorname{grad} u \in V,$$

(3.9) $$\|v - u\|_\infty \leqslant \varepsilon,$$

(3.10) $$\left| \int_\Omega f(x, \operatorname{grad} v(x))\, dx - \int_\Omega f^{**}(x; \operatorname{grad} u(x)\, dx \right| \leqslant \varepsilon.$$

Proof. We take $u \in \mathscr{W}$, piecewise affine, $\varepsilon > 0$, and a neighbourhood V of zero for $\sigma(L^1_n, L^\infty_n)$ of the form:

(3.11)
$$V = \left\{ p \in L^1_n \;\middle\|\; \left| \int_\Omega \langle h_m(x), p(x) \rangle\, dx \right| \leqslant \eta, m \in M \right\}$$

where η is a positive number and the h_m, $m \in M$, are a finite family of functions of L^∞_n.

By hypothesis, there exists a partition of Ω into a negligible set N and open subspaces Δ_i, $1 \leqslant i \leqslant p$, over which $\operatorname{grad} u$ is constant. By virtue of Lemma 3.1, we can choose a number $\rho \geqslant \|\operatorname{grad} u\|_\infty$ such that:

(3.12) $$\int_\Omega |f^{**}(x; \operatorname{grad} u) - f^{**}_\rho(x; \operatorname{grad} u)|\, dx \leqslant \varepsilon/6.$$

When (H_∞) applies, we put $\rho = k$. Under these conditions, f^{**}_l is a Carathéodory function which is non-negative over $\Omega \times \rho B$, and for each i there exists a compact subset $K_i \subset \Delta_i$ such that

(3.13) $$\int_{\Delta_i - K_i} [a(x) + c \max_{|\xi| \leqslant \rho} \varphi(\xi)]\, dx \leqslant \varepsilon/6p.$$

(3.14) The restrictions of f and f^{**}_l to $K_i \times \rho B$ are continuous.

By applying (3.14) and the compactness of ρB, for all $x \in K_i$ we can find an open ball ω_x with centre x contained in Ω such that:

(3.15)
$$\forall y \in \omega_x \cap K_i, \qquad |f^{**}_\rho(y, \operatorname{grad} u) - f^{**}_\rho(x, \operatorname{grad} u)| \leqslant \frac{\varepsilon}{6 \operatorname{meas} \Omega}$$

(3.16) $\forall y \in \omega_x \cap K_i, \ \forall \xi \in \rho B,' \qquad |f(y, \xi) - f(x, \xi)| \leqslant \dfrac{\varepsilon}{6 \operatorname{meas} \Omega}.$

From Lemma IX.3.3, for all $x \in \Delta_i$ fixed we can find positive numbers α_k, $1 \leqslant k \leqslant n+2$ with sum unity, and points $\xi_k \in \rho B$, $1 \leqslant k \leqslant n+1$, such that

$$\sum_{k=1}^{n+1} \alpha_k \xi_k = \operatorname{grad} u \tag{3.17}$$

$$\sum_{k=1}^{n+1} \alpha_k f(x, \xi_k) = f_\rho^{**}(x; \operatorname{grad} u). \tag{3.18}$$

And (3.15) and (3.16) then become:

(3.19)

$$\forall y \in \omega_x \cap K_i, \qquad \left| f_\rho^{**}(y, \operatorname{grad} u) - \sum_{k=1}^{n+1} \alpha_k f(y, \xi_k) \right| \leqslant \frac{2\varepsilon}{6 \operatorname{meas} \Omega}.$$

But we can cover the compact K_i by a finite number of those open balls ω_x, denoted by ω_j, $1 \leqslant j \leqslant l$. The open subspace $\omega'_j = \omega_j - \bigcup_{i=1}^{j-1} \bar{\omega}_i$ have a boundary $\partial \omega'_j$, of null measure, and $\bigcup_{j=1}^{l} \bar{\omega}_j = \bigcup_{j=1}^{l} \omega_j$. We take a family u_k, $1 \leqslant k \leqslant n+1$, such that $\operatorname{grad} u_k = \xi_k$ and $u = \sum_{k=1}^{n+1} \alpha_k u_k$, and we apply Proposition 1.3 to each of the open spaces ω'_j.

This is possible by virtue of (3.17), and this gives us a function $v_i \in W^{1,\infty}(\omega'_j)$ such that:

$$\left| \int_{\omega'_j} f(x, \operatorname{grad} v_i(x)) \, dx - \sum_{k=1}^{n+1} \alpha_k \int_{\omega'_j} f(x, \xi_k) \, dx \right| \leqslant \frac{\varepsilon}{6 p \ell} \tag{3.20}$$

$$\forall x \in \omega'_j, \qquad |\operatorname{grad} v_i(x)| \leqslant \rho \tag{3.21}$$

$$\forall x \in \omega'_j, \qquad |v_i(x) - u(x)| \leqslant \varepsilon \tag{3.22}$$

$$\forall x \in \partial \omega'_j, \qquad v_i(x) = u(x) \tag{3.23}$$

(3.24) $\forall m \in M$,

$$\left| \int_{\omega'_j} \langle h_m(x), \operatorname{grad} v_i(x) \rangle \, dx - \int_{\omega'_j} \langle h_m(x), \operatorname{grad} u(x) \rangle \, dx \right| \leqslant \eta / p \ell.$$

Finally we set:

$$v_i = u \qquad \text{on } \Delta_i - \bigcup_{j=1}^{l} \omega_j. \tag{3.25}$$

The function v_i is thus, by virtue of (3.23) a continuous function over Δ_i, coinciding with u over $\partial\Delta_i$. We can then define a continuous mapping v on Ω by:

(3.26)
$$\begin{vmatrix} v(x) = v_i(x) & \text{if } x \in \Delta_i \\ v(x) = u(x) & \text{if } x \in N. \end{vmatrix}$$

We then have $v = u$ over $\partial\Omega$, and $\|\operatorname{grad} v\|_\infty \leqslant \rho$ by (3.21) and (3.25). Hence $v \in \mathscr{W}$. From (3.22) and (3.25) we deduce (3.9) and from (3.24) and (3.25) we deduce (3.8). It remains for us to establish (3.10).

From the inequalities (3.1) or (3.3) and from condition (3.13) we deduce that, for $1 \leqslant i \leqslant p$:

(3.27)
$$\int_{\Delta_i - K_i} |f^{**}(x, \operatorname{grad} u)|\,\mathrm{d}x \leqslant \varepsilon/6p$$

(3.28)
$$\int_{\Delta_i - K_i} |f(x, \operatorname{grad} v(x))|\,\mathrm{d}x \leqslant \varepsilon/6p$$

Substituting (3.19) into (3.20), we obtain

(3.29)
$$\left| \int_{K_i \cap \omega'_j} f(x, \operatorname{grad} v(x))\,\mathrm{d}x - \int_{K_i \cap \omega'_j} f^{**}_\rho(x, \operatorname{grad} u(x))\,\mathrm{d}x \right| \leqslant \frac{\varepsilon}{6p} + \frac{\varepsilon \operatorname{meas} \omega'_j}{3 \operatorname{meas} \Omega}.$$

We sum this inequality from $j = 1$ to ℓ, taking into account the fact that the ω'_j cover K_i up to a negligible set:

(3.30)
$$\left| \int_{K_i} f(x, \operatorname{grad} v(x))\,\mathrm{d}x - \int_{K_i} f^{**}_\rho(x, \operatorname{grad} u(x))\,\mathrm{d}x \right| \leqslant \frac{\varepsilon}{6p} + \frac{\varepsilon \operatorname{meas} K_i}{3 \operatorname{meas} \Omega}.$$

On adding all the inequalities (3.27), (3.28), (3.30) from $i = 1$ to p and the inequality (3.12) we obtain the desired equation (3.10), and hence the result. ■

The central lemma of this section follows:

Theorem 3.3. *Under the hypotheses* (H_1) *or* (H_∞), *for all functions* $u \in \mathscr{W}$, *for all* $\varepsilon > 0$ *and all neighbourhoods* V *of zero in* $\sigma(L^1_n, L^\infty_n)$, *there exists a function* $v \in \mathscr{W}$ *such that*:

(3.31)
$$\operatorname{grad} v - \operatorname{grad} u \in V$$

$$(3.32)\qquad \left|\begin{array}{ll}(\mathrm{H}_1) & \left|\int \varphi \circ \operatorname{grad} u - \int \varphi \circ \operatorname{grad} v\right| \leqslant \varepsilon \\ (\mathrm{H}_\infty) & \|u - v\|_\infty \leqslant \varepsilon\end{array}\right.$$

$$(3.33)\qquad \left|\int_\Omega f(x, \operatorname{grad} v(x))\, dx - \int_\Omega f^{**}(x; \operatorname{grad} u(x)\, dx\right| \leqslant \varepsilon.$$

Proof. The case (H_1) can be treated by approximating u by a function w which is piecewise affine and null on the boundary such that (Proposition 2.8):

$$(3.34)\qquad \operatorname{grad} w - \operatorname{grad} v \in \tfrac{1}{2} V$$

$$(3.35)\qquad \left|\int_\Omega \varphi \circ \operatorname{grad} w - \int_\Omega \varphi \circ \operatorname{grad} u\right| \leqslant \varepsilon/2$$

$$(3.36)\qquad \left|\int_\Omega f^{**}(x; \operatorname{grad} w(x))\, dx - \int_\Omega f^{**}(x; \operatorname{grad} u(x))\, dx\right| \leqslant \varepsilon/2.$$

We obtain (3.36) by applying Proposition 2.10 to f^{**}, which is a Carathéodory function and which satisfies inequality (3.1). It only remains to apply Proposition 3.2 to w in such a way as to obtain a function $v \in \mathscr{W}$ such that:

$$(3.37)\qquad \operatorname{grad} v - \operatorname{grad} w \in \tfrac{1}{2} V$$

$$(3.38)\qquad \left|\int_\Omega \varphi \circ \operatorname{grad} v - \int_\Omega \varphi \circ \operatorname{grad} w\right| \leqslant \varepsilon/2$$

$$(3.39)\qquad \left|\int_\Omega f^{**}(x; \operatorname{grad} w(x))\, dx - \int_\Omega f(x, \operatorname{grad} v(x))\, dx\right| \leqslant \varepsilon/2.$$

The function v satisfies (3.31), (3.32) and (3.33) whence the result. The case (H_∞) is rather more difficult. We first note that if a sequence $(p_n)_{n\in\mathbf{N}}$ tends to p almost everywhere, with $\|p_n\|_\infty \leqslant k$ for all $n \in \mathbf{N}$, then:

$$(3.40)\qquad \int_\Omega |f^{**}(x; p_n(x)) - f^{**}(x; p(x))|\, dx \to 0.$$

Indeed the function f^{**} is of Carathéodory type (Lemma 3.1), hence the integrands converge almost everywhere and we can apply Lebesgue's theorem by virtue of the inequality (3.3).

We can always assume that $\|\operatorname{grad} u\|_\infty < k$. In fact, if this is not so, it is sufficient to approximate u by the functions $u + \lambda(u - u_0)$, $\lambda \to 0_+$, which are of the required type by virtue of (3.4).

But if $\|\operatorname{grad} u\|_\infty < k$, from Proposition 2.9 there exists a function $w \in \mathscr{W}$ and a Lipschitz open subset $\mathcal{O} \subset \Omega$ such that:

$$\int_{\Omega - \mathcal{O}} |a(x)| \, \mathrm{d}x \leqslant \varepsilon/4$$

$$w \text{ is piecewise affine over } \mathcal{O}$$

$$\|u - w\|_\infty \leqslant \varepsilon/2$$

$$\int_\Omega |f^{**}(x; \operatorname{grad} w(x)) - f^{**}(x; \operatorname{grad} u(x))| \, \mathrm{d}x \leqslant \varepsilon/4$$

$$\operatorname{grad} w - \operatorname{grad} v \in \tfrac{1}{2} V.$$

It only remains to apply Proposition 3.2 to the piecewise affine function w over $\mathcal{O}$: we obtain a function $v \in W^{1,\infty}(\mathcal{O})$ such that:

$$\forall x \in \partial\mathcal{O} \qquad v(x) = w(x) \tag{3.41}$$

$$\forall x \in \partial\mathcal{O}, \qquad |\operatorname{grad} v(x)| \leqslant k \tag{3.42}$$

$$\forall x \in \partial\mathcal{O}, \qquad |v(x) - w(x)| \leqslant \varepsilon/2 \tag{3.43}$$

$$\left| \int_\mathcal{O} f(x, \operatorname{grad} v(x)) \, \mathrm{d}x - \int_\mathcal{O} f^{**}(x; \operatorname{grad} w(x)) \, \mathrm{d}x \right| \leqslant \varepsilon/4 \tag{3.44}$$

$$\operatorname{grad} v - \operatorname{grad} w \in \tfrac{1}{2} V. \tag{3.45}$$

We can extend v by w outside $\mathcal{O}$: the function thus obtained belongs to $\mathscr{W}$ by virtue of (3.41) and (3.42) and satisfies (3.31), (3.32) and (3.33). ■

3.2. Calculation of F^{**} and of F^*

First of all, we have the following direct consequence:

Proposition 3.4. *Under hypotheses* (H_1) *or* (H_∞), *the Γ-regularization of F in the duality* $(W^{1,1}, L^\infty \times L^\infty_n)$ *satisfies*:

$$\forall u \in \mathscr{W} \qquad F^{**}(u) = \int_\Omega f^{**}(x; \operatorname{grad} u(x)) \, \mathrm{d}x. \tag{3.46}$$

Proof. The mapping $u \to \int f^{**}(x; \operatorname{grad} u(x)) \mathrm{d}x$ is convex and l.s.c. over $W^{1,1}$, is everywhere less than F and hence than F^{**}:

$$(3.47) \qquad \forall u \in W^{1,1}, \qquad \int_\Omega f^{**}(x; \operatorname{grad} u(x))\, dx \leqslant F^{**}(u).$$

As usual, we denote by $\bar{F}$ the l.s.c. regularization. From Theorem 3.3 we have:

$$(3.48) \qquad \forall u \in \mathscr{W}, \qquad \int_\Omega f^{**}(x, \operatorname{grad} u(x))\, dx \geqslant \bar{F}(u).$$

But $\bar{F} \geqslant F^{**}$; (3.47) and (3.48) then give us the required equality. ■

As usual, we can apply this theorem to the case where $\varphi(\xi) = |\xi|^\alpha$, $1 \leqslant \alpha < \infty$.

Corollary 3.5. *Let f be a Carathéodory function over $\Omega \times \mathbf{R}^n$ which satisfies:*

$$(3.49) \qquad 0 \leqslant f(x, \xi) \leqslant a(x) + c\,|\xi|^\alpha$$

where $a \in L^1(\Omega)$, $c \geqslant 0$, and $\alpha > 1$. We define a function F from $W_0^{1,\alpha}(\Omega)$ into $\mathbf{R}$ by:

$$(3.50) \qquad F(u) = \int_\Omega f(x, \operatorname{grad} u(x))\, dx.$$

In the duality $(W_0^{1,\alpha}, W^{-1,\alpha'})$, where $1/\alpha + 1/\alpha' = 1$, we have:

$$(3.51) \qquad F^{**}(u) = \int_\Omega f^{**}(x; \operatorname{grad} u(x))\, dx$$

$$(3.52) \quad F^*(u^*) = \min\left[\int_\Omega f^*(x; p^*(x))\, dx \mid p^* \in L_n^{\alpha'},\ -\operatorname{div} p^* = u^*\right].$$

Indeed, we obtain (3.52) by writing that $F^{**}(u) = \tilde{F} \circ A$ where $\tilde{F}$ is the mapping $p \rightsquigarrow \int_\Omega f^{**}(x; p(x))\, dx$ from L_n^α into $\mathbf{R}$ and A is the mapping $u \rightsquigarrow \operatorname{grad} u$ from $W_0^{1,\alpha}$ into L_n^α. Now, $\tilde{F}$ is a continuous convex function, whose polar we calculated in Chapter IX, Section 2, and it only remains to apply Proposition I.5.7.

3.3. Relaxation

We are now in a position to state a relaxation theorem as a result of Theorem 3.3.

Proposition 3.6. *Under hypotheses* (H_1) *or* $(\mathrm{H})_{,\infty}$ *the problems*

$$(\mathscr{P}) \qquad Inf \int_\Omega f(x, \operatorname{grad} u(x))\, dx \qquad for \quad u \in \mathscr{W}$$

$$(\mathscr{P}\mathscr{R}) \qquad \textit{Inf} \int_\Omega f^{**}(x; \operatorname{grad} u(x))\, dx \qquad \textit{for} \quad u \in \mathscr{W}$$

have the same value

$$\operatorname{Inf}(\mathscr{P}) = \operatorname{Inf}(\mathscr{P}\mathscr{R}).$$

We note that, in the case (H_1), we can no longer assert that the problem $(\mathscr{P}\mathscr{R})$ possesses solutions. For this to be so, we need further hypotheses; we shall now state a more precise theorem with hypotheses which depend on parameters $\alpha(\alpha = 1 \text{ or } +\infty)$ and $\beta(1 \leqslant \beta \leqslant \infty)$.

As always, Ω will be a Lipschitz bounded open subspace of $\mathbf{R}^n$. We take a function φ over $\mathbf{R}^n$ satisfying one of the following conditions:

(3.53) $(\alpha = 1)$, φ is a convex mapping from $\mathbf{R}^n$ into $\mathbf{R}_+$, hence non-negative and continuous, such that $\varphi(\xi)/|\xi| \to \infty$ when $|\xi| \to \infty$
$(\alpha = \infty)\varphi$ is the indicator function χ_{kB} of kB.

Let f be a mapping from $\Omega \times (\mathbf{R} \times \mathbf{R}^n)$ into $\overline{\mathbf{R}}$, such that:

(3.54) $(1 \leqslant \beta < \infty)$, there exists a_1 and $a_2 \in L^1(\Omega)$, $b \geqslant 0$ and $c \geqslant 1$ such that $a_2(x) + \varphi(\xi) \leqslant f(x,s,\xi) \leqslant a_1(x) + b|s|^\beta + c\varphi(\xi)$
$(\beta = \infty)$, there exists $a_2 \in L^1(\Omega)$ and, for all $\rho > 0$, there exists $a_1 \in L^1(\Omega)$ and $c \geqslant 1$ such that:
$a_2(x) + \varphi(\xi) \leqslant f(x,s,\xi) \leqslant a_1(x) + c\varphi(\xi)$ for $|s| \leqslant \rho$

(3.55) f is a Carathéodory function over its effective domain.

We note that, from the estimates (3.54), this effective domain is $\Omega \times (\mathbf{R} \times \mathbf{R}^n)$ if $\alpha = 1$ and $\Omega \times (\mathbf{R} \times kB)$ if $\alpha = \infty$.

We take $u_0 \in W^{1,1}(\Omega)$

(3.56)
$$\begin{array}{ll} (\alpha = 1) & u_0 = 0 \\ (\alpha = \infty) & \|\operatorname{grad} u_0\|_\infty < k. \end{array}$$

Theorem 3.7. *Hypotheses* (3.53) *to* (3.56). *If* $\alpha = 1$ *we assume that if a sequence* $(u_n)_{n\in\mathbf{N}}$ *in* $W_0^{1,\varphi}$ *converges weakly to* $\bar{u}$ *in* $W_0^{1,1}$ *and if* $\int \varphi \circ \operatorname{grad} n_n$ *converges to* $\int \varphi$ *grad u, then* u_n *converges to* $\bar{u}$ *in* L^β. *Under these conditions, the problems:*

$$(\mathscr{P}) \qquad \begin{array}{|l} \textit{Inf} \displaystyle\int_\Omega f(x, v(x), \operatorname{grad} u(x)\, dx \\ u \in u_0 + W_0^{1,1} \end{array}$$

$$(\mathscr{P}\mathscr{R}) \qquad \left| \begin{array}{l} Inf \displaystyle\int_\Omega f^{**}(x, u(x); \operatorname{grad} u(x))\, dx \\ u \in u_0 + W_0^{1,1} \end{array} \right.$$

have the same value:

$$\operatorname{Inf}(\mathscr{P}) = \operatorname{Min}(\mathscr{P}\mathscr{R}).$$

The problem $(\mathscr{P}\mathscr{R})$ *possesses optimal solutions. If* $\bar{u}$ *is a solution of* $(\mathscr{P}\mathscr{R})$, *there exists a minimizing sequence* $(u_n)_{n\in N}$ *of* $(\mathscr{P})$ *such that* $u_u \to \bar{u}$ *in* L^β *and that* $\operatorname{grad} u_n \to \operatorname{grad} u$ *in the* $\sigma(L^1, L^\infty)$ *topology. If* $(u_n)_{n\in N}$ *is a minimizing sequence of* $(\mathscr{P})$, *there exists an optimal solution* $\bar{u}$ *of* $(\mathscr{P}\mathscr{R})$ *and an extracted subsequence* $(u_{n_k})_{k\in N}$ *which converges to* $\bar{u}$ *in* L^β, *while* $(\operatorname{grad} u_{n_k})_{k\in N}$ *converges to* $\operatorname{grad}\bar{u}$ *in the* $\sigma(L^1, L^\infty)$ *topology.*

Proof. By virtue of the coerciveness (3.53), and Proposition VIII.2.5 the problem $(\mathscr{P}\mathscr{R})$ possesses solutions. We can thus proceed as in Section 4, Chapter IX.

Let $\bar{u}$ be a solution of $(\mathscr{P}\mathscr{R})$. From the estimates (3.54), we have $\int_\Omega \varphi \circ \operatorname{grad}\bar{u} < +\infty$.

There exists, by Lemma IX.4.2, an increasing convex function $\psi : [0, +\infty[\to [1, +\infty[$ such that:

$$\lim_{t\to\infty} \frac{\psi(t)}{t} = +\infty$$

$$(3.57) \qquad \int_\Omega \psi \circ \varphi \circ \operatorname{grad}\bar{u} < +\infty.$$

From the estimates (3.54) we deduce that:

$$(\beta < \infty), \qquad 0 \leqslant f(x, \bar{u}(x), \xi) - a_2(x) \leqslant a_1(x) - a_2(x) + b\,|\bar{u}(x)|^\beta + c\,\psi \circ \varphi(\xi)$$

$$(\beta = \infty), \qquad 0 \leqslant f(x, \bar{u}(x), \xi) - a_2(x) \leqslant a_1(x) - a_2(x) + c\,\psi \circ \varphi(\xi) \quad \text{for } |\xi| \leqslant k.$$

We can thus apply Theorem 3.3 to the integrand $f(x, \bar{u}(x), \xi) - a_2(x)$ and to the function $u \in W^{1, \psi\circ\varphi}$. There thus exists a sequence $(u_i)_{i\in N}$ in $W^{1,1}(\Omega)$ such that as $i \to \infty$:

$$(3.58) \qquad \operatorname{grad} u_i \to \operatorname{grad} u \qquad \text{for} \quad \sigma(L^1, L^\infty)$$

(3.59)
$$\left|\begin{array}{ll}(\alpha = 1) & \int \psi \circ \varphi \circ \operatorname{grad} u_i \to \int \psi \circ \varphi \circ \operatorname{grad} \bar{u} \\ (\alpha = \infty) & \|u_i - u\|_\infty \to 0\end{array}\right.$$

(3.60)
$$\left| \int_\Omega f(x, \bar{u}(x), \operatorname{grad} u_i(x))\, dx - \int_\Omega f^{**}(x, \bar{u}(x); \operatorname{grad} \bar{u}(x))\, dx \right| \to 0.$$

If $\alpha = 1$, by (3.59), as $\psi \geqslant 1$,

$$\int \varphi \circ \operatorname{grad} u_i \to \int \varphi \circ \operatorname{grad} u,$$

and from the stated hypothesis, u_i converges to u in L^β. If $\alpha = \infty$, this result comes directly from (3.59). Finally, by possible extraction of a subsequence, we have:

(3.61) $\qquad u_i \to u \quad$ in L^β and almost everywhere.

