****

Introduction to Variational Methods in Partial Differential Equations and Applications

A Summer Course at Michigan State University
(Math 890, Summer 2013)

by

BAISHENG YAN

Department of Mathematics
Michigan State University
yan@math.msu.edu

Contents

Introduction 1

A. Motivating Examples 1

B. Application to Some Physical Problems 4

C. Plan of Lectures 5

Chapter 1. Preliminaries 7

§1.1. Banach Spaces 7

§1.2. Bounded Linear Operators 11

§1.3. Weak Convergence and Compact Operators 13

§1.4. Spectral Theory for Linear Compact Operators 15

§1.5. Some Useful Results in Nonlinear Functional Analysis 17

Chapter 2. Sobolev Spaces 28

§2.1. Weak Derivatives and Sobolev Spaces 28

§2.2. Approximations and Extensions 31

§2.3. Sobolev Imbedding Theorems 44

§2.4. Additional Properties 51

Chapter 3. Second-Order Linear Elliptic PDEs in Divergence Form 55

§3.1. Second-Order PDEs in Divergence Form 55

§3.2. The Lax-Milgram Theorem 58

§3.3. Gårding's Estimates and Existence Theory 59

§3.4. Regularity of Weak Solutions 65

§3.5. Symmetric Elliptic Operators and Eigenvalue Problems 68

Chapter 4. Variational Methods for Nonlinear PDEs 71

§4.1. Variational Problems 71

§4.2. Multiple Integrals in the Calculus of Variations 72

§4.3. Direct Method for Minimization 77

§4.4. Minimization with Constraints 79

§4.5. Mountain Pass Theorem 90
§4.6. Nonexistence and Radial Symmetry 93

Chapter 5. Weak Lower Semicontinuity on Sobolev Spaces 99
§5.1. The Convex Case 99
§5.2. Morrey's Quasiconvexity 101
§5.3. Properties of Quasiconvex Functions 106
§5.4. Polyconvex Functions and Null-Lagrangians 112
§5.5. Existence in Nonlinear Elasticity 120
§5.6. Relaxation Principle and Existence for Nonconvex Problems 123

Bibliography 129

Index 131

Introduction

This course is intended to give an introduction to some important variational methods for certain problems in partial differential equations (PDE) and applications. It is suitable for graduate students with some knowledge of partial differential equations.

A. Motivating Examples

Variational methods provide a solid basis for the existence theory of PDE and other applied problems. They are the extension of methods of finding extreme values and crtical points in Calculus. We use some examples to introduce the main content of the course.

Example 1 – Dirichlet's Principle. The starting example of variational method for PDE is the Dirichlet principle for Laplace's equation:

$$\Delta u = 0, \quad u|_{\partial\Omega} = f,$$

where Ω is a given bounded domain in $\mathbb{R}^n$ and f is a given function on the boundary of Ω. This principle states that any *classical* solution u of this problem minimizes the Dirichlet integral:

$$I(u) = \int_\Omega |\nabla u(x)|^2 \, dx$$

among all smooth functions taking f on the boundary $\partial\Omega$. Therefore, in order to solve the problem, one tries to find a minimizer of the functional I among the mentioned class of smooth functions. One of the most important methods for such a minimization problem is the **direct method** of the calculus of variations, which originates from the **Weierstrass theorem**. By such a method, we take a minimizing sequence $\{u_j\}$ in the given class; *i.e.,*

$$\lim_{j\to\infty} I(u_j) = \inf I(u);$$

the infimum here is taken over all u in the given class. If this infimum is finite, then we know that each component of $\{\nabla u_j\}$ is a bounded sequence in $L^2(\Omega)$. The **weak convergence** theorem implies that there exist subsequence $\{\nabla u_{j_k}\}$ and $G = (g_1, \cdots, g_n) \in L^2(\Omega)$ such that

$$\nabla u_{j_k} \rightharpoonup G \quad \text{weakly in } L^2(\Omega) \text{ and hence } \textstyle\int_\Omega |G|^2 \, dx \le \inf I(u).$$

In fact, ∇u_{j_k} even converges to G *strongly* in $L^2(\Omega)$; that is,

$$\lim_{k\to\infty} \|\nabla u_{j_k} - G\|_{L^2(\Omega)} = 0.$$

(**Proof:** Let $\mu = \inf I(u)$. Since $\frac{u_{j_k}+u_{j_m}}{2}$ is in the class, one has $\int_\Omega |\nabla(\frac{u_{j_k}+u_{j_m}}{2})|^2\,dx \geq \mu$. This implies

$$\int_\Omega (|\nabla u_{j_k}|^2 + 2\nabla u_{j_k}\cdot\nabla u_{j_m} + |\nabla u_{j_m}|^2)\,dx \geq 4\mu.$$

Setting $k\to\infty$ and then $m\to\infty$, we have $2\mu + 2\int_\Omega G\cdot G\,dx \geq 4\mu$; hence $\int_\Omega |G|^2\,dx \geq \mu$ and so $\|G\|_{L^2(\Omega)} = \mu = \lim_{k\to\infty}\|\nabla u_{j_k}\|_{L^2(\Omega)}$. This proves the strong convergence. The main step is use of the **convexity** of functional $I(u)$ and the given class.)