Taking $\rho > 0$ we set:

(3.62) $$h_i(x) = |f(x, u_i(x), \operatorname{grad} u_i(x)) - f(x, \bar{u}(x), \operatorname{grad} u_i(x))|$$

(3.63) $$A^i_\rho = \{ x \in \Omega \,|\, |\varphi \circ \operatorname{grad} u_i(x)| \leqslant \rho \}$$

(3.64) $$B^i_\rho = \{ x \in \Omega \,|\, |\varphi \circ \operatorname{grad} u_i(x)| > \rho \}.$$

From the estimates given by (3.54), we have:

$$\begin{array}{ll}(\beta < \infty) & |h_i(x)| \leqslant |a_1(x) - a_2(x)| + b(|\bar{u}(x)|^\beta \\ & \qquad + |u_i(x)|^\beta) + (c-1)\varphi \circ \operatorname{grad} u_i(x) \\ (\beta = \infty) & |h_i(x)| \leqslant |a_1(x) - a_2(x)| + (c-1)\varphi \circ \operatorname{grad} u_i(x).\end{array}$$

From Theorem VIII.1.3, the conditions (3.59) imply that the family of mappings $x \mapsto \varphi \circ \operatorname{grad} u_i(x)$, for $i \in \mathbf{N}$, is equi-integrable. If $\beta < \infty$, the family $|u_i(x)|^\beta$ is equi-integrable by virtue of (3.61). The family of h_i, $i \in \mathbf{N}$, is thus equi-integrable. Theorem VIII.1.3 further serves to show that meas $B^i_\rho \to 0$ uniformly for $i \in \mathbf{N}$ as $\rho \to \infty$ and we can therefore choose $\rho > 0$ sufficiently large so that:

(3.65) $$\int_{B^i_\rho} h_i \leqslant \varepsilon/3 \qquad \forall i \in \mathbf{N}.$$

We now examine the sequence of functions $h_i 1_{A^i_\rho}$ for $i \in \mathbf{N}$. For almost all $x \in \Omega$, $u_i(x) \to \bar{u}(x)$, and hence $f(x, u_i(x), \xi) \to f(x, \bar{u}(x), \xi)$ uniformly for ξ

belonging to the compact set $\{\xi | \varphi(\xi) \leqslant \rho\}$. The functions $h_i \mathbf{1}_{A^i_\rho}$ thus converge to zero almost everywhere. Since they form an equi-integrable family, they converge to zero in L^1. We can therefore choose $i_0 \in \mathbf{N}$ sufficiently large so that:

$$\int_{A^i_\rho} h_i \leqslant \varepsilon/3 \qquad \forall i \geqslant i_0. \tag{3.66}$$

By adding (3.65) and (3.66) we obtain:

$$\int h_i \leqslant 2\varepsilon/3 \qquad \forall i \geqslant i_0. \tag{3.67}$$

Substituting (3.62) into this expression, we obtain for $i \geqslant i_0$

$$\int_\Omega |f(x, u_i(x), \text{grad } u_i(x)) - f(x, \bar{u}(x), \text{grad } u_i(x))| \,\mathrm{d}x \leqslant 2\varepsilon/3 \tag{3.68}$$

and by virtue of (3.60), for $i \geqslant i_i$:

(3.69)

$$\left| \int_\Omega f(x, \bar{u}(x), \text{grad } u_i(x)) \,\mathrm{d}x - \int_\Omega f^{**}(x, \bar{u}(x), \text{grad } \bar{u}(x)) \,\mathrm{d}x \right| \leqslant \varepsilon/3.$$

By adding together (3.68) and (3.69), we obtain for $i \geqslant \max(i_0, i_1)$

$$\left| \int_\Omega f(x, u_i(x), \text{grad } u_i(x)) \,\mathrm{d}x - \min(\mathscr{P}\mathscr{R}) \right| \leqslant \varepsilon. \tag{3.70}$$

Thus $\inf(\mathscr{P}) = \min(\mathscr{P}\mathscr{R})$ and the theorem results. ∎

In particular, we can use this method to deal with non-homogeneous problems (with non-zero boundary conditions). We give an important example in the case where $\varphi(\xi) = |\xi|^\alpha$, $1 \leqslant \alpha < \infty$.

Corollary 3.8. *Let f be a Carathéodory function from $\Omega \times (\mathbf{R} \times \mathbf{R}^n)$ into $\mathbf{R}$ which satisfies*

$$a_2(x) + c_2 |\xi|^\alpha \leqslant f(x, s, \xi) \leqslant a_1(x) + b|s|^\alpha + c_1 |\xi|^\alpha \tag{3.71}$$

where a_1 and a_2 $L^1(\Omega)$, $1 < \alpha < \infty$, $b \geqslant 0$ and $c_1 \geqslant c_2 > 0$. Let $u_0 \in W^{1,\alpha}(\Omega)$. The problems:

$$(\mathscr{P}) \qquad \left| \begin{array}{l} Inf \displaystyle\int_\Omega f(x, u(x), \text{grad } u(x)) \,\mathrm{d}x \\ u - u_0 \in W^{1,\alpha}_0 \end{array} \right.$$

$$(\mathscr{PR}) \qquad \left| \begin{array}{l} \mathit{Inf} \displaystyle\int_\Omega f^{**}(x; u(x), \operatorname{grad} u(x)) \, dx \\ u - u_0 \in W_0^{1,\alpha} \end{array} \right.$$

have the same value:

$$\operatorname{Inf}(\mathscr{P}) = \operatorname{Min}(\mathscr{PR}).$$

The problem $(\mathscr{PR})$ *possesses solutions: these are the weak cluster points in* $W^{1,\alpha}$ *of the minimizing sequences of problem* $(\mathscr{P})$.

Proof. Let us introduce the integrand:

$$g(x, s, \xi) = f(x, s + u_0(x), \xi + \operatorname{grad} u_0(x)) \tag{3.72}$$

and apply Theorem 3.7 to the problem:

$$(\mathscr{Q}) \qquad \left| \begin{array}{l} \operatorname{Inf} \displaystyle\int_\Omega g(x, v(x), \operatorname{grad} v(x)) \, dx \\ v \in W_0^{1,1}. \end{array} \right.$$

Indeed, g is a Carathéodory function over $\Omega \times (\mathbf{R} \times \mathbf{R}^n)$, and we have from convexity that:

$$\begin{aligned} g(x, s, \xi) &\geqslant a_2(x) + c_2 |\xi + \operatorname{grad} u_0(x)|^\alpha \\ &\geqslant (a_2(x) - c_2 |\operatorname{grad} u_0(x)|^\alpha) + \frac{c_2}{2^{\alpha-1}} |\xi|^\alpha \end{aligned} \tag{3.73}$$

$$\begin{aligned} g(x, s, \xi) &\leqslant a_1(x) + b |s + u_0(x)|^\alpha + c_1 |\xi + \operatorname{grad} u_0(x)|^\alpha \\ &\leqslant a_1(x) + 2^{\alpha-1}(b |u_0(x)|^\alpha + c_1 |\operatorname{grad} u_0(x)|^\alpha) \\ &\quad + 2^{\alpha-1} b |s|^\alpha + 2^{\alpha-1} c_1 |\xi|^\alpha. \end{aligned} \tag{3.74}$$

The integrand g has all the required properties. We now recall that the injection of $W^{1,\alpha}$ into L^α is compact. If therefore v_i is a sequence in $W_0^{1,1}$, the sequence of the grad v_i being bounded in L^α and converging to grad $\bar{v}$ for $\sigma(L^\infty, L^\alpha)$ (which coincides with $\sigma(L^\alpha, L^{\alpha'})$ on bounded subsets), then v_i converges to $\bar{v}$ in L^α. All the conditions of Theorem 3.7 are satisfied and thus $\operatorname{Inf}(\mathscr{Q}) = \min(\mathscr{QR})$, where:

$$(\mathscr{QR}) \qquad \left| \begin{array}{l} \operatorname{Inf} \displaystyle\int_\Omega g^{**}(x, v(x); \operatorname{grad} v(x)) \, dx \\ v \in W_0^{1,1}. \end{array} \right.$$

But clearly, $g^{**}(x,s;\xi) = f^{**}(x,s+u_0(x);\xi+\operatorname{grad}u_0(x))$. Hence:

$$\inf(\mathscr{P}) = \inf(\mathscr{Q}) = \min(\mathscr{Q}\mathscr{R}) = \min(\mathscr{P}\mathscr{R}). \quad \blacksquare \tag{3.75}$$

Again, we can put the relaxed problem into a second, equivalent, form:

$$(\mathscr{P}\mathscr{R}') \quad \left| \begin{array}{l} \text{minimize} \displaystyle\int_\Omega \sum_{i=1}^{n+1} l_i(x) f(x, u(x), p_i(x))\,\mathrm{d}x \\ \displaystyle\sum_{i=1}^{n+1} l_i(x) = 1 \text{ and } l_i(x) \geqslant 0 \qquad 1 \leqslant i \leqslant n, \quad \text{a.e.} \\ \operatorname{grad} u(x) = \displaystyle\sum_{i=1}^{n+1} \ell_i(x) p_i(x), \quad \text{a.e.} \\ u = u_0 \quad \text{on } \partial\Omega. \end{array}\right.$$

3.4. A maximum principle for the relaxed problem

Starting from Corollary 3.8, it is natural to investigate the properties of solutions of problem $(\mathscr{P}\mathscr{R})$. In the case where $\alpha > n$, it is clear from the Sobolev embedding theorems that such relaxed solutions—because they belong to $W^{1,\alpha}(\Omega)$—are bounded on Ω and Hölder continuous on every compact subset:

$$\forall K \text{ compact} \subset \Omega, \qquad \operatorname*{Sup}_{x,y\in K} \frac{|u(x)-u(y)|}{\|x-y\|^{1-n/\alpha}} < +\infty. \tag{3.76}$$

Scheurer [1] has shown how those results generalize to $1 < \alpha \leqslant n$. He proves a maximum principle, and then introduces an additional hypothesis (the "bounded slope condition" of Stampacchia) under which the relaxed solutions are Lipschitz continuous.

We shall resistrict ourselves to proving the maximum principle in the simpler case where $\alpha \neq n$ and f does not depends on s:

Theorem 3.9. *Let Ω be a regular bounded open subset of $\mathbf{R}^n$, and $f: \Omega \times \mathbf{R}^n \to \overline{\mathbf{R}}$ a normal integrand such that:*

$$c|\xi|^\alpha \leqslant f(x,\xi) \quad \textit{with} \quad c > 0 \quad \textit{and} \quad 1 < \alpha < n \tag{3.77}$$

$$f(.,0) \in L^\gamma(\Omega) \quad \textit{with} \quad \gamma > \frac{n}{\alpha}\,. \tag{3.78}$$

Let $u_0 \in W^{1,\alpha}(\Omega) \cap L^\infty(\Omega)$ and denote by $\mathscr{S}$ the set of solutions of the problem:

$$(\mathscr{P}\mathscr{R}) \qquad \begin{cases} \operatorname{Inf} \displaystyle\int_\Omega f^{**}(x; \operatorname{grad} u(x))\,\mathrm{d}x \\ u - u_0 \in W_0^{1,\alpha}(\Omega). \end{cases}$$

There exists a constant $e \geqslant 0$, not depending on u, such that:

(3.79) $$\|u\|_\infty \leqslant \|u_0\|_\infty + e \qquad \forall u \in \mathscr{S}.$$

Before proving this theorem, let us first show in what sense it yields a maximum principle. Recall that, for any constant $a \in \mathbf{R}_+$, and any function $u \in W_0^{1,\alpha}(\Omega)$, the truncated function:

(3.80) $$T_a\, u(x) = \begin{cases} a & \text{if } \; u(x) > a \\ u(x) & \text{if } \; |u(x)| \leqslant a \\ -a & \text{if } \; u(x) < -a \end{cases}$$

belongs to $W_0^{1,\alpha}(\Omega)$. We can define the sup norm of $u_0 \in W_0^{1,\alpha}(\Omega) \cap L^\infty(\Omega)$ on the boundary $\partial\Omega$ as:

(3.81) $$\|\partial u_0\|_\infty = \inf\{a \in \mathbf{R}_+ \,|\, u_0 - T_a\, u_0 \in W_0^{1,\alpha}(\Omega)\}.$$

Formula (3.79) then becomes:

(3.82) $$\|u\|_\infty \leqslant \|\partial u_0\|_\infty + e \qquad \forall u \in \mathscr{S}.$$

In the special case where $e = 0$, we recover the familiar form of the maximum principle. In the general case, the sup norm of u over the domain Ω can be estimated by its sup norm over the boundary $\partial\Omega$.

Proof. Take $u \in \mathscr{S}$. For any $k \geqslant \|u_0\|_\infty$, we consider the set:

(3.83) $$A(k) = \{x \in \Omega \,|\, u(x) > k\}$$

and the truncated function:

(3.84) $$T_a^+\, u(x) = \begin{cases} k & \text{if } \; u(x) > k \\ u(x) & \text{if } \; u(x) \leqslant k. \end{cases}$$

We know that $T_a^+ u \in W^{1,\alpha}$ and $T_a^+ u - u_0 \in W_0^{1,\alpha}$.
As u is a solution of problem $(\mathscr{P}\mathscr{R})$:

(3.85) $$\int_\Omega f^{**}(x; \operatorname{grad} u(x))\, \mathrm{d}x \leqslant \int_\Omega f^{**}(x; \operatorname{grad} T_a^+ u(x))\, \mathrm{d}x.$$

Using the fact that $u(x) = T_a^+\, u(x)$ over $\Omega - A(k)$, this yields:

(3.86) $$\int_{A(k)} f^{**}(x; \operatorname{grad} u(x))\, \mathrm{d}x \leqslant \int_{A(k)} f^{**}(x; 0)\, \mathrm{d}x.$$

Using the estimates (3.77) and (3.78):

(3.87) $$\int_{A(k)} |\operatorname{grad} u(x)|^\alpha\, \mathrm{d}x \leqslant \frac{1}{C} \int_{A(k)} f^{**}(x; 0)\, \mathrm{d}x$$

the integrand on the right-hand side being in $L^\gamma(\Omega)$. By the Hölder inequality:

$$(3.88)\qquad \int_{A(k)} |\operatorname{grad} u(x)|^{\alpha}\,\mathrm{d}x \leqslant \frac{1}{c}\,\|f^{**}(\,.\,;0)\|_{\gamma}\,(\operatorname{meas} A(k))^{1-1/\gamma}.$$

Note that the left-hand side can be written as $\|\operatorname{grad} v\|_{\alpha}^{\alpha}$, where $v = \max(u-k,0)$ belongs to $W_0^{1,\alpha}$. We can apply to v the Sobolev embedding theorems: there exists a constant S depending only on Ω and α such that:

$$\|v\|_{\alpha^*} \leqslant S\|\operatorname{grad} v\|_{\alpha}$$

where $\dfrac{1}{\alpha^*} = \dfrac{1}{\alpha} - \dfrac{1}{n}$. Taking (3.88) into account, this becomes:

$$(3.89)\qquad \left(\int_{A(k)} (u-k)^{\alpha^*}\,\mathrm{d}x\right)^{1/\alpha^*} \leqslant c(\operatorname{meas} A(k))^{1/\alpha-1/\alpha\gamma}$$

where $c = \|f^{**}(\,.\,;0)\|_{\gamma}^{\alpha}(S/c)^{1/\alpha}$. By Hölder's inequality:

$$(3.90)\qquad \int_{A(k)} (u-k)\,\mathrm{d}x \leqslant \left(\int_{A(k)} (u-k)^{\alpha^*}\,\mathrm{d}x\right)^{1/\alpha^*} (\operatorname{meas} A(k))^{1-1/\alpha^*}$$

and substituting (3.89) into (3.90):

$$(3.91)\qquad \begin{aligned}\int_{A(k)} (u-k)\,\mathrm{d}x &\leqslant c(\operatorname{meas} A(k))^{1-1/\alpha^*+1/\alpha-1/\alpha\gamma}\\ &= c(\operatorname{meas} A(k))^{1+\varepsilon} \quad\text{with}\quad \varepsilon = \frac{1}{n} - \frac{1}{\alpha\gamma} > 0.\end{aligned}$$

The inequality (3.91) holds for every $k \geqslant \|u_0\|_{\infty}$. Applying Lemma 3.10, which follows, to the function $w = u - \|u_0\|_{\infty}$, we obtain:

$$u(x) \leqslant \|u_0\|_{\infty} + d.$$

In a similar way, we can prove that:

$$u(x) \geqslant -\|u_0\|_{\infty} - d'$$

where d' is a non-negative constant. Whence the result, by setting $e = \max(d, d')$. ∎

Lemma 3.10. *There exists a constant* $e > 0$ *such that*:

$$\int_{\{u(x)>k\}} (u-k)\,\mathrm{d}x \leqslant c[\operatorname{meas} A(k)]^{1+\varepsilon} \qquad \forall k \geqslant 0 \Rightarrow u(x) \leqslant e \quad a.e.$$

Proof. Using the notation (3.83), we consider the function: $H: \mathbf{R}_+ \to \mathbf{R}_+$ defined by

$$H(k) = \int_{A(k)} (u-k)\,\mathrm{d}x.$$

Its derivative is easily computed: $H'(k) = \operatorname{meas} A(k)$. This enables us to write the assumption regarding u in the following way:

$$H(k) \leqslant c[-H'(k)]^{1+\varepsilon} \quad\text{or}\quad \frac{-H'(k)}{[H(k)]^{1/(1+\varepsilon)}} \geqslant c^{-1/(1+\varepsilon)}.$$

Set $m = \operatorname{ess\,sup} u$. Integrating the preceding inequality between $k = 0$ and $k = m$, and taking into account the fact that $H(m) = 0$, we obtain:

$$H(0)^{\varepsilon/(1+\varepsilon)} \geqslant \frac{\varepsilon}{1+\varepsilon} c^{1/(1+\varepsilon)} m.$$

Hence:

$$m \leqslant \frac{1+\varepsilon}{\varepsilon} c^{1/(1+\varepsilon)} H(0)^{\varepsilon/(1+\varepsilon)} \leqslant \frac{1+\varepsilon}{\varepsilon} c^{1/(1+\varepsilon)} c^{\varepsilon/(1+\varepsilon)} [\text{meas } A(0)]^{\varepsilon}$$

$$\operatorname{ess\,sup} u = m \leqslant \frac{1+\varepsilon}{\varepsilon} c[\text{meas } \Omega] \, \varepsilon = e. \quad \blacksquare$$

Note that this lemma is a refinement of Theorem VIII.1.3.

We refer to the work of Scheurer [1] for a treatment of the general case, as well as further results concerning the Hölder or Lipschitz continuity of solutions of the relaxed problem.

4. EULER EQUATIONS

4.1. Exact and approximate solutions

Let V be a Banach space and F a l.s.c. function from V into $\mathbf{R} \cup \{+\infty\}$ which is assumed to be Gâteaux-differentiable at all points of its effective domain. We consider the optimization problem without constraints:

$$(\mathscr{P}) \qquad \operatorname{Inf} F(v) \quad \text{for} \quad v \in v.$$

If there exists a solution $\bar{v}$, it necessarily satisfies the equation $(\mathscr{E})$

$$(\mathscr{E}) \qquad F'(v) = 0$$

called the *Euler equation* of the problem. Conversely, if in addition F is convex, all solutions of the Euler equation achieve the minimum of F over V.

But these hypotheses by themselves are far from sufficient to ensure the existence of a solution of problem $(\mathscr{P})$ or of equation $(\mathscr{E})$. However, we have approximate solutions in the following sense (as usual $\|.\|_*$ denotes the norm of V^* and B_* its unit ball).

Proposition 4.1. *If F is l.s.c. and* $\inf F > -\infty$, *there exists in V a sequence* v_n, $n \in \mathrm{N}$, *such that*:

$$(4.1) \qquad F(v_n) \to \inf F$$

$$\|F'(v_n)\|_* \to 0. \tag{4.2}$$

Proof. We take in V a sequence u_n, $n \in \mathbf{N}$, such that $F(u_n) - \inf F \leqslant 1/n$. From Corollary I.6.1, for all $n \in \mathbf{N}$ there exists a $v_n \in V$ such that:

$$F(v_n) \leqslant F(u_n)$$
$$\|F'(v_n)\|_* \leqslant \sqrt{1/n}.$$

This sequence v_n, $n \in \mathbf{N}$, satisfies the Proposition. ∎

Corollary 4.1. *If F is l.s.c. and if there exists $k > 0$ and $c \in R$ such that*

$$\forall v \in V, \qquad F(v) \geqslant k\,\|v\| + c \tag{4.3}$$

then $F'(V)$ is dense in kB_*.

Proof. Possibly replacing F by $F - c$, we can always assume that $c = 0$. It is sufficient to show that $F'(V)$ is dense in $k\mathring{B}_*$. We therefore take $v^* \in V^*$ such that $\|v^*\|_* < k$ and show that we can approximate it by elements of $F'(V)$.

Let us set $G(v) = F(v) - \langle v, v^*\rangle$. This is also a l.s.c. function from V into $\mathbf{R} \cup \{+\infty\}$, Gâteaux-differentiable over its effective domain, and $G'(v) = F'(v) - v^*$. From (4.3), we have:

$$\forall v \in V, \qquad G(v) \geqslant (k - \|v^*\|_*)\,\|v\|$$

and thus $\inf G \geqslant 0$. From Proposition 4.1, there exists a sequence v_n $n \in \mathbf{N}$, in V such that:

$$\|G'(v_n)\|_* \to 0$$
$$\|F'(v_n) - v^*\|_* \to 0. \quad ∎$$

Corollary 4.2. *If F is l.s.c. and non-negative and if:*

$$\frac{F(v)}{\|v\|} \to \infty \quad \textit{when} \quad v \to \infty \tag{4.4}$$

then $F' : V \to V^*$ possesses a dense image.