However, the main question is whether G renders a function in the given class; that is, does there exist a function u in the given class such that $G = \nabla u$? Any such function u would be a minimizer of $I(u)$ in the given class. Since, in principle, G is only in $L^2(\Omega)$, it is not clear whether such a u should exist or not. For this problem, the smoothness of u_j or even the fact that ∇u_j converges strongly in $L^2(\Omega)$ to G would not help much. The class in which we seek the minimizers (i.e., the **admissible class**) plays an important role in guaranteeing the existence of a minimizer.

We need to have a larger admissible class where $I(u)$ is defined and a minimizer u can be found through G as explained above; namely, $\nabla u = G$. As G is only in $L^2(\Omega)$, this leads us to the class of functions whose gradients (in certain sense) are in $L^2(\Omega)$. This motivates the study of **Sobolev spaces** such as $H^1(\Omega) = W^{1,2}(\Omega)$ or general $W^{m,p}(\Omega)$ spaces. Minimizers in such a generalized function space are only **weak solutions** to the Dirichlet problem for Laplace equation. Is it smooth and a classical solution of the Laplace equation? This is the **regularity** problem, which will also be covered in this course.

Example 2 – Lax–Milgram Method. The second example is on the Hilbert space method (**energy method**) for second-order linear elliptic equations in divergence form:

$$Lu = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega,$$

where f is a given function in $L^2(\Omega)$ and Lu is a *second-order linear elliptic operator*:

$$Lu = -\sum_{i,j=1}^n (a_{ij}(x)u_{x_i})_{x_j} + \sum_{i=1}^n b_i(x)u_{x_i} + c(x)u,$$

with **ellipticity condition**: for a constant $\theta > 0$

$$\sum_{i,j=1}^n a_{ij}(x)\xi_i\xi_j \geq \theta|\xi|^2, \quad \forall x\in\Omega,\ \xi\in\mathbb{R}^n.$$

A weak solution u is defined to be a function $u \in H = H_0^1(\Omega)$ for which

$$B(u,v) = (f,v)_{L^2(\Omega)} \quad \forall\ v\in H,$$

where B is the **bilinear form** associated with L:

$$B(u,v) = \int_\Omega \left(\sum_{i,j=1}^n a_{ij}u_{x_i}v_{x_j} + \sum_{i=1}^n b_i u_{x_i} v + cuv\right)dx, \quad u,v\in H.$$

Note that f can also be assumed in the **dual space** $H^* = H^{-1}(\Omega)$; in this case, $(f,v)_{L^2(\Omega)}$ above is replaced by the pairing $\langle f,v\rangle$ between H^* and H. Note also that $B(u,v)$ satisfies the **Gårding's estimates** (energy estimates):

$$|B(u,v)| \le \alpha \|u\| \|v\|, \quad B(u,u) \ge \beta \|u\|^2 - \gamma \|u\|^2_{L^2(\Omega)},$$

where $\alpha > 0, \beta > 0$ and $\gamma \ge 0$ are constants (assuming a^{ij}, b^i, c are $L^\infty(\Omega)$ functions). The **Lax-Milgram theorem** says that if $\gamma = 0$ in the estimate for $B(u,u)$ above then for each $f \in H^{-1}(\Omega)$ there exists a unique $u \in H$ such that $B(u,v) = \langle f, v\rangle$ for all $v \in H$. If B is *symmetric* then the Lax-Milgram theorem is simply the **Riesz representation theorem**.

Again, once we have the existence of weak solution in $H^1(\Omega)$, we would like to know whether it is more regular. This is the regularity problem to be studied for some special cases.

Example 3 – Mountain Pass Method. Our next example is use of **critical point theory** to find a *nontrivial* solution to the semilinear elliptic equations of the following type:

$$\Delta u + |u|^{p-1} u = 0 \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega,$$

where $1 < p < \frac{n+2}{n-2}$ (we will see why this special exponent is needed here later). If we define a functional on $H = H^1_0(\Omega)$ by

$$I(u) = \int_\Omega \left(\frac{1}{2} |\nabla u|^2 - \frac{1}{p+1} |u|^{p+1} \right) dx,$$

then any *critical point* of I on H will be a weak solution of the above problem. It can be seen that I does not have finite infimum or supremum on H.

Although there is another method of solving this problem based on minimizing functional I on a manifold of H (*i.e.*, via minimization with constraints, to be discussed later in the class), we use the **mountain pass method** to study the *saddle-point* critical points of I, which is a major contribution in **nonlinear functional analysis** in 1970s.