Proof. Taking $k > 0$, from (4.4) there exists $a > 0$ such that

$$F(v) \geqslant k\,\|v\| \quad \text{for} \quad \|v\| \geqslant a$$

and hence, since F is non-negative:

$$\forall v \in V, \qquad F(v) \geqslant k\,(\|v\| - a).$$

From Corollary 4.2, $F'(V)$ is dense in kB^*. Since $k > 0$ is arbitrary, $F'(V)$ is dense in V^*. ∎

Henceforth we shall assume that the Γ-regularization F^{**} of F coincides with its weak l.s.c. regularization $\bar{F}$, a hypothesis which we have proved in Chapters IX and X for certain functionals of the calculus of variations. We can then render Proposition 4.1 more precise:

Proposition 4.2. *With the assumptions that F is l.s.c., that $F^{**} = \bar{F}$ and that $\bar{v}$ minimizes F^{**} over V, then there exists in V a sequence v_n, $n \in \mathbf{N}$, such that*

$$v_n \to \bar{v} \ \ \textit{weakly}, \tag{4.5}$$

$$F(v_n) \to \inf F = F^{**}(\bar{v}), \tag{4.6}$$

$$F'(v_n) \to 0 \in \partial F^{**}(\bar{v}). \tag{4.7}$$

Proof. By definition:

$$\inf F = \bar{F}(\bar{v}) = \inf \{ F(v) \mid v \to \bar{v} \text{ weakly}\}.$$

There thus exists a sequence u_n, $n \in \mathbf{N}$ such that:

$$u_n \to \bar{v} \quad \text{weakly} \tag{4.8}$$

$$F(u_n) \to \bar{F}(\bar{v}) = \inf F. \tag{4.9}$$

From Corollary I.6.1, there exists for all $n \in \mathbf{N}$ a $v_n \in V$ such that:

$$F(v_n) \leqslant F(u_n) \tag{4.10}$$

$$\|v_n - u_n\| \leqslant \sqrt{1/n} \tag{4.11}$$

$$\|F'(v_n)\|_* \leqslant \sqrt{1/n}. \tag{4.12}$$

We obtain (4.5) from (4.8) and (4.11), (4.6) from (4.9) and (4.10), (4.7) from (4.12). ∎

4.2. Calculation of the differential: the convex case

Lemma 4.1. *Let $f : \Omega \times R^p \to \overline{\mathbf{R}}_+$ be a non-negative Carathéodory function, such that, for almost all $x \in \Omega$, $\xi \mapsto f(x, \xi)$ is convex and differentiable through-*

out its effective domain. We assume that there exist $v_0 \in L^\alpha(\Omega)$ *and* $v_0^* \in L^{\alpha'}(\Omega)$ *with* $1 \leqslant \alpha < \infty$ *and* $1/\alpha + 1/\alpha' = 1$, *such that*:

$$\int_\Omega f(x, v_0(x))\,\mathrm{d}x < \infty \quad \textit{and} \quad \int_\Omega f^*(x; v_0^*(x))\,\mathrm{d}x < \infty. \tag{4.13}$$

We define a function F over $L^\alpha(\Omega)$ *by*:

$$F(v) = \int_\Omega f(x, v(x))\,\mathrm{d}x. \tag{4.14}$$

F is subdifferentiable at $v \in V$ *if and only if the function*: $x \mapsto f'_\xi(x, v(x))$ *belongs to* $L^{\alpha'}(\Omega)$, *and it is then the unique subgradient of F at v.*

Proof. From Proposition IX.2.1, we have

$$\forall v^* \in V, \qquad F^*(v^*) = \int_\Omega f^*(x; v^*(x))\,\mathrm{d}x.$$

Let us fix $v \in V$. To say that $v^* \in \partial F(v)$ means that $v \in \operatorname{dom} F$ and that (Proposition I.5.1)

$$F(v) + F^*(v^*) - \langle v, v^* \rangle = 0.$$

On substituting the value for each term, we then have:

$$\int_\Omega [f(x, v(x)) + f^*(x; v^*(x)) - v(x)v^*(x)]\,\mathrm{d}x = 0.$$

Now, by definition of the polar $f^*(x; .)$, the term in the square brackets is $\geqslant 0$. If the integral is null and the integrand non-negative, then the integrand is zero almost everywhere:

$$f(x, v(x)) + f^*(x; v^*(x)) - v(x)v^*(x) = 0 \qquad \text{a.e.}$$

By applying Proposition IX.5.1 again

$$v^*(x) \in \partial f_\xi(x, v) \qquad \text{a.e.}$$
$$v^*(x) = f'_\xi(x, v) \qquad \text{a.e.} \quad \blacksquare$$

This Lemma allows us to find easily the Euler equations for convex problems of the calculus of variations. We give two examples:

Example 1. *Exact solution of a fourth-order equation.*

Let $g: \Omega \times \mathbf{R} \to \mathbf{R}_+$ be a non-negative Carathéodory function which satisfies the estimate:

(4.15) $$0 \leqslant g(x, s) \leqslant a_1(x) + b_1 |s|^2, \text{ where } a_1 \in L^1(\Omega) \text{ and } b_1 \geqslant 0$$

and such that, for almost all $x \in \Omega$, $g(x, .)$ is convex and differentiable.

Let $f: \Omega \times \mathbf{R} \to \overline{\mathbf{R}}$ be a Carathéodory function, which satisfies the estimate:

(4.16) $$a_2(x) + b_2 |\xi|^2 \leqslant f(x, \xi), \text{ where } a^2 \in L^1(\Omega) \text{ and } b_2 \geqslant 0$$

and such that, for almost all $x \in \Omega, f(x, .)$ is convex and differentiable throughout its effective domain. We assume that there exists $p_0 \in L^2(\Omega)$ such that:

(4.17) $$\int_\Omega f(x, p_0(x))\, \mathrm{d}x < +\infty.$$

We assume the open space Ω to be sufficiently regular for the Laplace operator Δ to be a bijection of $H_0^1 \cap H^2$ over L^2, and we examine the following problem in the calculus of variations:

(4.18) $$\left| \begin{array}{l} \operatorname{Inf} \displaystyle\int_\Omega [g(x, v(x)) + f(x, \Delta v(x))]\, \mathrm{d}x \\ v \in H_0^1 \cap H^2. \end{array} \right.$$

To put this problem into the usual form, we introduce the functions G and F on $L^2(\Omega)$ defined by:

(4.19) $$G(v) = \int_\Omega g(x, v(x))\, \mathrm{d}x$$

(4.20) $$F(p) = \int_\Omega f(x, p(x))\, \mathrm{d}x.$$

The canonical injection i from H_0^1 into L^2 is compact. The Green operator $\mathscr{G} = i \circ \Delta^{-1}$ is therefore compact from L^2 into L^2, and we can easily check that it is self-adjoint. We can thus put the problem under consideration into the following form:

(4.21) $$\left| \begin{array}{l} \operatorname{Inf} G \circ \mathscr{G}(p) + F(p) \\ p \in L^2. \end{array} \right.$$

These are the hypotheses of Theorem VIII.2.2 and the problem therefore has a solution $\bar{v} = \mathscr{G}\bar{p}$. The point $\bar{p}$ satisfies the Euler equation

$$0 \in \partial(G \circ \mathscr{G} + F)(\bar{p}). \tag{4.22}$$

But F is l.s.c. and $G \circ \mathscr{G}$ is continuous on L^2 (Proposition VIII.1.4). From Proposition I.5.6, we have:

$$\partial(G \circ \mathscr{G} + F)(\bar{p}) = \partial(G \circ \mathscr{G})(\bar{p}) + \partial F(\bar{p}). \tag{4.23}$$

From Lemma 4.1, and taking into account formula (2.30) of Chapter IX, we have:

$$\partial F(\bar{p}) = \{ x \mapsto f'_\xi(x, \bar{p}(x)) \} \tag{4.24}$$

$$\partial G(\mathscr{G}\bar{p}) = \{ x \mapsto g'_s(x, \mathscr{G}\bar{p}(x)) \}. \tag{4.25}$$

Since G is continuous, we obtain $\partial(G \circ \mathscr{G})(\bar{p})$ by applying the chain rule for subdifferentials (Proposition I.5.7).

$$\partial(G \circ \mathscr{G})(\bar{p}) = \mathscr{G}' \, \partial G(\mathscr{G}\bar{p}).$$

By identifying L^2 with its dual, we replace the transpose $\mathscr{G}'$ by the adjoint $\mathscr{G}^* = \mathscr{G}$.

$$\partial(G \circ \mathscr{G})(\bar{p}) = \mathscr{G} \, \partial G(\mathscr{G}\bar{p}). \tag{4.26}$$

Finally, the Euler equation can be written as

$$\mathscr{G} g'_s(x, \mathscr{G}\bar{p}(x)) + f'_\xi(x, \bar{p}(x)) = 0 \quad \text{in} \quad L^2(\Omega).$$

Or again, by definition of $\mathscr{G}$:

$$g'_s(x, \bar{v}(x)) + \Delta f'_\xi(x, \Delta\bar{v}(x)) = 0 \quad \text{in} \quad L^2(\Omega). \tag{4.27}$$

Or finally, on writing out the Laplacian explicitly:

$$g'_s(x, v(x)) + \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}\left[f'_\xi\left(x, \sum_{i=1}^{n} \frac{\partial^2 v}{\partial x_i^2}(x) \right)\right] = 0 \qquad \text{a.e.} \tag{4.28}$$

Conclusion. *Equation* (4.28) *has a solution $\bar{v}$ in $H_0^1 \cap H^2$. This solution is unique if $f(x,.)$ or $g(x,.)$ is strictly convex for almost all x.*

Example 2. *Approximate solution of the equation of minimal hypersurfaces.*

We take $v_0 \in W^{1,1}(\Omega)$, and we consider Plateau's problem:

$$\left| \begin{array}{l} \operatorname{Inf} \displaystyle\int_\Omega [1 + |\operatorname{grad} v|^2]^{1/2} \, dx \\ v - v_0 \in W_0^{1,1}(\Omega). \end{array} \right. \tag{4.29}$$

Setting $\operatorname{grad} v_0 = p_0$, we can write the problem in the following form:

$$\left| \begin{array}{l} \operatorname{Inf} \displaystyle\int_\Omega [1 + |\operatorname{grad} u(x) + p_0(x)|^2]^{1/2} \, dx \\ u \in W_0^{1,1}(\Omega). \end{array} \right. \tag{4.30}$$

Let us define the function F on $L_n^1(\Omega)$ by

$$F(p) = \int_\Omega [1 + |p(x) + p_0(x)|^2]^{1/2} \, dx. \tag{4.31}$$

It is convex and continuous and therefore subdifferentiable throughout L_n^1. From Lemma 4.1, $\partial F(p)$ is reduced to the function $(p + p_0)/[1 + |p + p_0|^2]^{1/2}$. Since F is continuous, we can apply the chain rule for subgradients to $F \circ \operatorname{grad}$ (Proposition I.5.7): $F \circ \operatorname{grad}$ possesses at every point $u \in W_0^{1,1}(\Omega)$ a unique subgradient:

$$- \operatorname{div} \frac{\operatorname{grad} u + p_0}{[1 + |\operatorname{grad} u + p_0|^2]^{1/2}}. \tag{4.32}$$

In particular, $F \circ \operatorname{grad}$ is Gâteaux-differentiable over $W_0^{1,1}(\Omega)$ (Proposition I.5.3), its differential being given by (4.32). The Euler equation can thus be written as:

$$- \operatorname{div} \frac{\operatorname{grad} u + p_0}{[1 + |\operatorname{grad} u + p_0|^2]^{1/2}} = 0, \qquad u \in W_0^{1,1}(\Omega) \tag{4.33}$$

or, by writing $u + v_0 = v$, as:

$$- \operatorname{div} \frac{\operatorname{grad} v}{[1 + |\operatorname{grad} v|^2]^{1/2}} = 0, \qquad v - v_0 \in W_0^{1,1}(\Omega). \tag{4.34}$$

However, we know that in general neither Plateau's problem (4.29) nor the equation of minimal hypersurfaces (4.34) have exact solutions. We must therefore seek approximate solutions.

Clearly, for $u \in W_0^{1,1}$:

(4.35)
$$\int_\Omega [1 + |\operatorname{grad} u(x) + p_0(x)|^2]^{1/2}\, dx \geqslant \int_\Omega |\operatorname{grad} u(x)|\, dx - \int_\Omega |p_0(x)|\, dx.$$

From Poincaré's inequality, there exists $k > 0$ such that:

(4.36)
$$\forall u \in W_0^{1,1}, \qquad \int_\Omega |\operatorname{grad} u(x)|\, dx \geqslant k\, \|u\|_{W^{1,1}}.$$

And thus (4.35) can be written as

(4.37)
$$F(u) \geqslant k\, \|u\| - \text{constant}.$$

We can therefore apply Corollary 4.1: there exists in the ball kB^* of $W^{-1,\infty}$ a dense subset $\mathscr{S}$ such that for all $T \in \mathscr{S}$ the equation

(4.38)
$$-\operatorname{div} \frac{\operatorname{grad} u + p_0}{[1 + |\operatorname{grad} u + p_0|^2]^{1/2}} = T, \qquad u \in W_0^{1,1}(\Omega)$$

has a solution $\bar{u}$. But equation (4.38) can also be written as:

(4.39)
$$\partial(F \circ \operatorname{grad})(\bar{u}) - \partial T(\bar{u}) = \{ 0 \}.$$

Since $F \circ \operatorname{grad}$ is continuous over $W_0^{1,1}(\Omega)$, the latter equation can also be written as Proposition I.5.6:

(4.40)
$$\partial(F \circ \operatorname{grad} - T)(\bar{u}) = \{ 0 \}$$

which means that the problem:

(4.41)
$$\left| \begin{array}{l} \operatorname{Inf} F \circ \operatorname{grad}(u) - \langle T, u \rangle \\ u \in W_0^{1,1}(\Omega) \end{array} \right.$$

has a solution. This is necessarily unique, by strict convexity, and the solution of (4.38) is thus unique. Writing $v = u + v_0$, we obtain the

Conclusion. *There exists in $W_0^{-1,\infty}(\Omega)$ an open subset around the origin which contains a dense subset $\mathscr{S}$ such that, for all $T \in \mathscr{S}$, the equation*:

$$-\operatorname{div} \frac{\operatorname{grad} v}{[1 + |\operatorname{grad} v|^2]^{1/2}} = T, \qquad v \in v_0 + W_0^{1,1}(\Omega)$$

and the problem:

$$\left| \begin{array}{l} \operatorname{Inf} \displaystyle\int_\Omega [1 + |\operatorname{grad} v|^2]^{1/2} \, \mathrm{d}x - \langle T, v \rangle \\ v \in v_0 + W_0^{1,1}(\Omega) \end{array} \right.$$

have a unique solution.

4.3. Calculation of the differential: the general case

In the non-convex case, the actual calculation of $F'(v)$ can be made by differentiation under the summation sign. We give an important example.

Let $f: \Omega \times \mathbf{R}^n \to \mathbf{R}_+$ be a non-negative Carathéodory function, such that for almost all x, $\xi \to f(x, \xi)$ is continuously differentiable (not necessarily convex) and satisfies the inequality:

$$(4.42) \qquad |f'_\xi(x, \xi)| \leqslant a + b\,|\xi|^{\alpha - 1}, \qquad \text{where } a \text{ and } b \geqslant 0 \text{ and } \alpha \geqslant 1.$$

We then define a function F on $W_0^{1,\alpha}(\Omega)$ by:

$$(4.43) \qquad \forall v \in W_0^{1,\alpha}, \qquad F(v) = \int_\Omega f(x, \operatorname{grad} v(x)) \, \mathrm{d}x$$

and finally we make the hypothesis that $\operatorname{dom} F \neq \varnothing$

$$(4.44) \qquad \exists v_0 \in W_0^{1,\alpha} : F(v_0) < +\infty.$$

F is non-negative and l.s.c. on $W_0^{1,\alpha}$ (Proposition VIII.1.4), but it is not necessarily convex and has no other reason for achieving its minimum. We shall show that it is finite everywhere and Gâteaux-differentiable and we shall apply Proposition 4.1 to it.

Thus, take $v \in W_0^{1,\alpha}$ such that $F(v) < +\infty$. We take $w \in W_0^{1,\alpha}$, and consider the function $t \mapsto F(v + tw)$ over $[0, 1]$:

$$(4.45) \qquad F(v + tw) = \int_\Omega f(x, \operatorname{grad} v(x) + t \operatorname{grad} w(x)) \, \mathrm{d}x.$$

For $0 \leqslant t \leqslant 1$, the inequality (4.42) gives us

$$(4.46) \qquad \left| \frac{\partial}{\partial t} f(x, \operatorname{grad} v(x) + t \operatorname{grad} w(x)) \right| =$$

$$= \langle \operatorname{grad} w(x), f'_\xi(x, \operatorname{grad} v(x) + t \operatorname{grad} w(x)) \rangle$$
$$\leqslant |\operatorname{grad} w(x)| (a + b\,|\operatorname{grad} v(x) + t \operatorname{grad} w(x)|^{\alpha - 1}).$$

We must thus consider the two cases: $1 \leqslant \alpha \leqslant 2$ and $\alpha > 2$.

First, let us assume that $1 \leqslant \alpha \leqslant 2$. Then $(\rho+\sigma)^{\alpha-1} \leqslant \rho^{\alpha-1} + \sigma^{\alpha-1}$ for ρ and σ real and positive, and (4.46) gives us:

$$(4.47) \qquad \left|\frac{\partial}{\partial t} f(x, \operatorname{grad} v(x) + t \operatorname{grad} w(x))\right| \leqslant$$
$$\leqslant |\operatorname{grad} w(x)| \, (a + b\,|\operatorname{grad} v(x)|^{\alpha-1} + b\,|\operatorname{grad} w(x)|^{\alpha-1}).$$

If now $\alpha > 2$, we have

$$(\rho + \sigma)^{\alpha-1} \leqslant 2^{\alpha-2}(\rho^{\alpha-1} + \sigma^{\alpha-1})$$

for ρ and σ real positive and (4.46) gives us:

$$(4.48) \qquad \left|\frac{\partial}{\partial t} f(x, \operatorname{grad} v(x) + t \operatorname{grad} w(x)\right| \leqslant$$
$$\leqslant |\operatorname{grad} w(x)| (a + 2^{\alpha-2} b\,|\operatorname{grad} v(x)|^{\alpha-1} + 2^{\alpha-2} b\,|\operatorname{grad} w(x)|^{\alpha-1}).$$

Bringing together (4.47) and (4.48) we obtain in every case for $0 \leqslant t \leqslant 1$:

$$(4.49) \qquad \left|\frac{\partial}{\partial t} f(x, \operatorname{grad} v(x) + t \operatorname{grad} w(x))\right| \leqslant g(x), \quad \text{where } g \in L^1(\Omega).$$

The inequality (4.49) is the standard condition for differentiability under the summation sign, applied to (4.45). We thus obtain for $0 \leqslant t \leqslant 1$

$$(4.50) \qquad F(v + tw) < +\infty$$

$$(4.51) \qquad \frac{\partial}{\partial t} F(v + tw)\big|_{t=0} = \int_\Omega \langle \operatorname{grad} w(x), f'_\xi(x, \operatorname{grad} v(x)) \rangle \, dx.$$

For all $v \in W_0^{1,\alpha}$, on taking $v = v_0$ and $w = v - v_0$ in (4.50), we obtain $F(v) < +\infty$. The functional F is thus finite everywhere: $\operatorname{dom} F = W_0^{1,\alpha}$. For all $v \in W_0^{1,\alpha}$, by virtue of (4.42), we have:

$$(4.52) \qquad \{ x \mapsto f'_\xi(x, \operatorname{grad} v(x)) \} \in L^{\alpha'}(\Omega), \quad \text{where } 1/\alpha + 1/\alpha' = 1.$$

Thus $-\operatorname{div} f'_\xi(x, \operatorname{grad} v(x)) \in W^{-1,\alpha}$ and (4.51) can be put into the form:

$$(4.53) \qquad \frac{\partial}{\partial t} F(v + tw)\big|_{t=0} = -\int_\Omega w(x) \operatorname{div} f'_\xi(x, \operatorname{grad} v(x)) \, dx$$

by writing, as usual, the canonical mapping from $W^{-1,\alpha}$ onto $W_0^{1,\alpha}$ (or from $\mathscr{D}'$ onto $\mathscr{D}$) in the integral form. The function F is thus Gâteaux-differentiable everywhere over $W_0^{1,\alpha}(\Omega)$, its differential at v being

$$-\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left[\frac{\partial f}{\partial \xi_i}(x, \text{grad } v(x))\right] \in W^{-1,\alpha'}. \tag{4.54}$$

It only remains to apply Proposition 1.1. We obtain the

Conclusion. *Under hypothesis* (4.42) *there exists in* $W_0^{1,\alpha}(\Omega)$ *a sequence* v_n, $n \in \mathbf{N}$ *such that*:

$$-\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left[\frac{\partial f}{\partial \xi_i}(x, \text{grad } v_n(x))\right] \to 0 \quad \text{in} \quad W^{-1,\alpha'}(\Omega). \tag{4.55}$$

If moreover $\alpha > 1$ and if we can find $a' \in L^1(\Omega)$ and $c > 0$ such that:

$$f(x, \xi) \geqslant a'(x) + c\,|\xi|^{\alpha} \tag{4.56}$$

then we can easily check that $F(v)/\|v\| \to \infty$ as v tends to infinity in $W_0^{1,\alpha}$, so that we can apply Corollary 4.2:

Conclusion. *Under hypotheses* (4.42) *and* (4.56), *the mapping*

$$v \mapsto -\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left[\frac{\partial f}{\partial \xi_i}(x, \text{grad } v(x))\right] \tag{4.57}$$

from $W^{1,\alpha}(\Omega)$ *into* $W^{-1,\alpha'}(\Omega)$ *possesses a dense image.*

An especially well-known case is the following

$$f(x, \xi) = \sum_{i=1}^{n} |\xi_i|^{\alpha} \qquad \text{with } \alpha > 1.$$

The inequalities (4.42) and (4.56) are satisfied. But this time the integrand is strictly convex and the space $W_0^{1,\alpha}$ reflexive. For all $T \in W^{-1,\alpha'}$ the function $F - T$ attains its minimum at a point which is the unique solution of the Euler equation

$$-\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(\left|\frac{\partial v}{\partial x_i}\right|^{\alpha-2} \frac{\partial v}{\partial x_i}\right) = 0 \qquad \text{in } \mathscr{D}'(\Omega). \tag{4.57}$$

Another special case where we can apply Proposition 4.2 is the following. We take:

$$f(x, \xi) = (1 - |\xi|^2)^2.$$

It is easily seen that f satisfies the foregoing hypotheses, in particular the inequality (4.42) is satisfied with $\alpha = 4$. But f also satisfies the hypotheses of Corollary 3.8, still with $\alpha = 4$ and we have:

$$f^{**}(x;\xi) = 0 \qquad \text{if } |\xi| \leqslant 1$$
$$f^{**}(x;\xi) = (1 - |\xi|^2)^2 \qquad \text{if } |\xi| \geqslant 1.$$

We take as boundary condition $u = 0$ over $\partial\Omega$, where $u \in W_0^{1,4}(\Omega)$. The optimal solution of the relaxed problem is then $u = 0$ over Ω. We deduce that, from Corollary 3.8, every minimizing sequence u_n of the problem

$$\left| \operatorname*{Inf}_{u \in W_0^{1,4}} \int_\Omega (1 - |\operatorname{grad} u(x)|^2)^2 \, dx \right.$$

converges to 0 weakly in $W_0^{1,4}(\Omega)$. Since furthermore

$$\inf(\mathscr{P}) = \min(\mathscr{P}\mathscr{R}) = 0$$

we have:

$$\int_\Omega (1 - |\operatorname{grad} u_n(x)|^2)^2 \, dx \to 0$$
$$\|1 - |\operatorname{grad} u_n|^2\|_{L^2} \to 0$$

and in particular, possibly by extracting a subsequence,

$$|\operatorname{grad} u_n(x)| \to 1 \quad \text{a.e.}$$

From Proposition 4.2, there exists a particular minimizing sequence v_n such that

$$\left| \begin{array}{ll} \displaystyle\sum_{i=1}^{n} \frac{\partial}{\partial x_i} [\operatorname{grad} v_n(x)(1 - |\operatorname{grad} v_n(x)|^2] \to 0 & \text{in } W^{-1,4/3}(\Omega). \\ v_n \to 0 \ \text{weakly in } W^{-1,4}(\Omega) \text{ faible}, & \\ \|1 - |\operatorname{grad} v_n|^2\|_{L^2} \to 0. & \end{array} \right.$$

This page intentionally left blank

An *a priori* Estimate in Non-convex Programming

Part II of this book has been devoted to the study of convex normal problems; recall that an optimization problem $(\mathscr{P})$ is called normal if:

$$\inf \mathscr{P} = \sup \mathscr{P}^* \neq \pm\infty.$$

It should be clear from Chapter III, particularly Lemma 2.3, that this relationship cannot be expected to hold for non-convex problems. It is replaced by the general inequality:

$$\inf (\mathscr{P}) \geqslant \sup (\mathscr{P}^*).$$

The non-negative number $\inf (\mathscr{P}) - \sup (\mathscr{P}^*)$ is called the *duality gap*. We shall give an *a priori* estimate of this duality gap in an important case. We begin by proving a theorem of Shapley and Folkman, which we shall use later on.

1. THE SHAPLEY–FOLKMAN THEOREM

Theorem 1. *Let K_i, $i \in I$, be a finite family of subsets of $\mathbf{R}^k$. For every $\hat{x} \in \sum_I \operatorname{co} K_i$, there exists a subfamily $J(\hat{x}) \subset I$ containing at most k elements, and such that*:

$$\hat{x} \in \sum_{I-J(\hat{x})} K_i + \sum_{J(\hat{x})} \operatorname{co} K_i.$$

Note that k is the dimension of the ambient space. By definition, $\hat{x}$ can be expressed, probably in many different ways, as a sum $\hat{x} = \sum_I \hat{x}_i$, with $\hat{x}_i \in \operatorname{co} K_i$ for every i. The Shapley–Folkman theorem states that it can also be expressed as a sum $\hat{x} = \sum_I x_i$, with all but k of the x_i belonging to the K_i.

Define a mapping Φ from $\mathbf{R}^{kI}$ to $\mathbf{R}^k$ by:

$$\Phi((x_i)_{i \in I}) = \sum_I x_i.$$

By definition

$$\sum_I K_i = \Phi\left(\prod_I K_i\right).$$

By linearity of the mapping Φ:

$$\operatorname{co}\Phi\left(\prod_I K_i\right) = \Phi \operatorname{co}\prod_I K_i = \Phi\left(\prod_I \operatorname{co} K_i\right)$$

$$\operatorname{co}\sum_I K_i = \sum_I \operatorname{co} K_i.$$

We now proceed to the proof of the Shapley–Folkman theorem. Note first that $\hat{x} \in \operatorname{co}\sum_I K_i$ if and only if $\hat{x}$ belongs to the convex hull of a finite number n of points of $\sum_I K_i$. It can be shown that $n \leqslant k+1$ (Carathéodory's theorem, see Rockafellar, [4]), but we shall not use this refinement. We may therefore express $\hat{x}$ as:

$$\hat{x} = \sum_{j=1}^{n} \alpha_j y_j, \quad \text{with} \quad y_j \in \sum_I K_i, \alpha_j > 0 \quad \text{and} \sum_{j=1}^{n} \alpha_j = 1.$$

Further, every y_j can be expressed as:

$$y_j = \sum_{i \in I} y_{ij}, \quad \text{with} \quad y_{ij} \in K_i.$$

Denote by F_i the n-set $\{y_{ij}\}_{1 \leqslant j \leqslant n}$. Clearly, $y_j \in \sum_I F_i$ for every j, so that:

$$\hat{x} \in \operatorname{co}\sum_I F_i.$$

We have thus replaced each set K_i by a finite subset $F_i \subset K_i$. It will be convenient for the rest of the proof to note that the co F_i are polyhedra in $\mathbf{R}^k$, and their product co $\prod_I F_i$ is a polyhedron in $\mathbf{R}^{kI}$.

Denote by H the inverse image of $\hat{x}$ under the mapping Φ. We are interested in the subset P of $\mathbf{R}^{kI}$:

$$P = H \cap \operatorname{co}\prod_I F_i = \left\{(x_i)_{i \in I} \mid x_i \in \operatorname{co} F_i \text{ and } \sum_I x_i = \hat{x}\right\}.$$

The assumption $\hat{x} \in \sum_I \operatorname{co} F_i$ means P is non-empty. Moreover, as co $\prod_I F_i$ is a polyhedron and H an affine subspace, P is a polyhedron. Let $(x_i)_{i \in I}$ be one of its vertices. We still have $\hat{x} = \sum_I x_i$, with $x_i \in \operatorname{co} F_i$, as the point $(x_i)_{i \in I}$ belongs to P. Moreover, we shall prove that all but k at most of the x_i are vertices of the corresponding co F_i. As any vertex of the polyhedron co F_i must belong to F_i, this will prove the theorem.

Otherwise suppose that there exist $(k+1)$ components of $(x_i)_{i \in I}$ which are not vertices in the corresponding co F_i. Let us denote them by $x_i, \ldots, x_{k+1}$. For every x_i, $1 \leqslant i \leqslant k+1$, there exist a vector $z_i \in \mathbf{R}^k$ and a number $\varepsilon_i > 0$ such that:

$$\text{(1)} \qquad \forall t \in [-\varepsilon_i, \varepsilon_i], \quad x_i + tz_i \in \text{co } F_i.$$

Denote $\varepsilon = \min_{1 \leqslant i \leqslant k+1} \varepsilon_i$.

Now, if we have $(k+1)$ vectors in a space of dimension k, there is a linear relationship between them. Hence, there exist numbers $\alpha_1, \ldots, \alpha_{k+1}$, not all zero such that:

$$\sum_{i=1}^{k+1} \alpha_i z_i = 0.$$

We may assume that $|\alpha_i| \leqslant 1$ for $1 \leqslant i \leqslant k+1$. Define now two points $(x_i')_{i \in I}$ and $(x_i'')_{i \in I}$ of $\mathbf{R}^{kI}$ by:

$$x_i' = x_i + \varepsilon \alpha_i z_i \quad \text{for} \quad 1 \leqslant i \leqslant k+1$$

$$x_i'' = x_i - \varepsilon \alpha_i z_i \quad \text{for} \quad 1 \leqslant i \leqslant k+1$$

$$x_i' = x_i = x_i'' \quad \text{otherwise.}$$

It follows from (1) that x_i' and x_i'' belong to co F_i. Moreover:

$$\sum_I x_i' = \sum_I x_i + \varepsilon \sum_{i=1}^{k+1} \alpha_i z_i = \hat{x}$$

$$\sum_I x_i'' = \sum_I x_i - \varepsilon \sum_{i=1}^{k+1} \alpha_i z_i = \hat{x}.$$

Hence, the points $(x_i')_{i \in I}$ and $(x_i'')_{i \in I}$ belong to P. But clearly:

$$(x_i)_{i \in I} = \tfrac{1}{2}(x_i')_{i \in I} + \tfrac{1}{2}(x_i'')_{i \in I}$$

and $(x_i)_{i \in I}$ cannot be a vertex of P, albeit it was assumed. ∎

2. ESTIMATING THE DUALITY GAP: SIMPLE CASE

Consider the following optimization problem:

$$\begin{cases} \operatorname{Inf} \sum_{i=1}^{n} f_i(x_i) + f_0\left(-\sum_{i=1}^{n} x_i\right) \\ x_i \in \mathbf{R}^k \quad \text{for} \quad 1 \leqslant i \leqslant n. \end{cases}$$

As usual, we allow $\pm\infty$ as values for the functions f_i. For instance, we might choose a subset K of $\mathbf{R}^k$, and define $f_0 = \chi_K$, the indicator of K. Such problems occur for instance in economics; we have seen some of them in Section VII.5. A striking feature is that n is usually very great with respect to k.

We make no convexity assumption at all on the functions f_i. We do not need any topological assumption either; nevertheless, we shall make one for simplicity's sake. By hypothesis, all the f_i, $0 \leqslant i \leqslant n$ are lower semi-continuous, and satisfy:

$$\forall i, f_i(x_i)/\|x_i\| \to +\infty \quad \text{as} \quad \|x_i\| \to \infty.$$

The usual approach to duality of Chapter III applies quite easily. We shall use the simplified theory of Section III.4. We take as perturbation function:

$$\Phi(x_1, \ldots, x_n; p) = \sum_{i=1}^{n} f_i(x_i) + f_0\left(p - \sum_{i=1}^{n} x_i\right)$$

with $p \in \mathbf{R}^k$. Denote as usual:

$$(\mathscr{P}_p) \qquad \operatorname*{Inf}_{x_i \in \mathbf{R}^k} \Phi(x_1; \ldots, x_n; p)$$

$$h(p) = \operatorname{Inf} \mathscr{P}_p.$$

The topological assumptions we have just made simplify matters in the following way:

Lemma 1. *h is a lower semi-continuous function on* $\mathbf{R}^k$.

Proof. Let p_r, $r \in \mathbf{N}$ be a sequence converging to $\bar{p}$ in $\mathbf{R}^k$. Using the topological assumptions on the functions f_i, we see that there exists for every $r \in \mathbf{N}$ a family (x_{ir}), $1 \leqslant i \leqslant n$, such that:

$$h(p_r) = \sum_{i=1}^{n} f_i(x_{ir}) + f_0\left(p_r - \sum_{i=1}^{n} x_{ir}\right).$$

Denote $\underline{\lim}_{r\to\infty} h_{(p_r)}$ by ℓ. If $\ell = +\infty$, the function h is lower semi-continuous at $\bar{p}$. If not, the subset of $\mathbf{R}^{kn}$:

$$K_\varepsilon = \left\{ (x_1, \ldots, x_n) \middle| \sum_{i=1}^{n} f_i(x_i) + f_0 \left(\bar{p} - \sum_{i=1}^{n} x_i \right) \leqslant \ell + \varepsilon \right\}$$

is compact. Note that there exists a subsequence $(x_{ir'})$ such that:

$$\sum_{i=1}^{n} f_i(x_{ir'}) + f_0 \left(p_{r'} - \sum_{i=1}^{n} x_{ir'} \right) \leqslant \ell + \varepsilon.$$

Using the lower semi-continuity of f_0 again:

$$\sum_{i=1}^{n} f_i(x_{ir'}) + f_0 \left(\bar{p} - \sum_{i=1}^{n} x_{ir'} \right) \leqslant \ell + \varepsilon.$$

Hence, the sequence (x_{ir}) has a cluster point $(\bar{x}_i)$ in K_ε. Now, the set of cluster points of the sequence (x_{ir}) is closed and intersects K_ε for every $\varepsilon > 0$. By compactness, it must intersect $\bigcap_{\varepsilon>0} K = K_0$. Hence, the sequence (x_{ir}) has a cluster point $(\bar{x}_i)$ such that:

$$\sum_{i=1}^{n} f_i(\bar{x}_i) + f_0 \left(- \sum_{i=1}^{n} \bar{x}_i \right) \leqslant \ell.$$

The result follows from the definition of ℓ, and the fact that the left-hand side is greater than $h(p)$. ■

To express the dual problem of $\mathscr{P}$ with respect to this set of perturbations, we shall have to compute:

$$\Phi^*(0, \ldots, 0; p^*) = \operatorname*{Sup}_{\substack{x_i; p \\ 1 \leqslant i \leqslant n}} \left\{ \langle p^*, p \rangle - \sum_{i=1}^{n} f_i(x_i) - f_0 \left(p - \sum_{i=1}^{n} x_i \right) \right\}.$$

Denote $x_0 = p - \sum_{i=1}^{n} x_i$:

$$\Phi^*(0, \ldots, 0; p^*) = \operatorname*{Sup}_{\substack{x_i, x_0 \\ 1 \leqslant i \leqslant n}} \left\{ \left\langle p^*, x_0 + \sum_{i=1}^{n} x_i \right\rangle - \sum_{i=1}^{n} f_i(x_i) - f_0(x_0) \right\}$$

$$= \sum_{i=0}^{n} \operatorname{Sup} \{ \langle p^*, x_i \rangle - f_i(x_i) \} = \sum_{i=0}^{n} f_i^*(p^*).$$

Hence the dual problem:

$$(\mathscr{P}^*) \qquad \operatorname*{Sup}_{p^* \in \mathbb{R}_k} \; - \sum_{i=0}^{n} f_i^*(p^*).$$

It follows from the definitions that:

$$\operatorname{Inf} \mathscr{P} = h(0) \geqslant h^{**}(0) = \operatorname{Sup} \mathscr{P}^*.$$

The non-negative number $\operatorname{Inf} \mathscr{P} - \operatorname{Sup} \mathscr{P}^*$ is called the duality gap. We are seeking to assign it an upper bound, which should be computed as easily as possible from the f_i. Moreover, this upper bound should be zero in the convex normal case.

Our starting point will be the following lemma:

Lemma 2. Graph $h^{**} \subset \operatorname{co} \sum_{i=1}^{n} \operatorname{graph} f_i$.

Proof. Recall the definition of h:

$$h(p) = \operatorname{Min} \left\{ \sum_{i=1}^{n} f_i(x_i) \,\middle|\, \sum_{i=1}^{n} x_i = p \right\}$$

$$= \operatorname{Min} \left\{ \alpha \,\middle|\, (p, \alpha) = \sum_{i=1}^{n} (x_i, \alpha_i) \quad \text{with} \quad (x_i, \alpha_i) \in \operatorname{graph} f_i \right\}.$$

It follows that:

$$\sum_{i=0}^{n} \operatorname{epi} f_i \subset \operatorname{epi} h \subset \overline{\sum_{i=0}^{n} \operatorname{epi} f_i}.$$

Recall that $\operatorname{epi} h^{**} = \overline{\operatorname{co}} \operatorname{epi} h$ (Proposition I.3.2). It follows from the definition of the graph as the lowest points on the epigraph that:

$$\operatorname{graph} h^{**} \subset \overline{\operatorname{co}} \operatorname{graph} h \subset \overline{\operatorname{co}} \sum_{i=0}^{n} \operatorname{graph} f_i.$$

Using the arguments of Lemma IX.3.3, it easily seen that each of the co graph f_i is closed, as well as their sum. Hence the result. ∎

Assume $\operatorname{Sup} \mathscr{P}^* \neq \pm\infty$. It follows from Lemma 1 that:

$$(0, \operatorname{Sup} \mathscr{P}^*) \in \operatorname{co} \sum_{i=1}^{n} \operatorname{graph} f_i.$$

We are now in a position to apply the Shapley–Folkman theorem. There exists a $(k+1)$-set $J \subset \{0, \ldots, n\}$ such that:

$$(0, \operatorname{Sup} \mathscr{P}^*) \in \sum_{i \notin J} \operatorname{graph} f_i + \sum_{i \in J} \operatorname{co} \operatorname{graph} f_i.$$

That means that there exist:

. for every $i \notin J$, a point x_i of $\mathbf{R}^k$

. for every $i \in J$, a finite subset x_{ij} of $\mathbf{R}^k$, and positive numbers α_{ij} with $\sum_j \alpha_{ij} = 1$

such that

$$0 = \sum_{i \notin J} x_i + \sum_{i \in J} \sum_j \alpha_{ij} x_{ij}$$

$$\operatorname{Sup} \mathscr{P}^* = \sum_{i \notin J} f_i(x_i) + \sum_{i \in J} \sum_j \alpha_{ij} f_i(x_{ij}).$$

Denote by x_i the barycentre $\sum_j \alpha_{ij} x_j$, for $i \in J$, so that:

$$\sum_j \alpha_{ij} f_i(x_{ij}) \geqslant f_i^{**}(x_i).$$

Putting this into the two preceding equations, we get:

$$0 = \sum_{i=0}^{n} x_i$$

$$\operatorname{Sup} \mathscr{P}^* \geqslant \sum_{i \in J} f_i(x_i) + \sum_{i \in J} f_i^{**}(x_i)$$

$$= \sum_{i=0}^{n} f_i(x_i) - \sum_{i \in J} (f_i(x_i) - f^{**}(x_i)).$$

The first equation yields $\sum_{i=0}^{n} f_i(x_i) \geqslant h(0)$, hence:

$$\operatorname{Sup} (\mathscr{P}^*) \geqslant h(0) - \sum_{i \in J} (f_i(x_i) - f_i^{**}(x_i). \tag{2}$$

It is in order to introduce a definition, before stating our final result:

Definition 1. $\alpha(f) = \operatorname{Sup}_x (f(x) - f^{**}(x))$.

Note that $\alpha(f)$ is a non-negative real number, which may take the value $+\infty$, and which does so as soon as the effective domain of f is not convex. Whenever $\alpha(f) < +\infty$ it provides a measure of the extent to which f fails to be convex; in particular, $\alpha(f) = 0$ means that f is convex.

Theorem 2. *Assume the functions f_i are lower semi-continuous and $f_i(x_i)/\|x_i\| \to \infty$ as $\|x_i\| \to \infty$. If* $\operatorname{Sup} \mathscr{P}^* \neq \pm\infty$, *we have*:

$$0 \leqslant \operatorname{Min} \mathscr{P} - \operatorname{Sup} \mathscr{P}^* \leqslant (k+1) \sup_{0 \leqslant i \leqslant n} \alpha(f_i).$$

Proof. Start from equation (2). Recall first that J has cardinality $(k+1)$:

$$\operatorname{Sup}(\mathscr{P}^*) \geqslant h(0) - (k+1) \sup_{0 \leqslant i \leqslant n} \alpha(f_i).$$

As $h(0) = \operatorname{Inf} \mathscr{P}$, this is the desired result. ∎

Theorem 2 yields an estimate of the duality gap. But only very mild non-convexities can be taken into account: the functions f_i have to have a convex domain. If, for instance, we take as f_0 the indicator of a subset K of $\mathbf{R}^k$, $\alpha(f_0)$ will be $+\infty$—and our estimation will be valueless—unless K is convex.

We shall therefore change slightly the formulation of problem ($\mathscr{P}$), to get an estimate which could be used in more general situations.

3. ESTIMATING THE DUALITY GAP: GENERAL CASE

Consider the following optimization problem:

$$(\mathscr{P}) \qquad \begin{cases} \operatorname{Inf} \sum_{i=1}^{n} f_i(x_i \\ \sum_{i=1}^{n} g_i^j(x_i) \leqslant c^j \quad \text{for} \quad 1 \leqslant j \leqslant k. \end{cases}$$

The functions f_i and g_i^j send some euclidian space E_i into $\overline{\mathbf{R}}$. For simplicity's sake again, we shall assume them to be lower semi-continuous and to satisfy:

$$\forall i, \left[f_i(x_i) + \sum_{j=1}^{k} g_i^j(x_i) \right] \Big/ \|x_i\| \to +\infty \quad \text{as} \quad \|x_i\| \to \infty.$$

Denote by g_i the mapping from E_i into $\mathbf{R}^k$ with components g_i^j. We shall use the perturbation function:

$$\Phi(x_1, \ldots, x_n; p) = \sum_{i=1}^{n} f_i(x_i) + \chi_{\mathbf{R}_+^k}\left(p + c - \sum_{i=1}^{n} g_i(x_i) \right)$$

with $p \in \mathbf{R}^k$. Hence $h(p) = \operatorname{Inf} \mathscr{P}_p$, with:

$$(\mathscr{P}_p) \qquad \begin{cases} \operatorname{Inf} \sum_{i=1}^{n} f_i(x_i) \\ \sum_{i=1}^{n} g_i^j(x_i) \leqslant c^j + p^j \quad \text{for} \quad 1 \leqslant j \leqslant k. \end{cases}$$

As in Lemma 1, it follows from the topological assumptions that h is a lower semi-continuous function on $\mathbf{R}^k$, and that problem $\mathscr{P}_p$ has an optimal solution whenever $h(p) \neq \pm\infty$. Instead of computing the dual $\mathscr{P}^*$, let us write down the Lagrangian:

$$L(x_1, \ldots, x_n; p) = \sum_{i=1}^{n} f_i(x_i) + \left\langle -p^*, \sum_{j=1}^{n} g_i(x_i) - c \right\rangle \quad \text{if} \quad -p^* \in \mathbf{R}_+^k$$

$$= +\infty \quad \text{if} \quad -p^{+j} < 0 \quad \text{for some } j.$$

Here again, our starting point will be:

$$\operatorname{Inf} \mathscr{P} = h(0) \geqslant h^{**}(0) = \operatorname{Sup} \mathscr{P}^*.$$

Using the Shapley–Folkman theorem and Definition 1, we obtain:

Theorem 3. *Assume that* $\operatorname{Sup} \mathscr{P} \neq \pm\infty$, *that the functions* f_i *and* g_c^j *are lower semi-continuous, and that*:

$$\forall i, \left[f_i(x_i) + \sum_{j=1}^{k} g_i^j(x_i) \right] \Big/ \|x_i\| \to +\infty \quad as \quad \|x_i\| \to \infty.$$

Set $\bar{p}^j = (k+1) \max_i \alpha(g_i^j)$, *for* $1 \leqslant j \leqslant k$. *Then*

$$\operatorname{Min} \mathscr{P}_{\bar{p}} - \operatorname{Sup} \mathscr{P}^* \leqslant (k+1) \max_i \alpha(f_i).$$

Proof. Any point of graph h can be written as $(p, \operatorname{Min} \mathscr{P}_p)$. By definition, there exist n points $x_i \in E_i$ such that:

$$p^j \geqslant \sum_{i=1}^{n} g_i^j(x_i) - c^j \quad \text{for} \quad 1 \leqslant j \leqslant k$$

$$\operatorname{Min} \mathscr{P}_p = \sum_{i=1}^{n} f_i(x_i).$$

That means that graph h can be decomposed into:

$$\text{graph } h \subset \sum_{i=1}^{n} C_i + (-c + \mathbf{R}_+^k, 0)$$

$$C_i = \{(g_i(x_i), f_i(x_i)) \,|\, x_i \in E_i\}.$$

Recall now that:

$$(0, \operatorname{Sup} \mathscr{P}^*) \in \text{graph } h^{**} \subset \overline{\text{co}} \text{ graph } h.$$

The hypothesis for the functions f_i and g_i^j implies that:

$$\overline{\text{co}} \text{ graph } h = \text{co graph } h.$$

Hence:

$$(0, \operatorname{Sup} \mathscr{P}^*) \in (-c + \mathbf{R}_+^k, 0) + \text{co} \sum_{i=1}^{n} C_i.$$

Using the Shapley–Folkman theorem, we can find a $(k+1)$-set $J \subset \{1, \ldots, n\}$ such that:

$$(0, \operatorname{Sup} \mathscr{P}^*) \in (-c + \mathbf{R}_+^k; 0) + \sum_{i \notin J} C_i + \sum_{i \in J} \text{co } C_i.$$

That means that there exists:

. for every $i \notin J$, some point x_i of E_i;

. for every $i \in J$, a finite subset x_{it} of E_i, and positive numbers α_{it} with $\sum_t \alpha_{it} = 1$

such that

$$0 = \epsilon - c_i + \mathbf{R}_+^k + \sum_{i \notin J} g_i(x_i) + \sum_{i \in J} \sum_t \alpha_{it} g_i(x_{ij})$$

$$\operatorname{Sup} \mathscr{P}^* = \sum_{i \notin J} f_i(x_i) + \sum_{i \in J} \sum_t \alpha_{it} f_i(x_{it}).$$

The first equation can be written in another way:

$$0 \geqslant -c^j + \sum_{i \notin J} g^j(x_i) + \sum_{i \in J} \sum_t \alpha_{it} g_i^j(x_{it}) \quad \text{for} \quad 1 \leqslant j \leqslant k.$$

Denote by x_i the barycentre of the x_{it}, for $i \in J$. By Definition 1:

$$\sum_t \alpha_{it} f_i(x_{it}) \geqslant f_i(x_i) - \alpha(f_i)$$

$$\sum_t \alpha_{it} g_i^j(x_{it}) \geqslant g_i^j(x_i) - \alpha(g_i^j).$$

Hence:

$$\sum_{i=1}^{n} f_i(x_i) - \sum_{i \in J} \alpha(f_i) \leqslant \operatorname{Sup} \mathscr{P}^*$$

$$\sum_{i=1}^{n} g_i^j(x_i) \leqslant c^j + \sum_{i \in J} \alpha(g_i^j) \quad \text{for} \quad 1 \leqslant j \leqslant k.$$

Taking into account the fact that J has cardinality $(k+1)$, we obtain:

$$\sum_{i=1}^{n} f_i(x_i) - (k+1) \max \alpha(f_i) \leqslant \operatorname{Sup} \mathscr{P}^*$$

$$\sum_{i=1}^{n} g_i^j(x_i) \leqslant c^j + (k+1) \max_i \alpha(g_i^j) \quad \text{for} \quad 1 \leqslant j \leqslant k.$$

Setting $\bar{p}^j = (k+1) \max_i \alpha(g_i^j)$, we see that the family $(x_i) \in \prod_{i=1}^{n} E_i$ satisfies the constraints of problem $\mathscr{P}_{\bar{p}}$. Hence,

$$\sum_{i=1}^{n} f_i(x_i) \geqslant \operatorname{Min} \mathscr{P}_{\bar{p}} \text{ and the result. } \blacksquare$$

Note that if $\operatorname{Min} \mathscr{P}_p \neq +\infty$ on a neighbourhood of c in $\mathbf{R}^k$, then the dual problem $\mathscr{P}^*$ has an optimal solution, so that we can replace $\operatorname{Sup} \mathscr{P}^*$ by $\operatorname{Max} \mathscr{P}^*$ in the preceding formulas.