Functional I can be proved to be C^1 on H with derivative $I' : H \to H$ being a locally Lipschitz function. A critical property of this functional is that it satisfies the so-called **Palais-Smale condition**: *every sequence $\{u_j\}$ in H with $\{I(u_j)\}$ bounded and $I'(u_j) \to 0$ in H is* **precompact** *in H.* Furthermore, there exist positive constants r, a such that

$$I(u) \ge a \quad \text{if } \|u\| = r,$$

and there exists an element $v \in H$ such that $\|v\| > r$ and $I(v) \le 0$. Note that $I(0) = 0$. Therefore, both inside and outside the mountain range $\|u\| = r$ there are points where I takes a smaller value than it takes on the mountain range. Let Γ be the set of all continuous passes connecting the two lower points 0 and v. Let

$$c = \inf_{g \in \Gamma} \max_{u \in g} I(u).$$

Then the **mountain pass theorem** says that c is a critical value of I; that is, there exists a critical point $u \in H$ at level c; namely, $I(u) = c$, $I'(u) = 0$. Note that $c \ge a$ and hence $u \ne 0$.

Example 4 – Weak Lower Semicontinuity. Direct method also works for the minimization problems of general integral functionals of the type

$$I(u) = \int_\Omega F(x, u(x), Du(x))\, dx,$$

where u may be even a vector valued function. An important question in this regard is whether and when the functional I is **lower semi-continuous** with respect to weak convergence of $W^{1,p}(\Omega)$; that is,

$$I(u) \le \liminf_{j\to\infty} I(u_j) \quad \text{whenever } u_j \rightharpoonup u \text{ in } W^{1,p}(\Omega).$$

This motivates the study of various notion of **convexity conditions**. In the scalar case, the **weak lower semicontinuity** of I is equivalent to the *convexity* of function $F(x,u,\xi)$ on the variable $\xi \in \mathbb{R}^n$. In the vectorial case (where, say, $u\colon \Omega \to \mathbb{R}^N$ for some $N \ge 2$), the weak lower semicontinuity problem is a difficult problem and involves Morrey's notion of **quasiconvexity**, which will also be discussed in the course. But, a sufficient condition for the semicontinuity will be given in terms of **null-Lagrangians**; this will be the **polyconvexity**. The vectorial case is closely related to the problems in **nonlinear elasticity**, **harmonic maps**, **liquid crystals**, and other physical problems.

B. Application to Some Physical Problems

Some problems in *nonlinear elasticity, liquid crystals* and *ferromagnetics* will be discussed as they can be formulated and solved by variational methods.

Problem 1 – Nonlinear Elasticity. In continuum mechanics, a material occupying a domain $\Omega \subset \mathbb{R}^3$ is deformed by a map u to another domain in the same space. The material at position $x \in \Omega$ is deformed to a point $u(x)$ in the deformed domain $u(\Omega)$. The *nonlinear elasticity* theory postulates that the *total stored energy* associated with the deformation u is given by

$$I(u) = \int_\Omega F(x, u(x), Du(x))\, dx,$$

where $Du(x) = (\partial u^i/\partial x_j)$ is the 3×3-matrix of *deformation gradient* and F is the *stored energy density function.* here, $\operatorname{adj} A$ and $\det A$ denote the cofactor matrix and the determinant of matrix A. For elasticity problems, the constraint $\det Du(x) > 0$ for a.e. $x \in \Omega$ is always assumed; this renders an additional difficulty for the variational problems. Also, material property and frame-indifference often prevent the function $F(x,u,A)$ from being convex in A. Nevertheless, the density function sometime can be written as (or is often assumed to be)

$$F(x,u,A) = W(x,u,A,\operatorname{adj} A, \det A),$$

where $\operatorname{adj} A$ and $\det A$ denote the cofactor matrix and the determinant of matrix A, and $W(x,u,A,B,t)$ is a convex function of (A,B,t). Exactly, this means F is *polyconvex.* For incompressible materials, the constraint $\det Du(x) = 1$ is assumed. The relationship between weak convergence and determinant involves the **compensated compactness**, and we will discuss this using the *null-Lagrangians* under a higher regularity assumption.

Problem 2 – Liquid Crystals. A liquid crystal is described by the orientation of the line-like (nematic) molecules. Such an orientation can be modeled by a unit vector $\mathbf{n}(x) \in S^2$ at each material point $x \in \Omega$, the domain occupied by the liquid crystal. The total energy, based on the **Oseen-Frank model**, is given by

$$\begin{aligned} I(\mathbf{n}) = \int_\Omega W_{OF}(\mathbf{n}, D\mathbf{n})\, dx \; &= \; \frac{1}{2}\int_\Omega (\kappa_1 (\operatorname{div} \mathbf{n})^2 + \kappa_2 (\mathbf{n}\cdot \operatorname{curl} \mathbf{n})^2 + \kappa_3 (\mathbf{n} \times \operatorname{curl} \mathbf{n})^2)\, dx \\ &+ \; \frac{1}{2}\int_\Omega \kappa_4 (\operatorname{tr}((D\mathbf{n})^2) - (\operatorname{div} \mathbf{n})^2)\, dx, \end{aligned}$$

where κ_1-term is the Frank **splay** energy, κ_2-term **twist** energy, κ_3-term **bend** energy; the κ_4-term is a **null-Lagrangian**, depending only on the boundary data of $\mathbf{n}$. If first three $\kappa_i > 0$, then it can be shown that the first part of $W_{OF}(\mathbf{n}, A)$ is convex in A. If all k_i are equal to a positive constant κ, then $I(\mathbf{n})$ reduces to the Dirichlet integral for **harmonic maps**: $I(\mathbf{n}) = \frac{1}{2}\int_\Omega \kappa |D\mathbf{n}|^2\,dx$. Note that the constraint $|\mathbf{n}(x)| = 1$ is lower-order in $H^1(\Omega; S^2)$ and thus presents no problem when using the direct method, but this nonconvex condition is the main obstacle for uniqueness and regularity.