4. THE LIAPUNOV EFFECT

Consider the following optimization problem:

$$(\mathscr{P}) \qquad \begin{cases} \operatorname{Inf} \displaystyle\int_0^1 f(t, x(t))\,\mathrm{d}t \\ \displaystyle\int_0^1 g^j(t, x(t))\,\mathrm{d}t \leqslant 0 \quad \text{for} \quad 1 \leqslant j \leqslant k. \end{cases}$$

Here f and the g^j are continuous functions from $[0, 1] \times K$ (where K is a compact metrizable space) into $\mathbf{R}$, and $x(.)$ is a measurable mapping from $[0, 1]$ into K. We assume no convexity at all.

For every $p \in \mathbf{R}^k$, consider the perturbed problem:

$$(\mathscr{P}_p) \qquad \begin{cases} \operatorname{Inf} \displaystyle\int_0^1 f(t, x(t))\,\mathrm{d}t \\ \displaystyle\int_0^1 g^j(t, x(t))\,\mathrm{d}t \leqslant p^j \quad \text{for} \quad 1 \leqslant j \leqslant k. \end{cases}$$

As usual, we shall denote by $h(p)$ the performance function, and by $L(t, x, p^*)$ the (pointwise) Lagrangian:

$$h(p) = \mathrm{Inf}\,(\mathscr{P}_p) \quad \text{for} \quad p \in \mathbf{R}^k$$

$$L(t, x, p^*) = f(t, x) + \sum_{j=1}^{k} p^{*j} g(t, x).$$

We shall make the following assumptions:

(a) $h(p)$ is finite on a neighbourhood $\mathscr{U}$ of the origin in $\mathbf{R}^k$. It follows already from our continuity assumptions that $h(p) \neq +\infty$ everywhere. Here we assume, moreover, that h does not take the value $-\infty$ near the origin, i.e. that problem $\mathscr{P}_p$ has a non-empty domain for small perturbations p.

(b) for every $p^* \in \mathbf{R}^k_+$, the set of $t \in [0, 1]$ such that the function $x \mapsto L(t, x, p^*)$ attains its minimum over K at several points, is negligible. In other words, for every $p^* \in \mathbf{R}^k_+$ and almost every $t \in [0, 1]$, there is a single point $x \in K$ where $L(t, x, p^*)$ is minimum. Note that if the functions f and g^j do not depend on t, this is tantamount to saying that they are convex. This is not so in the general case.

These assumptions are made mainly for simplicity's sake: Theorem 4 is true under much more general conditions.

Theorem 4. *Under the above assumptions, h is convex and continuous on $\mathscr{U}$. The dual problem $\mathscr{P}^*$ has an optimal solution $\bar{p}^*$ such that the measurable mapping $x(.)$ defined by:*

$$\bar{x}(t) \text{ minimizes } L(t, x, \bar{p}^*) \text{ for almost every } t$$

is an optimal solution of $(\mathscr{P})$. Hence:

$$\mathrm{Min}\, \mathscr{P} = \mathrm{Max}\, \mathscr{P}^*.$$

We shall now prove this theorem in several steps. We begin by writing down explicitly the dual problem $\mathscr{P}^*$ with respect to the given set of perturbations. Write first the Lagrangian:

$$\int_0^1 L(t, x(t), p^*)\, \mathrm{d}t$$

then minimize it with respect to $x(.)$. Using the measurable selection Theorem VIII.1.4, we see that it amounts to minimizing pointwise:

$$\underset{x(.)}{\mathrm{Min}} \int_0^1 L(t, x(t), p^*) = \int_0^1 \gamma(t, p^*)\, \mathrm{d}t$$

$$\gamma(t, p^*) = \underset{x}{\mathrm{Min}}\, L(t, x, p^*).$$

Note that the function γ is finite and continuous with respect to both variables t and p^*, and concave with respect to the variable p^*. The dual problem can now be written:

$$(\mathscr{P}^*) \qquad \begin{cases} \operatorname{Sup} \displaystyle\int_0^1 \gamma(t, p^*)\,\mathrm{d}t \\ p^{*j} \geqslant 0 \quad \text{for} \quad 1 \leqslant j \leqslant k. \end{cases}$$

We shall now approximate problem $\mathscr{P}$, by discretization. We divide [0, 1] in 2^n equal subintervals $T_i = [(i-1)2^{-n}, i2^{-n}]$, and we consider the problem ${}_n\mathscr{P}$ defined by:

$$({}_n\mathscr{P}) \qquad \begin{cases} \operatorname{Inf} \displaystyle\sum_{i=1}^{2^n} \int_{T_i} f(t, x_i)\,\mathrm{d}t \\ \displaystyle\int_{T_i} g^j(t, x_i)\,\mathrm{d}t \leqslant 0 \quad \text{for} \quad 1 \leqslant j \leqslant k \\ x_i \in K \quad \text{for} \quad 1 \leqslant i \leqslant 2^n. \end{cases}$$

Clearly:

$$\operatorname{Min} {}_n\mathscr{P} \geqslant \operatorname{Min} {}_{n+1}\mathscr{P} \geqslant \operatorname{Inf} \mathscr{P}.$$

Let us now write down the dual problem ${}_n\mathscr{P}^*$:

$$({}_n\mathscr{P}^*) \qquad \begin{cases} \operatorname{Sup} \Gamma_n(p^*) \\ p^{*j} \geqslant 0 \quad \text{for} \quad 1 \leqslant j \leqslant k \\ \Gamma_n(p^*) = \displaystyle\sum_{i=1}^{2^n} \min_{x_i} \int_{T_i} L(t, x_i, p^*)\,\mathrm{d}t. \end{cases}$$

Clearly, the functions Γ_n form a decreasing sequence, converging towards $\int_0^1 \gamma(t, .)\,\mathrm{d}t$ uniformly on every compact subset of $\mathbf{R}^k$ (use Dini's lemma, or a direct argument).

Denote by h_n the performance function of problem ${}_n\mathscr{P}$:

$$h_n(p) = \operatorname{Inf} {}_n\mathscr{P}_p.$$

Lemma 3. *As $n \to \infty$, $h_n^{**}(0)$ converges to $h^{**}(0)$, and any sequence of subgradients $p_n^* \in \partial h_n^{**}(0)$ has a cluster point $\bar{p}^* \in \partial h^{**}(0)$.*

Proof. Let $\mathscr{V}$ be a convex compact neighbourhood of the origin contained in $\mathscr{U}$. The assumptions of Theorem 4 imply that there exists an integer m such that h_m does not assume the value $+\infty$ on $\mathscr{V}$. Moreover, $\operatorname{Inf} \mathscr{P}_p \geqslant N$, where N denotes the minimum of f on the compact set $[0, 1] \times K$. Hence, for every $n \geqslant m$:

$$\forall p \in \mathscr{V}, \qquad N \leqslant h(p) \leqslant h_n(p) \leqslant h_m(p) < +\infty.$$

By Γ-regularization, for every $n \geqslant m$:

$$\forall p \in \mathscr{V}, \qquad N \leqslant h^{**}(p) \leqslant h_n^{**}(p) \leqslant h_m^{**}(p) < +\infty.$$

The functions h_n^{**} are finite-valued on a neighbourhood of the origin, hence continuous and subdifferentiable. Moreover, if $p_n^* \in \partial h_n^{**}(0)$, we must have:

$$\forall p \in \mathscr{V}, \qquad \langle p, p_n^* \rangle \leqslant h_n^{**}(p) - h_n^{**}(0).$$

We have chosen $\mathscr{V}$ compact, so that $\operatorname{Max}_{p \in \mathscr{V}} h_m^{**}(p) = M < +\infty$. Hence, for $n \geqslant m$:

$$\forall p \in \mathscr{V}, \qquad \langle p, p_n^* \rangle \leqslant M - N.$$

This proves that the subdifferentials $\partial h_n^{**}(0)$ are uniformly bounded in $\mathbf{R}^k$. But we know (Lemma III.2.4) that $\partial h_n^{**}(0)$ is the set of optimal solutions of ${}_n\mathscr{P}^*$. The problems $\mathscr{P}^*$ and ${}_n\mathscr{P}^*$, for $n \geqslant m$, thus have optimal solutions, all of which are contained in a fixed compact subset K^* of R^k. We restrict our attention to K^*, on which the functions Γ_n converge uniformly to $\int_0^1 \gamma(t, .)\,\mathrm{d}t$. Clearly:

$$h_n^{**}(0) = \operatorname*{Max}_{p^* \in K^*} \Gamma_n(p^*)$$

$$\partial h_n^{**}(0) = \{p_n^* \in K^* \mid \Gamma_n(p_n^*) \geqslant \Gamma_n(p^*), \qquad \forall p^* \in K^*\}$$

$$h^{**}(0) = \operatorname*{Max}_{p^* \in K^*} \int_0^1 \gamma(t, p^*)\,\mathrm{d}t$$

$$\partial h^{**}(0) = \left\{\bar{p}^* \in K^* \mid \int_0^1 \gamma(t, \bar{p}^*)\,\mathrm{d}t \geqslant \int_0^1 \gamma(t, p^*)\,\mathrm{d}t, \qquad \forall p^* \in K^*\right\}.$$

The result follows easily from the uniform convergence of the Γ_n and the compactness of K^*. ∎

Lemma 4. *Let p_n be an optimal solution of ${}_n\mathscr{P}^*$. There exists a family $\bar{x}_i$, $1 \leqslant i \leqslant n$, and a constant c such that:*

(a) $$\int_{T_i} L(t, \bar{x}_i, p_n^*)\,\mathrm{d}t = \min_{x \in K} \int_{T_i} L(t, x, p_n^*)\,\mathrm{d}t.$$

(b) $$\sum_{i=1}^{2} \int_{T_i} g^j(t, \bar{x}_i)\,dt \leqslant 2^{-n} c.$$

Proof. Define a mapping φ: $[0, 1]xK \to \mathbf{R}^{k+1}$ by:

$$\varphi(t, x) = (f(t, x), g^j(t, x)).$$

Denote by:

$$C_i = \left\{ \int_{T_i} \varphi(t, x)\,dt \,|\, x \in K \right\} \subset \mathbf{R}^{k+1}.$$

Problem ${}_n\mathscr{P}$ can be stated in the following way:

$$({}_n\mathscr{P}) \qquad \begin{cases} \operatorname{Inf} z^0 \\ z^1 \leqslant 0, \ldots, z^k \leqslant 0 \\ z \in \sum_{i=1}^{2^n} C_i. \end{cases}$$

Introduce the indicator function $\chi_{\Sigma C_i}$ of the set $\sum_{i=1}^{2^n} C_i$. Recall that its polar is the support function $\chi^*_{\Sigma C_i}$ of $\sum_{i=1}^{2^n} C_i$, and its bipolar $\chi^{**}_{\Sigma C_i}$ is the indicator function of the closed convex hull co $\sum_{i=1}^{2^n} C_i$. We have computed ${}_n\mathscr{P}^*$ with respect to the following perturbations (see III.1):

$$\Phi(z, p) = z^0 + \chi_{\Sigma C_i}(z) + \chi_{\mathbf{R}^k_+}(p^1 - z^1, \ldots, p^k - z^k).$$

We find easily:

$$\Phi^*(z^*, p^*) = \chi^*_{\Sigma C_i}(z^{*0} - 1, z^{*1} + p^{*1}, \ldots, z^{*k} + p^{*k}) + \chi_{\mathbf{R}^k_+}(-p^*)$$

$$\Phi^{**}(z, p) = z^0 + \chi_{\Sigma C_i}(z) + \chi_{\mathbf{R}^k_+}(p^1 - z^1, \ldots, p^k - z^k).$$

We can now state problem ${}_n\mathscr{P}^{**}$:

$$({}_n\mathscr{P}^{**}) \qquad \begin{cases} \operatorname{Inf} z^0 \\ z^1 \leqslant 0, \ldots, z^k \leqslant 0 \\ z \in \operatorname{co} \sum_{i=1}^{2^n} C_i. \end{cases}$$

Note that the convex hull of $\varphi([0, 1] \times K)$ is compact, and is the set of barycentres of probability measures bounded by $\varphi([0, 1] \times K)$. Hence:

$$\left(\int_{T_i} f(t, x)\, 2^n\, dt, \int_{T_i} g^j(t, x)\, 2^n\, dt \right) \in \operatorname{co} \varphi([0,1] \times K).$$

Denote by B the unit ball of $\mathbf{R}^{k+1}$. There exists a constant d large enough for $\varphi([0, 1]xK)$ to be contained in dB. It follows that:

$$C_i \subset 2^{-n} dB \quad \text{for} \quad 1 \leqslant i \leqslant 2^n.$$

Now let $\bar{z}$ be an optimal solution of ${}_n\mathscr{P}^{**}$ and p_n^* be an optimal solution of its dual ${}_n\mathscr{P}^*$. Then $\bar{z}$ must minimize the linear functional $z^0 + \sum_{i=1}^k p^{*j} z^j$ over co $\sum_{i=1}^{2n} C_i$, and must of course satisfy the constraints $\bar{z}^j \leqslant 0$. Using the Shapley–Folkman theorem:

$$\bar{z} = \sum_J \zeta_i + \sum_{I-J} z_i$$

where I denotes the set $\{1, 2, \ldots, 2^n\}$ and J a $(k+1)$-subset. We have $z_i \in C_i$ for $i \in I - J$ and $z_i \in \operatorname{co} C_i$ for $i \in J$, and every z_i (or ζ_i) minimizes the linear functional $z^0 + \sum_{j=1}^k p^{*j} z^j$ over C_i (or co C_i).

For $i \in J$, choose some point $z_i \in C_i$ where $z^0 + \sum_{j=1}^k p^{*j} z^j$ is minimum, and write:

$$\tilde{z} = \sum_J z_i + \sum_{I-J} z_i.$$

Clearly, $z_i = \int_{T_i} \varphi(t, \bar{x}_i)$, where $\bar{x}_i$ satisfies (a) for every i. Moreover:

$$\|\tilde{z} - \bar{z}\| \leqslant 2(k+1) \cdot 2^{-n} d.$$

Writing this inequality componentwise, we get (b). ∎

The proof of Theorem 4 follows easily. Let $\bar{p}^* \in \partial h$ be the limit of a sequence $p_n^* \in \partial h_n^{**}(0)$. Denote by $x_n(.)$ a piecewise constant mapping from $[0, 1]$ into K defined by:

$$\int_{T_i} L(t, x_n(t), p_n^*)\,\mathrm{d}t = \min_{x \in K} \int_{T_i} L(t, x, p_n^*)\,\mathrm{d}t \quad \text{for} \quad t \in T_i.$$

As $n \to \infty$ for fixed t, the sequence $x_n(t)$ has cluster points in the compact K, every one of which must minimize the limit function $x \to L(t, x, \bar{p}^*)$. By uniqueness assumption (b), we conclude that for almost every t, the sequence $x_n(t)$ converges to the unique solution $\bar{x}(t)$ in K of equation:

$$L(t, x(t), \bar{p}^*) = \min_{x \in K} L(t, x, \bar{p}^*).$$

From this definition of $\bar{x}(.)$, it follows that:

$$(\alpha) \qquad \int_0^1 L(t, \bar{x}(t), \bar{p}^*)\,\mathrm{d}t = \int_0^1 \gamma(t, \bar{p}^*)\,\mathrm{d}t$$

$$= \operatorname{Max} \mathscr{P}^* = h^{**}(0).$$

Moreover, by the Lebesgue convergence theorem:

$$\forall j, \qquad \int_0^1 g^j(t, \bar{x}(t))\,\mathrm{d}t = \lim_{n\to\infty} \int_0^1 g^j(t, x_n(t))\,\mathrm{d}t.$$

Lemma 4 states that the right-hand side goes to zero. Hence:

$$(\beta) \qquad \forall j, \qquad \int_0^1 g^j(t, \bar{x}(t))\,\mathrm{d}t \leqslant 0.$$

The result follows from (α) and (β). ■

5. COMMENTS

These results are due to the first author, partly in collaboration with J. P. Aubin (Ekeland [5], Ekeland and Aubin [1], Aubin [1]) and have been initiated by numerical experiments of Lemaréchal (I.R.I.A., 1973).

The continuous optimization problem of Section 4 plays an important role in mathematical economics, where it has been introduced by Aumann and Perles [1]. The existence of solutions and of Lagrange multipliers is related to the Lyapunov convexity theorem. We refer the reader to Berliocchi and Lasry [1] for an extensive treatment and bibliography. The approach developed here seems to be new, and has been discovered independently by Arstein ([1], [2]).

This page intentionally left blank

APPENDIX II

Non-convex Optimization Problems Depending on a Parameter

Introduction

Few general results are known concerning the existence and uniqueness of solutions of non-convex optimization problems. In the infinite dimensional case and without compactness[(1)], very few tools are available. It seems in fact that one cannot expect general existence and uniqueness results. Natural and simple examples show that many pathological situations may arise: it may happen that a continuous function bounded from below does not attain its minimum; the minimizing sequences of a problem may or may not converge; when they converge they may converge to a limit different from the solution even when such a solution exists.

The purpose of Chapters IX and X was to clarify the situation, in some cases, by introducing the relaxed form of a non-convex problem. Our purpose in this appendix is different. We will consider some families of non-convex problems depending on a parameter and prove that most often these problems are actually very regular: *for almost all values of the parameter* (in a sense to be made more precise later on), the considered non-convex problem possesses a unique solution which depends continuously of the parameter on the complement of the exceptional set.

The first results of this type seems to be that of Edelstein [1] [2] concerning farthest points (or closest points) of a non-convex set: let V be a Hilbert space and S a closed bounded set in V (not necessarily convex); then the points of S which are the farthest from some point v are solutions of the maximization problem ($\|.\|$ = norm in V)

(1) $$\sup_{u \in S} \|u - v\|.$$

Similarly the projection of v on S is a solution of the minimization problem

(2) $$\inf_{u \in S} \|u - v\|.$$

[(1)] The existence results in Chapter VII were based on compactness arguments.

Edelstein proved in [1] (or [2]) that a solution of (1) (or (2)) exists for all the v's of a dense subset of V. This result has then been extended in several directions which will be described below.

Our goal here is to establish a similar result for a family of optimization problems of the type

$$\operatorname*{Sup}_u \{F(u) + \omega(\|u - v\|)\}. \tag{3}$$

The hypotheses for V, F (a function from V into $\mathbf{R} \cup \{+\infty\}$) and ω (a function from $\mathbf{R}_+$ into $\mathbf{R}_+$) are specified in the text. The main result is given in Theorem 1.1.

1. THE MAIN RESULTS

1.1. Definitions and notations

We recall here a few definitions and notations.

(i) A Banach space V is called a strongly differentiable space (S.D.S. space) if the following property holds:

any convex continuous function F from V into $\mathbf{R} \cup \{+\infty\}$ is Frechet differentiable in a dense G_δ[1] subset of its domain

$$\operatorname{dom} F = \{u \in V \mid F(u) < +\infty\}.$$

In such a case (see Asplund [4] [5]), the mapping $u \to F'(u)$ is norm to norm continuous from the G_δ into V.

According to Trojanski [1] any reflexive Banach space is an S.D.S.

(ii) A strictly convex or rotund normed spaced is characterized by the property that

$$\|u + v\| = 2 \text{ and } \|u\| = \|v\| = 1 \text{ imply } u = v.$$

A locally uniformly normed space (L.U.R. space) is a space satisfying the following condition (see Lovaglia [1]):

$$\text{if} \quad \|u_n + u\| \to 2 \quad \text{as} \quad n \to \infty, \quad \text{and}$$

$$\|u_n\| = \|u\| = 1, \quad \text{then} \quad \|u_n - u\| \to 0.$$

(iii) Let Γ_E denote the set of even functions φ belonging to $\Gamma_0(\mathbf{R})$ and with values in $\mathbf{R}_+ = [0, +\infty[$ and such that $\varphi(0) = 0$. We consider also the following subsets of Γ_E

$$\Gamma_U = \{\varphi \in \Gamma_E \mid \varphi(t) > 0 \quad \text{for} \quad t > 0\}$$

$$\Gamma_L = \{\varphi \in \Gamma_E \mid \lim_{t \to 0} \varphi(t)/t = 0\}\cdot$$

[1] I.e. a denumerable intersection of open sets.

The functions of Γ_E are decreasing on $]-\infty, 0]$, increasing on $[0, +\infty[$ and attain their minimum at 0. We note the following property

$$\varphi \in \Gamma_U \Leftrightarrow \varphi^* \in \Gamma_L. \tag{1.1}$$

It is clear that φ is even ($\varphi \in \Gamma_E$) if and only if φ^* is even. Now for an arbitrary $\varepsilon > 0$ consider

$$\psi_\varepsilon(t) = \varepsilon|t|$$

whose conjugate is

$$\psi_\varepsilon^*(t) = 0 \quad \text{if} \quad |t| \leqslant \varepsilon, \quad +\infty \text{ otherwise.}$$

We note that

$$\varphi \notin \Gamma_U \Leftrightarrow \exists \varepsilon > 0, \qquad \varphi \leqslant \psi_\varepsilon^*$$
$$\varphi^* \notin \Gamma_L \Leftrightarrow \exists \varepsilon > 0, \qquad \psi_\varepsilon \leqslant \varphi^*.$$

Since the two relations to the right are equivalent, it is clear that $\varphi \notin \Gamma_U \Leftrightarrow \varphi^* \notin \Gamma_L$.

(iv) A characterization of Frechet differentiability of convex functions is given in Asplund [4] [5]:

a convex function F: $V \to \mathbf{R}$ is Frechet differentiable at the point u with differential $u^* \in V^*$ if and only if there exists a function $\varphi \in \Gamma_L$ such that

$$F(v) \leqslant F(u) + \langle u^*, v - u\rangle + \varphi(\|v - u\|). \tag{1.2}$$

This amounts to saying that the graph of F is below a parabolic type function centred at the point $\{u, F(u)\}$ of $V \times \mathbf{R}$.