Problem 3 – Micromagnetics. In the Landau-Lifshitz theory of micromagnetics, one seeks the magnetization $\mathbf{m} : \Omega \subset \mathbb{R}^N \to \mathbb{R}^N$ of a body occupying the region Ω that minimizes the total energy

$$I(\mathbf{m}) = \frac{\alpha}{2}\int_\Omega |\nabla \mathbf{m}(x)|^2\,dx + \int_\Omega \varphi(\mathbf{m}(x))\,dx - \int_\Omega H\cdot \mathbf{m}(x)\,dx + \frac{1}{2}\int_{\mathbb{R}^N} |F(z)|^2\,dz$$

among all admissible magnetizations $\mathbf{m}$ satisfying

$$\mathbf{m} \in L^\infty(\Omega), \quad |\mathbf{m}(x)| = 1 \ \text{ a.e. } x \in \Omega,$$

where $F \in L^2(\mathbb{R}^N; \mathbb{R}^N)$ is the unique field determined by the simplified Maxwell's equations:

$$\operatorname{curl} F = 0, \quad \operatorname{div}(-F + \mathbf{m}\chi_\Omega) = 0 \ \text{ in } \mathbb{R}^N.$$

Here, $\alpha > 0$ is a material constant which represents the so-called **exchange energy**, φ is density of the so-called **anisotropy energy** and is minimized along certain preferred crystallographic directions (*easy axises*), H is a given applied field accounting for the so-called **external interaction energy**, and finally F is the magnetic field induced by $\mathbf{m}$ on whole $\mathbb{R}^N$ via Maxwell's equations above, representing the so-called **magnetostatic energy**.

If $\alpha > 0$, except the last non-local energy term, the energy $I(\mathbf{m})$ is much similar to the harmonic map problem. In this case, the natural space for $\mathbf{m}$ is $H^1(\Omega; S^{N-1})$ and hence existence of minimizer is relatively easy because $I(\mathbf{m})$ is weakly lower semi-continuous on $H^1(\Omega; S^{N-1})$.

For large domains, the exchange energy is usually dropped from the total energy $I(\mathbf{m})$. In this case, with $\alpha = 0$, none of the terms in $I(\mathbf{m})$ is more dominating than others, and the nonconvex constraint $|\mathbf{m}(x)| = 1$ also becomes more troublesome; the energy $I(\mathbf{m})$ may not have a minimizer at all. We study a case that a minimizer does not exist, but minimizing sequences can have special structures. We avoid using the notion of **Young measures**, but refer to Young's monograph [**27**] and Tartar [**24**] for an introduction to this important and useful notion.

C. Plan of Lectures

This lecture note will be made available to all students in the class, which contains more detailed materials and some useful references. However, the lectures will only emphasize on some selective topics with more details and additional references; other materials may not be covered in lecture, but they are important part of the course. Students can truly learn the materials by reading the whole lecture note and working on some examples.

The following is a detailed list of materials contained in the lecture notes. The core materials are Chapters 3–5.

1. Preliminaries.

- Banach spaces
- Bounded linear operators
- Weak convergence and compact operators
- Spectral theory for linear compact operators
- Some useful results in nonlinear functional analysis

2. Sobolev Spaces.

- Weak derivatives and Sobolev spaces
- Approximations and extensions
- Sobolev imbedding theorems
- Additional properties

3. Second-Order Linear Elliptic PDEs in Divergence Form.

- Second-order PDEs and systems in divergence form
- Lax-Milgram theorem
- Gårding's inequality and existence theorem
- Regularity of weak solutions
- Symmetric elliptic operators and eigenvalue problems

4. Variational Methods for Nonlinear PDEs.

- Variational problems in PDEs
- Multiple integrals in the calculus of variations
- Direct method for minimization
- Minimization with constraints
- Mountain pass theorem
- Nonexistence and radial symmetry

5. Weak Lower Semicontinuity on Sobolev Spaces.

- The convex case
- Morrey's quasiconvexity
- Properties of quasiconvex functions
- Polyconvex functions and null-Lagrangians
- Existence in nonlinear elasticity
- Relaxation and existence for nonconvex problems

Chapter 1

Preliminaries

1.1. Banach Spaces

A (real) **vector space** is a set X, whose elements are called **vectors**, and in which two operations, **addition** and **scalar multiplication**, are defined as follows:

(a) To every pair of vectors x and y corresponds a vector $x + y$ in such a way that

$$x + y = y + x \qquad \text{and} \qquad x + (y + z) = (x + y) + z.$$

X contains a unique vector 0 (the **zero vector** or **origin** of X) such that $x+0 = x$ for every $x \in X$, and to each $x \in X$ corresponds a unique vector $-x$ such that $x + (-x) = 0$.