1.2. Statement of the results

Let V be a reflexive Banach space which may nor not satisfy the following property

(1.3) If a sequence u_n converges weakly to u and $\|u_n\|$ converges to $\|u\|$, then $\|u_n - u\| \to 0$.

It is well known that any uniformly convex space is reflexive and satisfies (1.3).

We consider a l.s.c. F from V into $\mathbf{R} \cup \{-\infty\}$, which is not identically equal to $-\infty$, and a scalar even function ω which is convex continuous and strictly increasing on $[0, +\infty[$. We are interested in the family of maximization problems

$$\sup_{u \in V} \{F(u) + \omega(\|u - v\|)\}. \tag{1.4}$$

The supremum in (1.4) is denoted $F_\omega(v)$ and we assume that

(1.5) $$F_\omega(v) < +\infty \quad \text{for each} \quad v \in V,$$

and

(1.6) Every maximizing sequence of problem (1.4) is bounded.

As usual, property (1.6) holds if, for each fixed v,

$$F(u) + \omega(\|u - v\|) \to -\infty, \quad \text{as} \quad \|u\| \to +\infty, \quad u \in \operatorname{dom} F.$$

As a particular case of (1.4), we may consider a function F

$$F(u) = \begin{vmatrix} J(u), & u \in S \\ -\infty & \text{otherwise} \end{vmatrix}$$

where $S \subset V$ is closed convex and J is a l.s.c. function from S in $\mathbf{R}$. Problem (1.4) then becomes

(1.7) $$\operatorname*{Sup}_{u \in S} \{J(u) + \omega(\|u - v\|)\}.$$

The problem (1.4) is generally not convex. Although property (1.6) guarantees the existence of weakly convergent maximizing sequences, we do not have any information on the limit: by lack of a weak semi-continuity property of the functional, the limit of a weakly convergent maximizing sequence is not necessarily a solution of (1.4). For this reason a reasoning similar to that used for the proof of Proposition II.1.2 is not sufficient to get the existence of a solution of (1.4). Using completely different methods we will prove the following results due to J. Baranger and R. Temam [1] [2]:

Theorem 1.1. *Let V be a reflexive Banach space; under the preceding assumption and in particular* (1.5) (1.6), *there exists a dense G_δ subset of V such that for each v in this set, problem* (1.4) *possesses a solution $\bar{u}$,*

(1.8) $$F(\bar{u}) + \omega(\|\bar{u} - v\|) = F_\omega(v),$$

and any weakly convergent maximizing sequence of (1.4) *converges weakly to such a solution* (*or strongly if property* (1.3) *holds*).

Theorem 1.2. *The assumptions are those of Theorem* 1.1, *including* (1.3). *We assume, moreover, that V is strictly convex, that ω is everywhere differentiable with derivative ω',*

(1.9) $$\omega' > 0$$

and that

(1.10) $$\operatorname{dom} \omega^* = \mathbf{R}.^{(1)}$$

(1) This holds in particular if $\omega(s)/s \to +\infty$ as $s \to +\infty$.

Then the preceding G_δ set, denoted $\mathscr{A}$, can be chosen so that for each $v \in \mathscr{A}$, the solution $\bar{u}$ of (1.4) *exists and is unique, the mapping*

$$v \in \mathscr{A} \to \bar{u}(v) \in V,$$

being norm continuous.

Remark 1.1. Theorem 1.1 extends some results of Edelstein [1] ((1.7) with $J \equiv 0$ $\omega(s) = s^2$ and V is uniformly convex), Asplund [4] (where $\omega(s) = s^2$ and V is a L.U.R. reflexive Banach space), Baranger [2] [3] ((1.7) with V uniformly convex and $\omega(s) = |s|$), Bidaut [1] (V uniformly convex and $\omega(s) = |s|^\alpha$, $1 \leqslant \alpha < \infty$) and Zizler [1] ((1.7) with $J = 0$, $\omega(s) = |s|$). Uniqueness and continuity results appear also in Asplund [4] and Bidaut [1].

Remark 1.2. Similar results are proved by different methods for minimization problems (this amounts to replacing ω by $-\omega$). This type of results should be compared with a result of Aronszajn [1] on the differentiability, "almost everywhere" of Lipschitz functions. ∎

1.3. Proof of Theorem 1.1

With (1.5), and since F is not identical to $-\infty$, the function $F_\omega: v \to F_\omega(v)$ is defined from V into $\mathbf{R}$; this function is convex l.s.c. as an upper bound of such functions. According to Corollary I.2.5 the function F_ω is continuous on X and then Proposition I.5.2 shows that F_ω is everywhere subdifferentiable. From point (i) of Section 1.1, V is an S.D.S. space and therefore there exists a G_δ dense subset of V, denoted $\mathscr{A}$, on which F_ω is Frechet differentiable. We will prove that the conclusions of Theorems 1.1 and 1.2 hold for this set $\mathscr{A}$.

Let ξ denote a point of $\mathscr{A}$ and η be the differential $F'_\omega(\xi)$. From point (iv) of Section 1.1 we get the existence of $\varphi \in \Gamma_L$ (which depends on ξ and F_ω) such that

$$0 \leqslant F_\omega(v) - F_\omega(\xi) - \langle \eta, v - \xi \rangle \leqslant \varphi(\|v - \xi\|) \quad \text{for all} \quad v \in V. \tag{1.11}$$

Since the function $\theta_u: v \mapsto \omega(\|v - u\|)$ is convex and continuous, it is subdifferential everywhere and in particular at the point ξ. Let t be some element of this subdifferential

$$0 \leqslant \omega(\|u - v\|) - \omega(\| - \xi\|) - \langle t, v - \xi \rangle. \tag{1.12}$$

It follows from (1.11) and (1.12) that, for each $u \in V$

$$\begin{aligned} &\omega(\|u - \xi\|) + \langle t - \eta, v - \xi \rangle + F(u) - F_\omega(\xi) \\ &\leqslant F_\omega(v) - F_\omega(\xi) - \langle \eta, v - \xi \rangle \leqslant \varphi(\|v - \xi\|) \end{aligned}$$

and hence

$$\omega(\|u-\xi\|)+\langle t-\eta, v-\xi\rangle-\varphi(\|v-\xi\|)-F_\omega(\xi)\leqslant -F(u).$$

Taking the supremum with respect to v of the left-hand side, we obtain

(1.13) $\omega(\|u-\xi\|)-F_\omega(\xi)+\varphi^*(\|t-\eta\|_*)\leqslant F(u)$, for all $u\in V$,

where $\varphi^*\in\Gamma_U$ (see (1.1)).

Now let u_n be a maximizing sequence of problem (1.4) where $v=\xi$:

$$\omega(\|u_n-\xi\|)+F(u_n)\to F_\omega(\xi),\quad\text{as}\quad n\to\infty.$$

For each n, let $t_n\in\partial\theta_{u_n}(\xi)$. Relation (1.13) gives

$$0\leqslant\varphi^*(\|t_n-\eta\|_*)\leqslant F_\omega(\xi)-F(u_n)-\omega(\|u_n-\xi\|).$$

Hence $\varphi^*(\|t_n-\eta\|_*)\to 0$, as $n\to\infty$, and since $\varphi^*\in\Gamma_U$, this implies

(1.14) $$\|t_n-\eta\|_*\to 0,\quad\text{as}\quad n\to\infty.$$

From (1.6) the sequence u_n is bounded. Extracting a subsequence, we may assume that u_n is weakly convergent in V to some limit $\bar u$. We write (1.12) with u replaced by u_n and v by $u_n-\bar u+\xi$:

$$\omega(\|\bar u-\xi\|)\geqslant\omega(\|u_n-\xi\|)+\langle t_n, u_n-\bar u\rangle.$$

As $n\to\infty$, $\langle t_n, u_n-\bar u\rangle\to 0$ and we get by weak l.s.c.:

$$\begin{aligned}\omega(\|\bar u-\xi\|)&\geqslant\overline{\lim}\,\omega(\|u_n-\xi\|)\\&\geqslant\omega(\overline{\lim}\,\|u_n-\xi\|)\\&\geqslant\text{(as }\omega\text{ is increasing)}\\&\geqslant\omega(\underline{\lim}\,\|u_n-\xi\|)\\&\geqslant\omega(\|\bar u-\xi\|).\end{aligned}$$

Since ω is strictly increasing, all these inequalities are in fact equalities,

$$\|\bar u-\xi\|=\underline{\lim}\,\|u_n-\xi\|=\overline{\lim}\,\|u_n-\xi\|,$$

the sequence $\|u_n-\xi\|$ converges to $\|\bar u-\xi\|$. If condition (1.3) is satisfied this means that u_n converges strongly in V to $\bar u$.

Finally we will see that, in all cases, $\bar u$ is a solution of (1.4). Indeed

$$\begin{aligned}F_\omega(\xi)&=\lim_{n\to\infty}\{\omega(\|u_n-\xi\|)+F(u_n)\}\\&=\omega(\|\bar u-\xi\|)+\lim_{n\to\infty}F(u_n)\\&\leqslant\text{(by u.s.c.)}\\&\leqslant\omega(\|\bar u-\xi\|)+F(\bar u)=F_\omega(\xi).\end{aligned}$$

The maximizing sequence u_n is completely arbitrary, and Theorem 1.1 is thus established. ■

1.4. Proof of Theorem 1.2

We continue the proof of Theorem 1.1 with the same notations and in particular the same set $\mathscr{A}$. As mentioned before, we infer from (1.3) that u_n converges to $\bar{u}$ strongly in V.

We write (1.12) with u and t replaced by u_n and t_n:

$$\omega(\|u_n - v\|) \geqslant \omega(\|u_n - \xi\|) + \langle t_n, v - \xi\rangle.$$

At the limit

$$\omega(\|\bar{u} - v\|) \geqslant \omega(\|\bar{u} - \xi\|) + \langle \eta, v - \xi\rangle.$$

Setting $w = v - \bar{u}$, we get

$$\omega(\|w\|) \geqslant \omega(\|\bar{u} - \xi\|) + \langle \eta, w - (\xi - \bar{u})\rangle, \qquad \forall w \in V, \tag{1.15}$$

which amounts to saying

$$\eta \in \partial(\omega \circ \|.\|)(\xi - \bar{u}), \tag{1.16}$$

where $\sigma = \omega \circ \|.\|$ denotes the function

$$w \mapsto \omega(\|w\|).$$

Relation (1.16) is equivalent to

$$\xi - \bar{u} \in \partial\sigma^*(\eta)$$

or

$$\bar{u} \in \xi - \partial\sigma^*(\eta). \tag{1.17}$$

We recall that $\eta = F'_\omega(\xi)$ and that (point (i), Section 1.1), the mapping $\xi \to \eta = F'_\omega(\xi)$ is norm continuous on $\mathscr{A}$. Theorem 1.2 is thus proved if we demonstrate these two points

(1.18) $\partial\sigma^*$ is reduced to one point, i.e., σ^* is everywhere Gâteaux-differentiable.

(1.19) $\eta \to (\sigma^*)'(\eta)$ is norm continuous.

We may assume without loss of generality that $\omega(0) = 0$. Then $\omega(t) > 0$ for $t > 0$, i.e., $\omega \in \Gamma_U$ and $\omega^* \in \Gamma_L$.

Lemma 1.1. *ω^* is continuously differentiable on* $\mathbf{R}$ *and* $(\omega^*)'(0) = 0$.

Proof. ω^* is convex and continuous on $\overset{0}{\widehat{\operatorname{dom} \omega^*}}$ and thus on $\mathbf{R}$ due to (1.10).

The function ω^* is everywhere subdifferentiable; $\partial\omega^*(t)$ is reduced to one point for each $t \in \mathbf{R}$, for if $x_1, x_2 \in \partial\omega^*(t)$ then

$$\omega(x) \geqslant \omega(x_i) + t(x - x_i), \qquad \forall x \in \mathbf{R}, \quad i = 1, 2.$$

In particular

$$\omega(x_2) \geqslant \omega(x_1) + t(x_2 - x_1)$$
$$\omega(x_1) \geqslant \omega(x_2) + t(x_1 - x_2)$$

hence

$$\omega(x_2) - \omega(x_1) = t(x_2 - x_1)$$

but ω is strictly convex and this equality is impossible unless $x_1 = x_2$. The function ω^* is differentiable on $\mathbf{R}$.

The function $(\omega^*)'$ is increasing as ω^* is convex; thus $(\omega^*)'$ is continuous except perhaps at a countable set of discontinuity points; since ω^* is everywhere differentiable, $(\omega^*)'$ has no discontinuity.

Lemma 1.2. $\partial\sigma^*(v) = \left| \begin{array}{l} (\omega^*)'(\|v\|) . \phi(v) \quad \textit{if} \quad v \in V^*, v \neq 0, \\ 0 \quad \textit{if} \quad v = 0 (\in V^*), \end{array} \right.$

ϕ the duality mapping from V^* into V with gauge 1.

The mapping $v \to (\sigma^*)'(v)$ is norm continuous.

Proof. If $v = 0$,

$$\frac{\omega^*(\|\lambda w^*\|)}{\lambda} \to 0 \quad \text{as} \quad \lambda \to 0, \quad (\omega^* \in \Gamma_L)$$

and then $(\sigma^*)'(0) = 0$.

If $v \neq 0$, the result follows from the chain rule of differentiation. For the continuity we observe that $(\omega^*)'$ is continuous and also the gauge function ϕ except at the origin where $\phi(v)$ remains bounded while $(\omega^*)'(\|v\|) \to 0$. ∎

The points (1.18) (1.19) are proved and Theorem 1.2 is demonstrated. ∎

2. APPLICATIONS AND EXAMPLES

2.1. Farthest points and projection on a non-convex set

Let V be as in Theorem 1.2 and S a closed bounded subset of V. We set $\omega(s) = s^2$. The point of S the farthest from v is solution of

$$\sup_{u \in S} \|u - v\|^2. \tag{2.1}$$

For all the v of a dense G_δ subset of V, there exists a unique farthest point $\bar{u}$ and the mapping $v \to \bar{u}(v)$ is norm continuous.

A similar result can be deduced from Theorem 1.2 for closest points if X is a Hilbert space. The projection of v on S is a solution of

$$\inf_{u \in S} \|u - v\|^2. \tag{2.2}$$

Since

$$\|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2 - \|u + v\|^2,$$

the problem (2.2) can be written as

$$2\|v\|^2 \sup_{u \in S} \{\|u + v\|^2 - 2\|u\|^2\}. \tag{2.3}$$

Setting $F(u) = -2\|u\|^2$ if $u \in S$ and $-\infty$ otherwise, $\omega(s) = s^2$, we deduce from Theorem 1.2 that, *for all the v's of a dense G_δ subset of V, the projection of v on S exists, is unique and depends continuously on v.*

A similar result can be proved for more general spaces using a different method.

2.2. Examples in calculus of variation

A variational problem

We consider here a problem of the type considered in Chapter X.

Let f denote a scalar continuous function such that

$$f(\xi)/\xi^2 \to c \neq 0 \quad \text{as} \quad |\xi| \to \infty.$$

Let Ω be an open bounded set of $\mathbf{R}^n$. For each $u \in H_0^1(\Omega)$, the function

$$x \mapsto f(\operatorname{grad} u(x))$$

is summable and the functional

$$u \in H_0^1(\Omega) \mapsto F(u) = \int_{\Omega} f(\operatorname{grad} u(x))\, dx \in \mathbf{R}$$

is continuous on $H_0^1(\Omega)$ (Proposition IV.1.1).

Let g belong to $L^2(\Omega)$; a natural variational problem is the following

$$\inf_{u \in H_0^1(\Omega)} \left\{ \int_{\Omega} f(\operatorname{grad} u(x))\, dx - \int_{\Omega} g(x)\, u(x)\, dx \right\}. \tag{2.4}$$

Theorem 1.2 applies and shows that for all the g's of a dense G_δ subset, the problem (2.4) possesses a unique solution which depends continuously on g.

In order to show this we introduce $v \in H_0^1(\Omega)$, satisfying $-\Delta v = g/2$ in Ω. Then

$$-\int_\Omega gu \, dx = -2((v, u))$$

$((.,.))$ = scalar product of $H_0^1(\Omega)$ and (2.4) is nothing else than

$$(2.5) \qquad \|v\|^2 - \operatorname*{Sup}_{u \in H_0^1(\Omega)} \left\{ \|u + v\|^2 - \int_\Omega [f(\nabla u(x)) + |\nabla u(x)|^2] \, dx \right\}.$$

It becomes now very easy to check the assumptions of Theorem 1.2.

An application to optimal control

Let Ω denote again an open bounded set in $\mathbf{R}^n$. For each given function $u \in L^2(\Omega)$, with

$$(2.6) \qquad 0 < \alpha \leqslant u(x) \leqslant \beta \quad \text{a.e.}$$

there exists a unique function $y = y(u)$ in $H_0^1(\Omega)$ such that

$$(2.7) \qquad \sum_{i=1}^n \frac{\partial}{\partial x_i} \left(u \frac{\partial y}{\partial x_i} \right) = g(\in L^2(\Omega), \text{ given}).$$

We then consider the following control problem: for a given $y_d \in H_0^1(\Omega)$ find a measurable function u satisfying (2.6) which maximizes

$$\|y(u) - y_d\|$$

($\|.\|$ = norm in $H_0^1(\Omega)$).

This problem has no solution in general, see for instance Lions [4] and Murat [1].

Let ω be as in Theorem 1.2. For each $\varepsilon > 0$ and $v \in L^2(\Omega)$, we consider the following perturbation problem

$$(2.8) \qquad \operatorname*{Sup}_{\alpha \leqslant u \leqslant \beta} \{\|y(u) - y_d\| + \varepsilon\omega(|u - v|_{L^2(\Omega)})\}.$$

Theorem 1.2 asserts that for all $\varepsilon > 0$, for all the v's of a dense G_δ subset of $L^2(\Omega)$, the maximum in (2.8) is attained by a unique function $\bar{u}$. ■

Comments

CHAPTER I

After a very brief summary of the properties of topological vector spaces Chapter I develops the broad outline of the theory of convex functions as set out in the work of FENCHEL, MOREAU and ROCKAFELLAR.

The main works of reference on this subject are FENCHEL [2], MOREAU [1] and ROCKAFELLAR [4], where a comprehensive bibliography will naturally be found.

The concept of convex conjugate functions is due to FENCHEL [1] [2] and this concept was further developed by BRØENSTED and MOREAU to cover the case of infinite dimension. Initially, FENCHEL only examined finite functions defined on subsets. MOREAU introduced functions which take the value $+\infty$ and which are defined over the whole space. Here we have followed MOREAU's presentation.

Much of ROCKAFELLAR's work concerns the subdifferentiality of convex functions, which is only developed here in a very restricted sense. Nor have we proceeded with the study of subgradient mappings (which are in general multi-valued). Among those works which treat the subdifferentials of convex functions, we single out that of ROCKAFELLAR [13] which completely characterizes these operators by introducing the concept of m-cyclic monotone operators.

Theorem 6.2 concerning ε subdifferentials plays an important role in the study of subdifferentiability of convex functions: it is due to BRØNDSTED and ROCKAFELLAR [1]. In what follows it is important for the study of minimizing sequences, especially in Chapters V and VII.

Theorem 6.1, which is new, is a partial extension of Theorem 6.2 to non-convex cases and is of interest in the study of Euler equations (Chap. X, §4). For further generalizations and applications, we refer to EKELAND [2].

CHAPTER II

The criteria for the existence and uniqueness of the minimum of a convex function, given in Proposition 1.2, are completely standard and are sufficient for the applications we have in mind. Other general theorems are given in J. CEA [1], for example.

The characterization of the solutions of a convex optimization problem by inequalities such as (2.2) or (2.3) is also standard.

Section 3 gives an existence criterion for the solution of variational inequalities of "elliptic" type. The notion of variational inequality is due to G. STAMPACCHIA [5]; the proof given here of Theorem 3.1, which uses the proximity mappings of J. J. MOREAU [2], is due to H. BREZIS [1]. For a deeper study of variational inequalities, see H. BREZIS [1] [2], F. BROWDER [3] [4], J. L. LIONS [3], G. STAMPACCHIA [4], E. ZARANTONELLA [2] and the bibliography contained therein.

CHAPTER III

We are indebted to R. T. Rockafellar for the concept of duality in convex optimization as developed here. It gives a unified presentation of two different approaches to duality:

First, a presentation using conjugate convex functions which was introduced by W. FENCHEL [2] and then extended by R. T. ROCKAFELLAR (see especially [5] [6] [7]).

In [2], W. FENCHEL studied problems of the type

$$\operatorname*{Inf}_{u \in V} \{ F(u) + G(u) \},$$

where F and $G \in \Gamma_0(V)$ and in [7], R. T. ROCKAFELLAR considered in a more general fashion problems of the type

$$\operatorname*{Inf}_{u \in V} \{ F(u) + G(\Lambda u) \},$$

where $F \in \Gamma_0(V)$, $G \in \Gamma_0(Y)$, $\Lambda \in \mathscr{L}(V, Y)$.

A slight extension of ROCKAFELLAR [7] is given in R. TEMAM [2]. This class of problems is dealt with as the special cases of Section 4.

Secondly, duality is presented using Lagrangian functions which is, classically, the standpoint adopted in Mathematical Economics and which is the subject of an abundant literature; see, among others, ARROW, HURWICZ and UZAWA [1], and the theorem of KUHN and TUCKER [1]. The corresponding problems are dealt with as special cases in Section 5.

R. T. ROCKAFELLAR's unified presentation (*cf.* Sections 1 and 2) which uses the very elegant concept of perturbed problems and conjugate convex functionals, was introduced in [8] and developed for the finite dimensional case in [4].

The stability criterion given in Proposition 2.3 plays an essential part in the application of the calculus of variations to variational problems.(1) This criterion is sufficient in the majority of cases; see however the examples of Sections 3 and 4 of Chapter VII. The concept of extremality relations is also essential for application to the calculus of variations.

After the special cases of Sections 4 and 5, the concept of a generalized solution given in Section 6.1, is useful for certain problems which have no solution in the classical sense. Section 6.2 describes one of the ways of introducing duality into variational inequalities.

CHAPTER IV

Duality in the calculus of variations has been appearing in the literature in various forms for a long time. In particular, a heuristic presentation was given by R. COURANT and D. HILBERT [1] using the LEGENDRE transformation: the dual problems can be written down for several variational problems but it is not always possible to establish the relationship between primal and dual by this method. Duality as given in Chapter III enables us to obtain more information: the equality inf = sup and the extremality relations.

After the simple examples in Section 2, the examples in Section 3 all arise out of diverse applications; MOSSOLOV's problem (no. 3.1) is linked with the problem of the

(1) In fact we could give a speedier proof in a more direct context.

flow of Bingham fluids studied by G. DUVAUT and J. L. LIONS [1] from the standpoint of no. III.6.2 and IV.3.1. The problems in no. 3.2 and 3.3 arise in filtering theory, while the problem of 3.4 is the classical problem of elasto-plastic torsion of a rod; the problem of the existence of a solution to the dual problem in $L^1(\Omega)^n$ [(1)] remains open: this solution when linked to the constraints has an important solution in mechanics; the problem has been solved by H. BREZIS [3] when f is constant.

Section 4 takes up and completes a result of H. BREZIS [2]. Section 5 points out a general method for applying duality to the calculus of variations. It would be interesting to state the exact meaning of the solution of the dual problem, when the primal problem does not possess one.

CHAPTER V

This chapter takes up and develops the work of one of the authors. The application of duality to the problem of minimal hypersurfaces was presented in R. TEMAM [2] and developed in R. TEMAM [3]. Section 3 which develops the application of duality to a class of variational problems of minimal hypersurface type takes up, with the same presentation, a part of R. TEMAM [3]. Sections 1 and 2 give the same result in a more simple manner but only for the single special case of minimal hypersurfaces: this simpler presentation was suggested to one of the authors by S. AGMON during a seminar.

It was not considered useful to prove explicitly the theorem of singular perturbation used in Chapter II.I This proof, which is given in R. TEMAM [3], requires techniques for partial differential equations which are very different from those used in this book.

Sections 1, 2 and 3 are the source of many unsolved problems:

Find necessary or necessary and sufficient conditions for given ϕ and Ω such that the generalized solution of the problem is a real solution. Some partial results have been given in Section 2 for minimal hypersurfaces and in Section 3 for more general problems.

Describe in particular the set of the ϕ such that the problem admits of a real solution.

Study the correspondence between the trace of ϕ over $\partial\Omega$ and the trace over $\partial\Omega$ of the generalized solution.

The extension of these techniques to other problems connected with minimal hypersurfaces seems easy: hypersurfaces with obstacles, hypersurfaces with given mean curvature, or the capillary problem. These problems were not studied here to avoid too long an exposition; we have similarly not discussed the numerical approximation of minimal hypersurfaces, for which we refer the reader to VANENDE [1] and for a more systematic treatment to C. JOURON [1].

The existence result for problems of NEUMANN type studied in no. 4.1 was stated in R. TEMAM [4]; like the problems in Section 3, these problems are only coercive in a non-reflexive space, the space $W^{1,1}(\Omega)$, and it is noteworthy that a single change in boundary conditions leads to a far more simple situation from the point of view of the existence of solutions of the primal problem.

To be sure, in this chapter we have not tackled the many other aspects of the problem of minimal hypersurfaces and for these we refer the reader to F. J. ALMGREN [1]

[(1)] A solution function and not merely a solution of Radon measure as in no. 3.5.

[2], E. BOMBIERI, E. DE GIORGI and M. MIRANDA [1], P. CONCUS and R. FINN [1] [2], E. DE GIORGI [1] [2], M. EMMER [1], H. FEDERER [1], R. FINN [1] [2], W. H. FLEMMING [1], D. GILBARG [1], H. JENKINS and J. SERRIN [1], D. H. KINDERLEHRER [1] to [4], O. A. LADYZENSKAY and N. N. URALCEVA [1], J. LERAY [1], M. MIRANDA [1] to [8], C. B. MARREY [1], P. P. MOSSOLOV [1], J. C. C. NITSCHE [1] to [5], R. OSSERMAN [1], E. RADO [1], E. SANTI [1], J. SERRIN [1] [2], C. STAMPACCHIA [3] [6], L. C. YOUNG [1] [2] [3], among many other references.

CHAPTER VI

It seemed useful to us to develop concisely the minimax approach to duality which is simpler on occasion and which corresponds to current practice in economics.

The results given in Section 1 are completely standard and can be found for example in S. KARLIN [1], except for properties 1.6 and 1.7 which are due to J. L. LIONS. The same is also true of the majority of the results in Section 2, Proposition 2.1 being essentially the theorem of KY-FAN [1] and SION [1]. Propositions 2.2 and 2.4 are natural extensions to the infinite dimensional case. The proof of Proposition 2.1 presented here is due to H. BREZIS; moreover H. BREZIS, L. NIREMBERG and G. STAMPACCHIA give in [1] an extension of the KY-FAN–SION theorem.

The study of saddle points is also developed in LEMAIRE [1] [2] where many applications to saddle point problems in partial differential equations are considered.

Section 3 shows how the existence results for saddle-points are useful in duality; the connection between duality and saddle points is examined completely in MACLINDEN [1].

CHAPTER VII

The numerical algorithms described in Sections 1 and 2 are those of UZAWA and of ARROW and HURWICZ respectively; see ARROW and HURWICZ [1], ARROW, HURWICZ and UZAWA [1]. The convergence proof given here is that of GLOWINSKI, LIONS and TRÉMOLIÈRES [1] where great use is made of duality and of this type of algorithm for the numerical solution of variational inequalities arising in mechanics and physics.

We are indebted to D. EDERY [1] and R. TEMAM [5] for the example of Section 3, and to J. MOSSINO [1] for that in Section 4. For other numerical applications of duality, see among others A. AUSLENDER [1] [2], J. CÉA, R. GLOWINSKY and J. C. NEDELEC [1], HAUGAZEAU [1], M. FORTIN [1], M. FORTIN, R. PEYRET and R. TEMAM [1], M. FREMOND [1] [3], B. MARTINET [1] where the author studies the systematic generation of algorithms using duality, R. TEMAM [5], R. TRÉMOLIÈRES [1]. The most systematic applications of duality to optimal control theory are developed in W. HEINS and K. S. MITTER [1], J. L. LIONS [4], J. MOSSINO [1], R. T. ROCKAFELLAR [1] [2] [3] [4] [10] [11], UHLENBECK [1]. The application of duality to mechanics given in Section 5 are taken from M. FREMOND [2].

Some applications of duality to the approximation of functions are given in J. L. JOLY and P. J. LAURENT [1], P. J. LAURENT [1].

CHAPTER VIII

Convex normal integrands were introduced by R. T. ROCKAFELLAR in two articles ([1] [2]) where he studied in depth their polar properties. The general concept of a non-convex normal integrand (Def. 1.1) is due to H. BERLIOCCHI and J. M. LASRY [1] (see also [3] and [4]) who characterized it completely; we have borrowed from them Theorem 1.2 and Proposition 1.4. They also showed that if a normal integrand $f(x,a)$ is convex in a then it is a convex integrand in the sense of ROCKAFELLAR and that if the indicator function of $G \subset \Omega \times \mathbf{R}^p$ is a normal integrand, then G is the graph of a measurable multi-function. We recover in this way the now standard theory of measurable multi-functions to which belong the section theorems of type 1.6 (K. KURATOWSKI and C. RYLL-NARDZEWSKI [1], C. CASTAING [1] [2]; for the convex case see also R. T. ROCKAFELLAR [3]).

The proof of Theorem 1.11 (criteria of weak compactness in L^1) may be found in the book by P. A. MEYER [1], Chapter II, Section 2.

The second section examines the existence of the solution of a problem in the calculus of variations for a convex and coercive integrand. The result obtained (Th. 2.7) is standard (L. CESARI [1] [2] [3]), originating with TONELLI. C. CASTAING appears to have been the first person to have used MAZUR's lemma for this type of problem; the proof of Theorem 2.3 which we give here was discovered independently by L. CESARI.

CHAPTERS IX AND X

In 1900, D. HILBERT stated his twentieth problem in the following way: "ob nicht jedes reguläre Variationsproblem eine Lösung besitzt, sobald hinsichtlich der gegebenen Grenzbedingungen gewisse Annahmen erfüllt sind und nötigenfalls der Begriff der Lösung eine sinngemässe Erweiterung erfährt."

In their work, L. YOUNG [1] [2] [3] [4] and E. MACSHANE [1] [2] solved this problem, completely for the one-dimensional case, and partially for dimension $n > 1$. With the advent of control theory, the problem was carried over to this new field with analogous results (R. GAMKRELIDZE [1] [2], J. WARGA [2], A. GHOUILAHOURI [1]). A special mention is owed to L. CESARI who, in a series of papers, examined the existence and relaxation of solutions of problems of optimal control for systems governed by partial differential equations (see [1] [2] [3] and the bibliography in these papers). His method is to reduce the state equations locally to first order systems by a transformation due to DIEUDONNE and RASHEVSKY. The relaxed problems are obtained by suitably completing the set of controls: thus they are all of the second form ($\mathscr{P}\mathscr{R}$).

Simultaneously, however, A. IOFFE and V. TIHOMIROV [1], following an idea of N. BOGOLYUBOV [1], introduced the problem ($\mathscr{P}\mathscr{R}$) without establishing all its properties. The standpoint is entirely different: the set of controls is unchanged, only the criterion is made convex with respect to the control. The synthesis of these two points of view was made by I. EKELAND [1] who introduced the use of convex analysis. His results have been generalized by H. BERLIOCCHI and J. M. LASRY [3] (see also [1] and [2]) who instead appealed to measure theory.

Finally, what is relaxation?

Starting with a problem for which there is no solution, we seek to formulate explicitly a second problem which has the same values as the first and whose optimal solutions are precisely the cluster points of the minimizing sequences of the first.

We set out essentially two construction procedures which lead to two equivalent problems $(\mathscr{P}\mathscr{R})$ and $(\mathscr{P}\mathscr{R}')$ (see Chsp. IX; see also I. EKELAND [1], Sections 9 and 10).

After this brief historical survey, we return to look at Chapter IX in detail. The results of Section 1 were established by I. EKELAND [1], by appealing to Lyapunov's theorem: the framework is more general but the method remains the same. For the case of convex normal integrands we refer to R. T. ROCKAFELLAR [1] [2]. Finally, the results of Section 4 are due to H. BERLIOCCHI and J. M. LASRY [3]. In I. EKELAND's paper [1], a special case is considered and it is shown how to extend the results to problems of optimal control of the type:

$$\left| \begin{array}{l} \text{minimize} \displaystyle\int_\Omega f(x, u(x), p(x))\, \mathrm{d}x \\ Au(x) = \gamma(x, p(x)), \text{ where } A = \mathscr{G}^{-1} \end{array} \right.$$

Passing on to Chapter X, we are indebted to L. C. YOUNG [2] [3] for Theorem 1.2. In these papers he considers completely the case $W^{1,\infty}$ for a continuous integrand (Prop. 2.13 and Th. 3.6 with $\alpha = \beta = \infty$). The general case seems to be new as do the polarity results (Cor. 3.5).

APPENDICES I & II

The appendices develop recent results related to some other aspects of non-convex optimization.

Appendix I gives an estimation of the duality gap in some finite-dimensional cases, i.e., an estimation of the difference $\inf \mathscr{P} - \sup \mathscr{P}^*$, which is positive.

Appendix II deals with families of non-convex problems depending on a parameter. It is shown that for almost all values of the parameter the non-convex problem has a unique solution.

Bibliography

F. J. ALMGREN Jr. [1], The theory of varifold. A variational calculus in the large for the k dimensional area integrand, *Notes multigraphiées*, Princeton, 1965.

[2], Geometric measure theory and elliptic variational problems, *Actes du Congrès Intern. des Math.*, Nice 1970, **2**, 813–819.

N. ARONSZAJN [1], *Théorèmes de différentiabilité*, Cours de 3ème cycle, Université de Paris, 1974.

K. J. ARROW and L. HURWICZ [1], Reduction of constrained maxima to saddle-point problems, *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* (J. Neyman, editor), Univ. of California Press, 1956, vol. V, pp. 1–20.

K. J. ARROW, L. HURWICZ and H. UZAWA [1], *Studies in Linear and Non Linear Programming*, Stanford University Press, 1958.

[2], Constraint qualifications in maximization problems, *Naval Res. Logist. Quart.*, **8**, 1967, 175–191.

Z. ARTSTEIN [1], The MacShane Integral and the theory of set-valued mappings (to be published).

Z. ARTSTEIN and R. BURNS [2], A law of large numbers for subsets of vector spaces (to be published).

E. ASPLUND [1], Farthest points in reflexive locally uniformly rotund Banach spaces, *Isr. J. of Math.*, **4**, 1966, 213–216.

[2], Averaged norms, *Isr. J. Math.*, **5**, 1967, 227–233.

[3], Frechet differentiability of convex functions, *Acta Math.*, **121**, 1968, 31–47.

[4], Topics in the theory of convex functions, in *Theory and Applications of Monotone Operators*, A. Ghizetti, Italie, 1969, pp. 1–33.

[5], Positivity of duality mappings, *Bull. Am. Math. Soc.*, **73**, 1967, 200–203.

E. ASPLUND and R. T. ROCKAFELLAR [1], Gradients of convex functions, *Trans. A.M.S.*, **139**, 1969, 443–467.

J. P. AUBIN [1], *Optimization Game Theory* (to be published).

J. P. AUBIN and I. EKELAND [1], Estimates of the duality gap in non-convex programming, *Math. Operations Research* (to be published).

R. J. AUMANN and M. PERLES [1], A variational problem arising in economics, *J. Math. An. Appl.*, **11**, 1965, 488–503.

A. AUSLENDER [1], *Problèmes de minimax via l'analyse convexe et les inégalités variationnelles: théorie et applications*, Lecture Notes in Economics and Math. Syst, Springer Verlag, 1972.

[2], *Méthodes numériques pour la résolution de problèmes d'optimisation avec contraintes*, Thèse, Faculté des Sciences de Grenoble, 1970.

M. S. BAOUENDI [1], Sur une classe d'opérateurs elliptiques dégénérés, *Bull. Soc. Math. Fr.*, **95**, 1967, 45–87.

J. BARANGER [1], Existence de solutions pour des problèmes d'optimisation non convexe, *C. R. Acad. Sci., Paris*, **274**, 1973, 307–311.

[2] Existence de solutions pour des problèmes d'optimisation non convexes, *J. Math. Pures Appl.* **52**, 1973, 377–406.

[3], Thèse, Université de Grenoble, 1973.

J. BARANGER and R. TEMAM [1], Nonconvex optimization problems depending on a parameter, *Siam J. Control*, **13**, 1974,

[2], Problèmes d'optimisation non convexes dependants d'un paramètre, in *Proceedings of a Conference on Convex Analysis, held at Saint Pierre de Chartreuse January 1974*, Springer-Verlag, 1975.

L. D. BERKOVITZ and H. POLLARD [1], A non classical variational problem arising from an optimal filter problem, I, *Arch. Rat. Mech. Anal.*, **26**, 1967, 281–304.

[2], A non classical variational problem arising from an optimal filter problem, II, *Arch. Rat. Mech. Analysis*, **38**, 1972, 161–172.

[3], A variational problem related to an optimal filter problem with self-correlated noise, *Trans. Am. Math. Soc.*, **142**, 1969, 153–175.

[4] Addenda to A variational problem related to an optimal filter problem with self-correlated noise, *Trans. Am. Math. Soc.*, **157**, 1971; 499–504.

H. BERLIOCCHI and J. M. LASRY [1], Intégrandes normales et mesures paramètrées en calcul des variations, *Bull. S.M.F.*, **101**, 1973, 129–184.

[2], Intégrandes normales et mesures paramètrées en calcul des variations, *C. R. Acad. Sci. Paris*, **274**, série A, 1971, 839–842.

[3], Diverses applications des mesures paramètrées, *C. R. Acad. Sci. Paris*, **274**, série A, 1972, 1623–1626.

S. BERNSTEIN [1], Sur la Généralisation du problème de Dirichlet, *Math. Am.*, **69**, 1910, 82–136.

[2], Sur les équations de calcul des variations. *Ann. Sc. Ecole Norm. Sup.*, **29**, 1912, 431–485.

M. F. BIDAUT [1], Thèse, Université de Paris, 1973.

E. BISHOP and R. R. PHELPS, The support functionals of a convex set, *Proc. Symp. Pure Math.*, 7, 1962, 27–35, AMS publ.

N. BOGOLYUBOV [1], Sur quelques méthodes nouvelles dans le calcul des variations, *Annali di matematica pura e applicata*, ser., **4**, no. 7, 1930, 249–271.

E. BOMBIERI, E. DE GIORGI and M. MIRANDA [1], Una Maggiorazione *a priori* relative alle ipersurficï minimali non parametriche, *Arch. Rat. Mech. Anal.*, **32**, 1969, 255–267.

N. BOURBAKI [1], *Éléments de Mathématiques*, Livre III, *Topologie Générale*, Hermann, Paris, 1964.

[2], *Éléments de Mathématiques*, Livre V, *Espaces vectoriels topologiques*, Hermann, Paris, 1974.

H. BREZIS, L. NIRENBERG and G. STAMPACCHIA [1], A remark on Ky-Fan's minimax dualité. *Ann. de l'Inst. Fourier*, **18**, I, 1968, 115–175.

[2], Problèmes unilatéraux, *Journ. de Math. Pures et Appliquées* (to be published).

[3], Multiplicateur de Lagrange en torsion élasto-plastique, *Arch. Rat. Mech. Anal.* (to be published).

[4], Fonctions duales relativement à une forme bilinéaire. *C. R. Acad. Sci., Paris*, **264**, série A, 284–286.

H. Brezis, L. Nirenberg and G. Stampacchia [1], A remark on Ky-Fan's minimax principle, *Bolletino della Unione Matematica Italina*, **6** (4), 1972, 293–300.

H. Brezis and G. Stampacchia [1], Sur la régularité de la solution d'inéquations elliptiques, *Bull. Soc. Math. Fr.*, **96**, 1968, 153–180.

F. Brezzi [1], *On the Existence, Uniqueness and Approximation of Saddle-point Problems Arising from Lagrangian Multipliers*, Laboratorio di Analisi Numerica del CNR, Universita di Pavia, Italy, 1974.

A. Brøndsted and R. T. Rockafellar [1], On the subdifferentiability of convex functions, *Proceedings AMS*, **16**, 1965, 605–611.

F. Browder [1], On the Fredholm alternative for nonlinear operators, *Bulletin AMS*, **76**, 1970, 993–998.

[2], Normal solvability and the Fredholm alternative for mappings into infinite dimensional manifolds, *J. Func. An.*, **8**, 1971, 250–274.

[3], Non linear operators and non linear equations of evolution in Banach spaces, in *Proc. Symposium on Non Linear Functional Analysis*, Chicago, 1968; A.M.S., vol.

[4], Editor, *Proc. Symposium on Non Linear Functional Analysis*, Chicago, 1968; A.M.S.,

[5], Normal solvability for nonlinear mappings into Banach spaces, *Bull. Am. Math. Soc.*, **77**, 1971, 73–77.

R. Burns, see Z. Artstein and R. Burns.

C. Castaing [1], *Sur les multi-applications mesurables*, thèse Caen, 1967.

[2], Sur les multi-applications mesurables, *Revue Française d'Informatique et de Recherche Opérationnelle*, **1**, 1967, 91–126, Dunod.

C. Castaing and M. Valadier [1], Sur les équations différentielles multivoques, *Revue d'Informatique et de Recherche Opérationnelle*, **16**, 1969, 3–16.

J. Céa [1], *Optimisation, Théorie et Algorithmes*, Dunod, Paris, 1971.

[2], Remarques sur la dualité dans les inéquations variationnelles (to be published).

J. Céa, R. Glowinski and J. C. Nedelec [1], Dualité et problème aux dérivées partielles (to be published).

L. Cesari [1], Existence theorems for weak and usual optimal solutions in Lagrange problems with unilateral constraints, I–II, *Transactions A.M.S.*, **124** (3), 1966, 369–430.

[2], Multidimensional Lagrange problems of optimization in a fixed domain, and an application to a problem of magnetohydrodynamics, *Archive for Rational Mechanics and Analysis*, **29**, 1968, 81–104.

[3], Existence theorems for problems of optimization, *Actes du Congrès International des Mathématiciens*, Nice, 1970, vol. 3, pp. 157–161.

P. Concus and R. Finn [1], On the behaviour of a capillary surface in a wedge, *Proc. Nat. Acad. Sc.*, **63**, 1969, 292–299.

[2], On a class of Capillary surfaces *J. d'An. Math.*, 1970, 65–70.

R. Courant [1], *Dirichlet's Principle*, Interscience Publishers, New York, 1950.

R. Courant and D. Hilbert [1], *Mathematical Methods of Physics*, Interscience Publishers, New York, 1953, 2 vols.

E. de Giorgi [1], Sulla differenziabilita e l'analiticà delle estremali degli integrali multipli regolari, *Memorie Acc. Sci. Torina*, S. 3, **3**, 1957, 25–43.

[2], *Nouveaux résultats dans la théorie des hypersurfaces minima*, Conférences faites au Collège de France, juin 1970.

[3], *Teoremi di semi continuita nel calcolo dello variazioni*, Seminaire, Istituto Nazionale di Alta Matematica, Rome 1968–1969.

G. DUVAUT and J. L. LIONS [1], *Les inéquations en mécanique et en physique*, Dunod, Paris, 1972.

M. EDELSTEIN [1], Farthest points of sets in uniformly convex Banach spaces, *Isr. J. Math.*, 4(3), 1966, 171–176.

[2], On nearest points of sets in uniformly Convex Banach Spaces *J. Lond. Math. Soc.*, **43**, 1968, 375–377.

D. EDERY [1], Report C.E.A., Saclay, France, 1972.

I. EKELAND [1], Sur le contrôle optimal de systèmes gouvernés par des équations elliptiques, *Journal of Functional Analysis*, **9**(1), 1972, 1–62.

[2], On the variational principle, *J. Math. An. Appl.*, **47** (2), 1974, 324–353.

[3], Remarques sur les problèmes variationnels, *C. R. Acad. Sci., Paris*, **275**, 1972, 1057–1059.

[4], Remarques sur les problèmes variationnels, *C. R. Acad. Sci., Paris*, **276**, 1973, 1347–1348.

[5], Une estimation a priori en programmation non convexe, *C. R. Acad. Sci., Paris*, 1974, 149–151.

I. EKELAND and J. P. AUBIN, see J. P. AUBIN and I. EKELAND.

M. EMMER [1], *Existenza, unicita e regolarita delle Superficie di Equilibro nei Capillari*, Univ. di Ferrara, 1971.

H. FEDERER [1], *Geometric Measure Theory*, Springer-Verlag, Berlin, 1969.

H. FEDERER and W. H. FLEMING [1], Normal and Integral currents, *Ann. Math.*, **72**, 1960.

W. FENCHEL [1], On Conjugate Convex Functions, *Canad. J. Math.*, **1**, 1949, 73–77.

[2], *Convex Cones, Sets and Functions*, Notes de Cours polycopiées, Princeton University, 1951.

[3], A remark on convex sets and polarity, *Medd. Lunds. Univ. mat. Sem.* (Supplement volume, 1952), 82–89.

R. FINN [1], New estimates for equations of minimal surface type, *Arch. Rat. Mech. Anal.*, **14**, 1963, 337–375.

[2], Remarks relevant to minimal surfaces and to surfaces of constant mean curvature, *J. d'An. Math.*, **14**, 1965, 139–160.

R. FINN, see P. Concus and R. Finn.

W. H. FLEMING, see H. Federer and W. H. Fleming.

M. FORTIN [1], *Calcul numerique des écoulements de fluides de Bingham et des fluides newtoniens incompressibles par la méthode des éléments finis*, Thèse, Université de Paris, 1972.

M. FORTIN, R. PEYRET and R. TEMAM [1], Calcul des écoulements d'un fluide visqueux incompressible, *Proceedings of the 2nd Int. Conf. on Num. Meth. in Fluid Mech. Lecture Notes in Physiques*, **8**, 337–342, Springer Verlag, 1971.

[2], Résolution numérique des équations de Navier-Stokes pour un fluide incompressible, *Jour. de Méc.*, **10**(3), 1971, 357–390.

A. FOUGERES [1], Thèse Université de Paris, 1972, to be published.

M. FRÉMOND [1], *Étude des Structures visco-élastiques stratifiées soumises à des charges harmoniques et de solides élastiques reposant sur des structures*, Thèse, Université de Paris, 1971.

[2], Formulation duales des énergies potentielles et complémentaires. Applidation à la méthode des éléments finis, *C. R. Acad. Sci., Paris*, série A, **273**, 775–777.

[3], *Sur la méthode frontale* (to be published).