(b) To every pair (α, x), with $\alpha \in \mathbb{R}$ and $x \in X$, corresponds a vector αx in such a way that

$$1x = x, \qquad \alpha(\beta x) = (\alpha\beta)x$$

and such that the two distributive laws

$$\alpha(x + y) = \alpha x + \alpha y, \qquad (\alpha + \beta)x = \alpha x + \beta x$$

hold.

A nonempty subset M of a vector space X is called a **subspace** of X if $\alpha x + \beta y \in M$ for all $x, y \in M$ and all $\alpha, \beta \in \mathbb{R}$. A subset M of a vector space X is said to be **convex** if $tx + (1 - t)y \in M$ whenever $t \in (0, 1)$, $x, y \in M$. (Clearly, every subspace of X is convex.)

Let $x_1, \ldots, x_n$ be elements of a vector space X. The set of all $\alpha_1 x_1 + \cdots + \alpha_n x_n$, with $\alpha_i \in \mathbb{R}$, is called the **span** of $x_1, \ldots, x_n$ and is denoted by $\operatorname{span}\{x_1, \ldots, x_n\}$. The elements $x_1, \ldots, x_n$ are said to be **linearly independent** if $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$ implies that $\alpha_i = 0$ for each i; otherwise, they are said to be **linearly dependent**. An arbitrary collection of vectors is said to be linearly independent if every finite subset of distinct elements is linearly independent.

The **dimension** of a vector space X, denoted by $\dim X$, is either 0, a positive integer or ∞. If $X = \{0\}$ then $\dim X = 0$; if there exist linearly independent $\{u_1, \ldots, u_n\}$ such that each $x \in X$ has a (unique) representation of the form

$$x = \alpha_1 u_1 + \cdots + \alpha_n u_n \qquad \text{with} \qquad \alpha_i \in \mathbb{R}$$

then $\dim X = n$ and $\{u_1, \dots, u_n\}$ is a **basis** for X; in all other cases $\dim X = \infty$.

A (real) vector space X is said to be a **normed space** if to every $x \in X$ there is associated a nonnegative real number $\|x\|$, called the **norm** of x, in such a way that

(a) $\|x + y\| \le \|x\| + \|y\|$ for all x and y in X (**Triangle inequality**)

(b) $\|\alpha x\| = |\alpha|\|x\|$ for all $x \in X$ and all $\alpha \in \mathbb{R}$

(c) $\|x\| > 0$ if $x \ne 0$.

Note that (b) and (c) imply that $\|x\| = 0$ iff $x = 0$. Moreover, it easily follows from (a) that

$$|\|x\| - \|y\|| \le \|x - y\| \quad \text{for all} \quad x, y \in X.$$

A sequence $\{x_n\}$ in a normed space X is called a **Cauchy sequence** if, for each $\epsilon > 0$, there exists an integer N such that $\|x_m - x_n\| < \epsilon$ for all $m, n \ge N$. We say $x_n \to x$ in X if $\lim_{n\to\infty} \|x_n - x\| = 0$ and, in this case, x is called the limit of $\{x_n\}$. X is called **complete** if every Cauchy sequence in X converges to a limit in X.

A complete (real) normed space is called a (real) **Banach space**. A Banach space is **separable** if it contains a countable dense set. It can be shown that a subspace of a separable Banach space is itself separable.

Example 1.1. Let Ω be an open subset of $\mathbb{R}^n$, $n \ge 1$. The set $C(\Omega)$ of (real-valued) continuous functions defined on Ω is an infinite dimensional vector space with the usual definitions of addition and scalar multiplication:

$$(f + g)(x) = f(x) + g(x) \quad \text{for} \quad f, g \in C(\Omega),\ x \in \Omega$$

$$(\alpha f)(x) = \alpha f(x) \quad \text{for} \quad \alpha \in \mathbb{R},\ f \in C(\Omega),\ x \in \Omega.$$

$C(\bar\Omega)$ consists of those functions which are uniformly continuous on Ω. Each such function has a continuous extension to $\bar\Omega$. $C_0(\Omega)$ consists of those functions which are continuous in Ω and have compact support in Ω. (The **support** of a function f defined on Ω is the closure of the set $\{x \in \Omega : f(x) \ne 0\}$ and is denoted by $\operatorname{supp}(f)$.) The latter two spaces are clearly subspaces of $C(\Omega)$.

For each n-tuple $\alpha = (\alpha_1, \dots, \alpha_n)$ of nonnegative integers, we denote by D^α the partial derivative

$$D_1^{\alpha_1} \cdots D_n^{\alpha_n}, \quad D_i = \partial/\partial x_i$$

of order $|\alpha| = \alpha_1 + \cdots + \alpha_n$. If $|\alpha| = 0$, then $D^0 = I$(identity).

For integers $m \ge 0$, let $C^m(\Omega)$ be the collection of all $f \in C(\Omega)$ such that $D^\alpha f \in C(\Omega)$ for all α with $|\alpha| \le m$. We write $f \in C^\infty(\Omega)$ iff $f \in C^m(\Omega)$ for all $m \ge 0$. For $m \ge 0$, define $C_0^m(\Omega) = C_0(\Omega) \cap C^m(\Omega)$ and let $C_0^\infty(\Omega) = C_0(\Omega) \cap C^\infty(\Omega)$. The spaces $C^m(\Omega), C^\infty(\Omega), C_0^m(\Omega), C_0^\infty(\Omega)$ are all subspaces of the vector space $C(\Omega)$. Similar definitions can be given for $C^m(\bar\Omega)$ etc.

For $m \ge 0$, define X to be the set of all $f \in C^m(\Omega)$ for which

$$\|f\|_{m,\infty} \equiv \sum_{|\alpha| \le m} \sup_\Omega |D^\alpha f(x)| < \infty.$$

Then X is a Banach space with norm $\|\cdot\|_{m,\infty}$. To prove, for example, the completeness when $m = 0$, we let $\{f_n\}$ be a Cauchy sequence in X, i.e., assume for any $\varepsilon > 0$ there is a number $N(\varepsilon)$ such that for all $x \in \Omega$

$$\sup_{x\in\Omega} |f_n(x) - f_m(x)| < \varepsilon \quad \text{if} \quad m, n > N(\varepsilon).$$

But this means that $\{f_n(x)\}$ is a uniformly Cauchy sequence of bounded continuous functions, and thus converges uniformly to a bounded continuous function $f(x)$. Letting $m \to \infty$ in the above inequality shows that $\|f_n - f\|_{m,\infty} \to 0$.

Note that the same proof is valid for the set of bounded continuous scalar-valued functions defined on a nonempty subset of a normed space X.

EXAMPLE 1.2. Let Ω be a nonempty Lebesgue measurable set in $\mathbb{R}^n$. For $p \in [1, \infty)$, we denote by $L^p(\Omega)$ the set of equivalence classes of Lebesgue measurable functions on Ω for which

$$\|f\|_p \equiv \left(\int_\Omega |f(x)|^p dx\right)^{\frac{1}{p}} < \infty.$$

(Two functions belong to the same equivalence class, i.e., are equivalent, if they differ only on a set of measure 0.) Let $L^\infty(\Omega)$ denote the set of equivalence classes of Lebesgue measurable functions on Ω for which

$$\|f\|_\infty \equiv \text{ess-sup}_{x\in\Omega}|f(x)| < \infty.$$

Then $L^p(\Omega), 1 \leq p \leq \infty$, are Banach spaces with norms $\|\cdot\|_p$. For $p \in [1, \infty]$ we write $f \in L^p_{loc}(\Omega)$ iff $f \in L^p(K)$ for each compact set $K \subset \Omega$.

For the sake of convenience, we will also consider $L^p(\Omega)$ as a set of functions. With this convention in mind, we can assert that $C_0(\Omega) \subset L^p(\Omega)$. In fact, if $p \in [1, \infty)$, then as we shall show later, $C_0(\Omega)$ is dense in $L^p(\Omega)$. The space $L^p(\Omega)$ is also separable if $p \in [1, \infty)$. This follows easily, when Ω is compact, from the last remark and the Weierstrass approximation theorem.

EXAMPLE 1.3. (**Hölder's inequality**) We recall that if $p, q, r \in [1, \infty]$ with $p^{-1}+q^{-1} = r^{-1}$, then **Hölder's inequality** implies that if $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$, then $fg \in L^r(\Omega)$ and

$$\|fg\|_r \leq \|f\|_p\|g\|_q.$$

By induction, we also have the following **generalized Hölder's inequality**:

$$\|f_1f_2 \cdots f_k\|_r \leq \|f_1\|_{p_1}\|f_2\|_{p_2} \cdots \|f_k\|_{p_k} \tag{1.1}$$

for $f_i \in L^{p_i}(\Omega)$, $i = 1, 2, \cdots, k$, with $p_i, r \in [1, \infty]$ satisfying $p_1^{-1} + p_2^{-1} + \cdots + p_k^{-1} = r^{-1}$.

EXAMPLE 1.4. The **Cartesian product** $X \times Y$, of two vector spaces X and Y, is itself a vector space under the following operations of addition and scalar multiplication:

$$[x_1, y_1] + [x_2, y_2] = [x_1 + x_2, y_1 + y_2]$$
$$\alpha[x, y] = [\alpha x, \alpha y].$$

If in addition, X and Y are normed spaces with norms $\|\cdot\|_X$, $\|\cdot\|_Y$ respectively, then $X \times Y$ becomes a normed space under the norm

$$\|[x, y]\| = \|x\|_X + \|y\|_Y.$$

Moreover, under this norm, $X \times Y$ becomes a Banach space provided X and Y are Banach spaces.

Let H be a real vector space. H is said to be an **inner product space** if to every pair of vectors x and y in H there corresponds a real-valued function (x, y), called the **inner product** of x and y, such that

(a) $(x, y) = (y, x)$ for all $x, y \in H$

(b) $(x + y, z) = (x, z) + (y, z)$ for all $x, y, z \in H$

(c) $(\lambda x, y) = \lambda(x, y)$ for all $x, y \in H$, $\lambda \in \mathbb{R}$

(d) $(x, x) \geq 0$ for all $x \in H$, and $(x, x) = 0$ if and only if $x = 0$.