E. Gagliardo [1], Caratterizzazioni delle trace sulla frontiera relative alcune classi di funzioni in *n* variabili, *Rend. Semin. Mat. Padova*, **27**, 1957, 284–305.

R. Gamkrelidze [1], Régimes optimaux glissants, *D.A.N. S.S.S.R.*, **134** (6), 1962, 106–128 (in Russian); *Soviet math. dokl.* **3**, 1962, 559–561 (in English).

[2], On some extremal problem in the theory of differential equations, *Journal S.I.A.M. Control*, 1965, 106–128.

R. Gamkrelidze and G. Kharatishvili [1], Conditions nécessaires du premier ordre dans les problèmes d'extremum, *Actes du Congrès international des Mathématiciens*, Nice, 1970, vol. 3, pp. 169–176.

A. Ghouila-Houri [1], Généralisation de la notion de commande d'un système guidable. *Revue d'Informatique et de Recherche Opérationnelle*, **4**, 1967, 7–32, Dunod.

D. Gilbarg [1], Boundary value problems for non linear elliptic equations in *n* variables, *Non-linear Problems*, University of Wisconsin Press, 1963, pp. 151–159.

E. de Giorgi, see E. Bombieri, E. de Giorgi and M. Miranda.

E. Giusti [1], Boundary behavior of nonparametric minimal surfaces, *Indiana Univ. Math. J.*, **22**, 1972.

R. Glowinski, see J. Céa, R. Glowinski and J. C. Nédelec.

R. Glowinski, J. L. Lions, R. Trémolières. Book to be published.

A. Grothendieck, *Espaces vectoriels topologiques*, Publication de la Société de Mathématiques de Sao Paulo, Brésil.

A. Ioffe and V. Tihomirov [1], Extension des problèmes en calcul des variations, *Trudy Moskovskovo matematiceskovo obcestva*, **18**, 1968, 188–246 (in Russian).

Ph. Hartman and G. Stampacchia [1], On some non linear elliptic differential functional equations, *Acta Math.*, **115**, 1966, 271–310.

Y. Haugazeau [1], Sur la minimisation dd formes quadratiques avec contraintes, *C. R. Acad. Sci., Paris*, **264**, 1967, 322–324.

W. Heins and K. S. Mitter [1], *Conjugate Convex Functions, Duality and Optimal Problems (I) (II) (III)*.

D. Hilbert, see R. Courant and D. Hilbert.

E. Hope [1], Über den funktionalen, insbesondere den Analytischen Charakter, der Lösungen elliptischer Differential-gleichungen zweiter Ordnung, *Math. Z.*, **34**, 1931, 194–233.

L. Hurwicz, see K. J. Arrow and L. Hurwicz; K. J. Arrow, L. Hurwicz and H. Uzawa.

H. Jenkins and J. Serrin [1], Variational Problems of Minimal Surface type I, II, III, *Arch. Rat. Mech. Anal.*, **12**, 1963, 185–212; **21**, 1966, 321–342; **29**, 1968, 304–322.

[2], The Dirichlet problem for the minimal surface equation in higher dimensions, *Journ, Reine und Angewandte Math.*, **229**, 1968, 170–187.

J. L. Joly and P. J. Laurent [1], Stability and duality in convex minimization problems, *R.I.R.O.*, **2**, 1971, 3–42.

C. Jouron [1], Approximation des surfaces minima, *Arch. Rat. Mech. Anal.* (to be published).

S. KARLIN [1], *Mathematical Methods and Theory in Games, Programming, and Economics*, Addison-Wesley, London, 1959, 2 volumes.

G. KHARATISHVILI, see R. Gamkrelidzé and G. Kharatishvili.

D. H. KINDERLEHRER [1], The coincidence set of solutions of certain variational inequalities, *Arch. Rat. Mech. Anal.*, **40**, 1971, 231–250.

[2], The regularity of the solution to a certain variational inequality, *Proc. Symp. Pure Math., A.M.S.*, **23**.

[3], Variational inequalities with lower dimensional obstacles, *Isr. Journ. Math.*, **10**, 1971, 339–348.

[4], Minimal surfaces whose boundary contains a pike, *J. Math. Mech.*, 1973.

J. J. KOHN and L. NIRENBERG [1], Degenerate elliptic parabolic equations of second order, *Comm. Pure Appl. Math.*, **20**, 1967, 797–892.

M. A. KRASNOSEL'SKII [1], *Topological Methods in the Theory of Non Linear Integral Equations*, Pergamon Press, London, 1964.

M. A. KRASNOSEL'SKII and RUTICKI [1], *Convex functions and Orlicz Spaces*, North Holland.

W. KRELLE, see H. P. Künzi and W. Krelle.

H. W. KUHN and A. W. TUCKER [1], Non linear programming, *Proceedings of the Second Berkeley Symposium on Mathematics, Statistics, and Probability*, Univ of Calif. Press, 1950.

[2], Editor, Linear inequalities and related systems, *Annals of Mathematics Studies*, no. 38, Princeton University Press, 1956.

H. P. KUNZI and W. KRELLE [1], *Non linear Programming*, Blaisdell Publishing Company, Waltham, Massachusetts, 1966.

K. KURATOWSKI and C. RYLL-NARDZAWSKI [1], A general theorem on selectors. *Bulletin de l'Académie polonaise des Sciences*, **13**, 1965, 273–411.

KY-FAN [1], Minimax theorems, *Proc. Mat. Acad. Sc.*, **39**, 1953, 42–47.

O. A. LADYZENSKAYA and N. N. URALCEVA [1], Local estimates for gradients of solutions of non uniformly elliptic and parabolic equations, *Comm. Pure Appl. Math.*, **18**, 1970, 677–703.

[2], *Équations aux dérivées partielles de type elliptique*, Dunod, Paris, 1968.

J. M. LASRY, see H. Berliocchi and J. M. Lasry.

P. J. LAURENT [1], *Approximation et Optimisation*, Hermann, Paris, 1972.

P. J. LAURENT, see J. L. Joly and P. J. Laurent.

B. LEMAIRE [1], Thesis, Paris 1970, Université de Montpellier II, Publication no. 89.

[2], Saddle point problems in partial differential equations and aoplications to linear quadratic differential games, *Annali Scuola Norm. Sup. Pisa*, **27**, 1973, 105–160.

J. LERAY [1], Discussion d'un problème de Dirichlet, *J. Math. Pures Appl.*, **18**, 1939, 249–284.

H. LEWY and G. STAMPACCHIA [1], On the regularity on the solution of a variational inequality, *Comm. Pure Appl. Math.*, **22**, 1969, 153–188.

[2], On the smoothness of superharmonics which solve a minimum problem, *Journ. d'An. Math.*, **23**, 1970, 227–236.

[3], On Existence and smoothness of solutions of some non-coercive variational inequalities, *Arch. Rat. Mech. Anal.*, **41**, 1971, 241–253.

A. LICHNEWSKY [1], Principe du maximum local et solutions généralisées de problèmes du types hypersurfaces minimales, *Bull. Soc. Math. Fr.*, 1974.

[2], Sur le comportement au bord des solutions généralisées du problème non paramétrique des surfaces minimales, *J. Math. Pures Appl.*, **53**, 1974, 397–425.

J. L. LIONS [1], *Equations différentielles opérationnelles et problèmes aux limites*, Springer-Verlag, Berlin, 1961.

[2], *Problèmes aux limites dans les équations aux dérivées partielles*, Presses de l'Université de Montréal, 1962.

[3], *Quelques Méthodes de résolution des problèmes aux limites non linéaires*, Dunod and Gauthier-Villars, Paris, 1969.

[4], *Contrôle optimal de systèmes gouvernés par des équations aux Dérivées Partielles*, Dunod and Gauthier-Villars, Paris, 1968.

J. L. LIONS, see G. Duvaut and J. L. Lions; R. Glowinski, J. L. Lions and R. Trémolières.

J. L. LIONS and E. MAGENES [1], *Problèmes aux limites non homogènes*, Dunod, Paris, 1968, 3 volumes.

A. R. LOVAGLIA [1], Locally uniformly convex Banach spaces, *Trans. Am. Math. Soc.*, **78**, 1955, 225–238.

L. MACLINDEN [1], *Minimax Problems, Saddle Functions and Duality*. Thesis, Univ. of Washington, Seattle, 1971.

E. MACSHANE [1], Generalized curves, *Duke mathematical Journal*, **6**, 1940, 513–536.

[2], *Necessary conditions in the generalized-curve problem of the calculus of variations, Duke mathematical Journal*, **7**, 1940, 1–27.

[3], Extension of range of functions, *Bulletin A.M.S.*, **40**, 1934, 837–842.

B. MARTINET [1], *Algorithmes pour la résolution de problèmes d'optimisation et de minimax*, Thèse, Université de Grenoble II, 1972.

P. A. MEYER [1], *Probabilité et potentiel*, Hermann, Paris, 1966.

V. P. MIASNIKOV, see P. P. Mossolov and V. P. Miasnikov.

G. J. MINTY [1], On a monotonicity method for the solution of non linear equations in Banach spaces, *Proc. Nat. Acad. Sc., U.S.A.*, **50**, 1963, 1038–1042.

M. MIRANDA [1], Un teorama di Extenza e unicità per il problema dell area minima in *n* variabili, *Ann. Scuola Norm. Sup. Pisa*, **19**, 1965, 233–249.

[2], Diseguaglianze di Sobolev sulle ipersurfici minimali, *Rend. Sem. Uni. di Padova*, 69–79.

[3], Comportamento delle successioni convergenti di frontiere minimali, *Rend. Semin. Univ. di Padova*, 1967.

[4], Un principio di massimo forte per le frontiere minimali e una sua applicazione alla resoluzione del problema al contorno per l'equazione delle superficiadi area minima, *Rend. Semin. Univ. di Padova.*

[5], Frontiere minimali con ostacoli, *Ann. Univ. di Ferrara*, Sez VII, **16**, 1971, 29–37.

[6], El problem de Dirichlet para la ecuacion de las superficies de area minima, *Seminario Matematico di Barcelona*, 1971.

[7], Nouveaux résultats pour les hypersurfaces Minima, *Actes du Congrès Intern. des Math.*, Nice 1970, **2**, 853–858.

[8], *Minimal hypersurfaces*, Summer Institute on P. D. E., Berkeley, August 1971.

M. MIRANDA, see E. Bombieri, E. de Giorgi and M. Miranda.

K. S. MITTER, see W. Heins and K. S. Mitter.

J. J. MOREAU [1], Fonctionnelles convexes, *Séminaire Equations aux dérivées partielles*, Collège de France, 1966.
[2], Proximité et dualité dans un espace hilbertien, *Bull. Soc. Math. France*, **93**, 1965, 273–299.
[3], Weak and strong solutions of dual problems, in *Contributions to Non Linear Functional Analysis* (E. Zarantonello, ed.), Academic Press, 1971.

C. B. MORREY [1], *Multiple Integrals in the Calculus of Variations*, Springer-Verlag, Berlin, 1966.

U. MOSCO [1], A remark on a theorem of F. E. Browder, *J. Math. Anal. Appl.*, **20**, 1967, 90–93.
[2], Approximation of the solutions of some variational inequalities, *Ann. Scuol. Norm. Sup. Pisa*, **21**, 1967, 373–394.
[3], Perturbation of Variational inequalities, in *Proc. Symp. on Non Linear Functional Analysis*, vol. 18 (F. E. Browder, ed.), 1968.
[4], Convergence of convex sets and solutions of variational inequalities (to be published).
[5], *Cours du C.I.M.E.*, Italie, 1971.

J. MOSSINO [1], An application of duality to distributed optimal control problems with constraints on the control and the state, *J. Math. Anal. Appl.* (to be published).
[2], Approximation numérique de problèmes de contrôle optimal avec contrainte sur le contrôle et sur l'état, *Calcolo* (to be published).

P. P. MOSSOLOV [1], Sur une fonctionnelle liée aux surfaces de courbure moyenne fixée, *Matematik Sbornik U.S.S.R.*, **78**, (120) 1969, 51–64.

P. P. MOSSOLOV and V. P. MIASNIKOV [1], Variational methods in the theory of the fluidity of a viscous-plastic medium, *P.M.M.*, **29**, 1965, 468–492.

F. MURAT [1], Un contre exemple pour le problème du contrôle dans les coefficients, *C. R. Acad. Sci., Paris*, série A, **273**, 708–711.

L. NIRENBERG, see J. J. Kohn and L. Nirenberg; H. Brezis, L. Nirenberg and G. Stampacchia.

J. C. C. NITSCHE [1], Variational problems with inequalities as boundary conditions or how to fashion a cheap hat for Giacometti's brother, *Arch. Rat. Mech. Anal.*, **35**, 1969, 83–113.
[2], On the non-solvability of Dirichlet's problem for the minimal surface equation, *J. Math. Mech.*, **14**, 1965, 779–788.
[3], On new results in the theory of minimal surfaces, *Bull. Ann. Math. Soc.*, **71** (2), 1965, 195–270.
[4], Elementary proof of Bernstein's theorem on minimal surfaces, *Ann. Math.*, 1957.
[5], A Book, to appear.

R. OSSERMAN [1], Minimal varieties, *Bull. A.M.S.*, 1969, 1092–1120.

M. C. PÉLISSIER [1], Thesis, University of Paris South, 1975.

M. C. PÉLISSIER and L. REYNAUD [1], Etude d'un modèle mathématique d'écoulement de glacier, *C. R. Acad. Sci., Paris*, série A, 1975.

R. PEYRET, see M. Fortin, R. Peyret and R. Temam.

E. POLAK [1], *Computational Methods in Optimization: A Unified Approach*, Academic Press, New York, 1971.

H. POLLARD, see L. D. BERKOVITZ and H. POLLARD.

E. RADO [1], On the problem of plateau, *Eg. d. Math.*, Berlin, 1933.

L. REYNAUD, see M. C. Pélissier and L. Reynaud.

R. T. ROCKAFELLAR [1], Integrals which are convex functionals, *Pacific J. Math.*, **24** (3), 1968, 525–539.

[2], Convex integral functionals, II, *Pacific J. Math.*, **39**, 1971, 439–469.

[3], Measurable dependence of convex sets and functions on parameters, *J. Math. Analysis Applic.*, **28**, 1969, 4–25.

[4], *Convex Analysis*, Princeton University Press, 1970.

[5], Extension of Fenchel's duality theorem for convex functions, *Duke Math. Journ.*, **33**, 1960, 81–90.

[6], Minimax theorems and conjugate saddle functions, *Math. Scand.*, **14**, 1964, 151–173.

[7], Duality and stability in extremum problems involving convex functions, *Pac. J. Math.*, **21**, 1967, 167–187.

[8], Convex functions and duality in optimization problems and dynamics, *Lect. Notes in Oper. Res. and Math. Ec.*, vol. II, Springer, Berlin, 1969.

[9], On the maximal monotonicity of subdifferential mappings, *Pac. J. Math.*, 1970, 209–216.

[10], Existence and duality theorems for convex problems of Bolza, *Trans. Amer. Math. Soc.*, **159**, 1971.

[11], Conjugate convex functions in optimal control and the calculus of variations, *J. Math. Analysis Applic.*, **32**, 1970, 174–222.

[12], Convex integral functionals and duality, in *Contributions to Non Linear Functional Analysis* (E. Zarantonello, ed.), Academic Press, 1971.

[13], Characterization of the subdifferentials of convex functions, *Pac. J. Math.*, **17**, 1966, 497–510.

R. T. ROCKAFELLAR, see E. Asplund and R. T. Rockafellar; A. Brøndsted and R. T. Rockafellar.

RUTICKI, see Krasnosel'skii and Rutick.

C. RYLL-NARDZEWSKI, see K. Kuratowski and C. Ryll-Nardzewski.

E. SANTI [1], Sul problema al contorno per l'equazione delle superfici di area minima su donmini limatiti qualunque, *Ann. Univ. Rerrara*, Sez VII, **17**, 13–26.

B. SCHEURER [1], Regularité des solutions de problèmes variationnels relaxés, thèse 3eme cycle, Orsay 1974.

L. SCHWARTZ [1], *Théorie des Distributions*, Hermann, Paris, 1967.

J. SERRIN [1], The problem of Dirichlet for quasilinear elliptic differential equations with many independent variables, *Phil. Trans. Royal Soc. London*, **A264**, 1969, 413–496.

[2], Boundary curvatures and the solvability of Dirichlet's problem, *Actes du Congrès Inter. des Math.*, Nice, 1970, **2**, 867–875.

J. SERRIN, see H. Jenkins and J. Serrin.

M. SION [1], On general min-max theorems, *Pac. J. Math.*, **8**, 1958, 171–176.

M. SLATER [1], *Lagrange Multiplier Revisited: A Contribution to non linear Programming*. Santa Monica, Calif., The Rand Corporation, report R. M. 676, 1951.

V. L. SMULLIAN [1], Sur la dérivabilité de la norme dans l'espace de Banach, *Dokl. Ak. Nauk SSSR*, **37**, 1940, 643–648.

S. L. SOBOLEV [1], Sur les problèmes aux limites pour les équations polyharmoniques, *Mat. Sbornik*, **44** (2), 1937, 467–500.

[2], *Applications de l'analyse fonctionelle aux équations de la physique Mathématique*, Leningrad, 1950 (English translation: Amer. Math. Soc.).

G. Stampacchia [1], *Équations elliptiques du second ordre à coefficients discontinus*, Presses de l'Université de Montréal.

[2], Regularity of solutions of some variational inequalities, in *Proceedings Symp. Non Linear Functional Analysis*, A.M.S., vol. 18, part I, 1968, pp. 271–281.

[3], On some regular multiple integral problems in the calculus of variations, *Comm. Pure Appl. Math.*, **16**, 1963, 383–421.

[4], Variational inequalities, *Actes du Congrès Int. des Math.*, Nice 1970, **2**, 877–883.

[5], Formes bilinéaires coercitives sur les ensembles convexes, *C. R. Acad. Sci., Paris*, **258**, 1964, 4413–4416.

[6], Recent results in the theory of variational inequalities.

G. Stampacchia, see H. Brézis and G. Stampacchia; H. Brézis, I. Nirenberg and G. Stampacchia; Ph. Hartman and G. Stampacchia; H. Lewy and G. Stampacchia.

R. Temam [1], Remarques sur la dualité en calcul des variations, *C. R. Acad Sci., Paris*, **270**, 1970, 754–757.

[2], Solutions généralisées de certains problèmes de calcul des variations, *C. R. Acad Sci., Paris*, **271**, 1970, 1116–1119.

[3], Solutions généralisées de certaines équations du type hypersurfaces minima, *Arch. Rat. Mech. Anal.*, **44**, 1971, 121–156.

[4], Existence et unicité de solutions pour des problèmes du type de Neuman, coercifs dans des espaces non reflexifs, *C. R. Acad. Sci. Paris*, **273**, serie A, 1971, 609–611.

[5], Sur un problème non linéaire aux équations de Maxwell, *Actes du Congrès d'Analyse Numérique*, Rome, 1972. Editions Zanichelli and Academic Press.

[6], Cours de 3e cycle. Université de Paris-Sud, Orsay, 1970–1971 and 1971–1972.

R. Temam, see M. Fortin, R. Peyret and R. Temam; J. Baranger and R. Temam.

R. Trémolières [1], *Algorithmes pour la résolution de problèmes posés sous la forme min-max f. Cas non linéaire*, Report no. 6909, I.R.I.A., Rocquencourt, France, 1969.

R. Trémolières, see R. Glowinski, J. L. Lion and R. Trémolières.

S. L. Troyanski [1], On locally uniformly convex and differentiable norms in certain non separable Banach spaces, *Studia Math.*, **37**, 1971, 173–180.

A. W. Tucker, see H. W. Kuhn and A. W. Tucker.

N. N. Uralceva, see O. A. Ladyzenskaya and N. N. Urelceva.

H. Uzawa, see K. J. Arrow, L. Hurwicz and H. Uzawa.

M. Valadier, see C. Castaing and M. Valadier.

S. Vanende [1], Thèse de 3e cycle. Université de Paris, 1971.

J. Warga [1], Relaxed variational problems, *J. Math. Analysis Applic.* **4**, 1962, 111–128.

[2], *Optimal Control of Differential and Functional Equations*, Academic Press, 1972.

L. C. Young [1], Generalized curves and the existence of an attained absolute minimum in the calculus of variations, *C. R. Soc. Sci. Lett. Varsovie*, classe III, **30**, 1937, 212–234.

[2], Generalized surfaces in the calculus of variations I, *Ann. Math.*, **43**, (1), 1942, 84–103.

[3], Generalized surfaces in the calculus of variations, II, *Ann. Math.* **43**, (3), 1942, 530–544.
[4], *Lectures on the Calculus of Variations and Optimal Control Theory*, W. B. Saunders, Philadelphia, Pa, 1969.
E. H. ZARANTONELLO [1], The closure of the numerical range contains the spectrum, *Bull. A.M.S.*, **70**, 1964, 781–787.
[2], Editor, *Contributions to Non-Linear Functional Analysis*, Academic Press, 1971.
V. ZIZLER [1], On some extremal problems in Banach spaces, *Math. Scand.*, **32**, 1973.

ADDITIONAL REFERENCES AND COMMENTS TO THE CLASSICS EDITION

We provide hereafter a few references related to the new developments in the field mentioned in the Preface to the Classics Edition; as usual, the interested reader can also consult the references quoted in these books and articles. We follow the order in this preface.

1. Concerning duality in plasticity related problems, see, e.g.,

G. STRANG and R. TEMAM, Duality and relaxation in the variational problems of plasticity, *J. Mécanique*, **19**, 1980, 493–527.
P. SUQUET, Sur les équations de la plasticité: existence et regularité des solutions, *J. Mécanique*, **20**, 1981, 3–39.
R. TEMAM, *Mathematical problems in plasticity*, Gauthier-Villars, Paris 1985.

Concerning nonconvex problems, see, e.g.,

B. DACOROGNA, *Direct methods in the calculus of variations*, Applied Mathematical Sciences, 78, Springer-Verlag, New York, 1989.
B. DACOROGNA, *Weak continuity and weak lower semicontinuity of nonlinear functionals*, Lecture Notes in Mathematics, 922. Springer-Verlag, New York, 1982.

2. Concerning the time dependent minimal surfaces and related problems, see, e.g.,

P. MARCELLINI and K. MILLER, Elliptic versus parabolic regularization for the equation of prescribed mean curvature, *J. Diff. Equ.*, **137**, 1997, 1–53.
A. LICHNEWSKY and R. TEMAM, Pseudo-solutions of the time-dependent minimal surface problem, *J. Diff. Equ.*, **30**, 1978, 340–364.
V. OLIKER and N. N. URALTSEVA, Evolution of nonparametric surfaces with speed depending on curvature. II. The mean curvature case, *Comm. Pure Appl. Math.*, **46**, 1993, 97–135.
V. OLIKER and N. N. URALTSEVA, Long time behavior of flows moving by mean curvature (English), Nonlinear evolution equations, 163–170, *Amer. Math. Soc. Transl.* Ser., 2, **164**, *Amer. Math. Soc.*, Providence, RI, 1995.

3. Application of the minimax theorem to robust control appears, e.g., in

T. BEWLEY, R. TEMAM and M. ZIANE, A general framework for robust control in fluid mechanics, submitted to *Physica D*.

4. Concerning the work of the school of E. De Giorgi, the subject is now so rich, we cannot even think of producing a bibliography of it; see, however, the articles dedicated to his memory and appearing in the *Annali della Scuola Normale Superiore de Pisa* in 1997–1999.

Index