For $x \in H$ we set

$$\|x\| = (x, x)^{1/2}. \tag{1.2}$$

Theorem 1.5. *If H is an inner product space, then for all x and y in H, it follows that*

(a) $|(x, y)| \leq \|x\| \, \|y\|$ (**Cauchy-Schwarz inequality**);

(b) $\|x + y\| \leq \|x\| + \|y\|$ (**Triangle inequality**);

(c) $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$ (**Parallelogram law**).

Proof. (a) is obvious if $x = 0$, and otherwise it follows by taking $\delta = -(x, y)/\|x\|^2$ in

$$0 \leq \|\delta x + y\|^2 = |\delta|^2\|x\|^2 + 2\delta(x, y) + \|y\|^2.$$

This identity, with $\delta = 1$, and (a) imply (b). (c) follows easily by using (1.2). □

Furthermore, by (d), equation (1.2) defines a norm on an inner product space H. If H is complete under this norm, then H is said to be a **Hilbert space**.

Example 1.6. The space $L^2(\Omega)$ is a Hilbert space with inner product

$$(f, g) = \int_\Omega f(x)g(x)\,dx \quad \text{for all} \quad f, g \in L^2(\Omega).$$

Theorem 1.7. *Every nonempty closed convex subset S of a Hilbert space H contains a unique element of minimal norm.*

Proof. Choose $x_n \in S$ so that $\|x_n\| \to d \equiv \inf\{\|x\| : x \in S\}$. Since $(1/2)(x_n + x_m) \in S$, we have $\|x_n + x_m\|^2 \geq 4d^2$. Using the parallelogram law, we see that

$$\|x_n - x_m\|^2 \leq 2(\|x_n\|^2 - d^2) + 2(\|x_m\|^2 - d^2) \tag{1.3}$$

and therefore $\{x_n\}$ is a Cauchy sequence in H. Since S is closed, $\{x_n\}$ converges to some $x \in S$ and $\|x\| = d$. If $y \in S$ and $\|y\| = d$, then the parallelogram law implies, as in (1.3), that $x = y$. □

If $(x, y) = 0$, then x and y are said to be **orthogonal**, written sometimes as $x \perp y$. For $M \subset H$, the **orthogonal complement** of M, denoted by $M^\perp$, is defined to be the set of all $x \in H$ such that $(x, y) = 0$ for all $y \in M$. It is easily seen that $M^\perp$ is a closed subspace of H. Moreover, if M is a dense subset of H and if $x \in M^\perp$, then in fact, $x \in H^\perp$ which implies $x = 0$.

Theorem 1.8. (Projection) *Suppose M is a closed subspace of a Hilbert space H. Then for each $x \in H$ there exist unique $y \in M$, $z \in M^\perp$ such that $x = y + z$. The element y is called the* **projection** *of x onto M.*

Proof. Let $S = \{x - y : y \in M\}$. It is easy to see that S is convex and closed. Theorem 1.7 implies that there exists a $y \in M$ such that $\|x - y\| \leq \|x - w\|$ for all $w \in M$. Let $z = x - y$. For an arbitrary $w \in M$, $w \neq 0$, let $\alpha = (z, w)/\|w\|^2$ and note that

$$\|z\|^2 \leq \|z - \alpha w\|^2 = \|z\|^2 - |(z, w)/\|w\||^2$$

which implies $(z, w) = 0$. Therefore $z \in M^\perp$. If $x = y' + z'$ for some $y' \in M$, $z' \in M^\perp$, then $y' - y = z - z' \in M \cap M^\perp = \{0\}$, which implies uniqueness. □

Remark. In particular, if M is a proper closed subspace of H, then there is a nonzero element in $M^{\perp}$. Indeed, for $x \in H \backslash M$, let y be the projection of x on M. Then $z = x - y$ is a nonzero element of $M^{\perp}$.

1.2. Bounded Linear Operators

Let X, Y be real vector spaces. A map $T\colon X \to Y$ is said to be a **linear operator** from X to Y if

$$T(\alpha x + \beta y) = \alpha Tx + \beta Ty$$

for all $x, y \in \mathcal{D}(T)$ and all $\alpha, \beta \in \mathbb{R}$.
Let X, Y be normed spaces. A linear operator T from X to Y is said to be **bounded** if there exists a constant $m > 0$ such that

$$\|Tx\| \leq m\|x\| \quad \text{for all} \quad x \in X. \tag{1.4}$$

We define the **operator norm** $\|T\|$ of T by

$$\|T\| = \sup_{x \in X,\, \|x\|=1} \|Tx\| = \sup_{x \in X,\, \|x\| \leq 1} \|Tx\|. \tag{1.5}$$

The collection of all bounded linear operators $T : X \to Y$ will be denoted by $\mathcal{B}(X, Y)$. We shall also set $\mathcal{B}(X) = \mathcal{B}(X, X)$ when $X = Y$. Observe that

$$\|TS\| \leq \|T\|\|S\| \quad \text{if} \quad S \in \mathcal{B}(X, Y),\ T \in \mathcal{B}(Y, Z).$$

Theorem 1.9. *If X and Y are normed spaces, then $\mathcal{B}(X, Y)$ is a normed space with norm defined by equation* (1.5). *If Y is a Banach space, then $\mathcal{B}(X, Y)$ is also a Banach space.*

Proof. It is easy to see that $\mathcal{B}(X, Y)$ is a normed space. To prove completeness, assume that $\{T_n\}$ is a Cauchy sequence in $\mathcal{B}(X, Y)$. Since

$$\|T_n x - T_m x\| \leq \|T_n - T_m\|\|x\| \tag{1.6}$$

we see that, for fixed $x \in X$, $\{T_n x\}$ is a Cauchy sequence in Y and therefore we can define a linear operator T by

$$Tx = \lim_{n \to \infty} T_n x \quad \text{for all} \quad x \in X.$$

If $\varepsilon > 0$, then the right side of (1.6) is smaller than $\varepsilon\|x\|$ provided that m and n are large enough. Thus, (letting $n \to \infty$)

$$\|Tx - T_m x\| \leq \varepsilon\|x\| \quad \text{for all large enough } m.$$

Hence, $\|Tx\| \leq (\|T_m\| + \varepsilon)\|x\|$, which shows that $T \in \mathcal{B}(X, Y)$. Moreover, $\|T - T_m\| < \varepsilon$ for all large enough m. Hence, $\lim_{n\to\infty} T_n = T$. □

The following theorems are important and can be found in any standard texbook of real analysis or functional analysis; e.g., [**3, 26**].

Theorem 1.10. (Banach-Steinhaus) *Let X be a Banach space and Y a normed space. If $A \subset \mathcal{B}(X, Y)$ is such that $\sup_{T \in A} \|Tx\| < \infty$ for each fixed $x \in X$, then $\sup_{T \in A} \|T\| < \infty$.*

Theorem 1.11. (Bounded Inverse) *If X and Y are Banach spaces and if $T \in \mathcal{B}(X, Y)$ is one-to-one and onto, then $T^{-1} \in \mathcal{B}(Y, X)$.*

When X is a (real) normed space, the Banach space $\mathcal{B}(X, \mathbb{R})$ will be called the (normed) **dual space** of X and will be denoted by X^*. Elements of X^* are called **bounded linear functionals** or **continuous linear functionals** on X. Frequently, we shall use the notation $\langle f, x\rangle$ to denote the value of $f \in X^*$ at $x \in X$. Using this notation we note that $|\langle f, x\rangle| \leq \|f\| \, \|x\|$ for all $f \in X^*$, $x \in X$.

Example 1.12. Suppose $1 < p, q < \infty$ satisfy $1/p + 1/q = 1$ and let Ω be a nonempty Lebesgue measurable set in $\mathbb{R}^n$. Then $L^p(\Omega)^* = L^q(\Omega)$. The case of $p = \infty$ is different. The dual of L^∞ is much larger then L^1.

The following results are also standard.

Theorem 1.13. (Hahn-Banach) *Let X be a normed space and Y a subspace of X. Assume $f \in Y^*$. Then there exists a bounded linear functional $\tilde{f} \in X^*$ such that*

$$\langle \tilde{f}, y\rangle = \langle f, y\rangle \quad \forall\, y \in Y, \quad \|\tilde{f}\|_{X^*} = \|f\|_{Y^*}.$$

Corollary 1.14. *Let X be a normed space and $x_0 \neq 0$ in X. Then there exists $f \in X^*$ such that*

$$\|f\| = 1, \quad \langle f, x_0\rangle = \|x_0\|.$$

The dual space X^{**} of X^* is called the **second dual space** of X and is again a Banach space. Note that to each $x \in X$ we can associate a unique element $F_x \in X^{**}$ defined by $F_x(f) = \langle f, x\rangle$ for all $f \in X^*$. From Corollary 1.14, one can also show that $\|F_x\| = \|x\|$. Thus, the (canonical) mapping $J : X \to X^{**}$, given by $Jx = F_x$, is a linear isometry of X onto the subspace $J(X)$ of X^{**}. Since J is one-to-one, we can identify X with $J(X)$.

A Banach space X is called **reflexive** if its canonical map J is onto X^{**}. For example, all L^p spaces with $1 < p < \infty$ are reflexive.

We shall need the following properties of reflexive spaces.

Theorem 1.15. *Let X and Y be Banach spaces.*

(a) *X is reflexive iff X^* is reflexive.*
(b) *If X is reflexive, then a closed subspace of X is reflexive.*
(c) *Let $T : X \to Y$ be a linear bijective isometry. If Y is reflexive, then X is reflexive.*

The following theorem characterizes all bounded linear functionals on a Hilbert space.

Theorem 1.16. (Riesz Representation) *If H is a Hilbert space and $f \in H^*$, then there exists a unique $y \in H$ such that*

$$f(x) = \langle f, x\rangle = (x, y) \quad \textit{for all} \quad x \in H.$$

Moreover, $\|f\| = \|y\|$.

Proof. If $f(x) = 0$ for all x, take $y = 0$. Otherwise, there is an element $z \in \mathcal{N}(f)^\perp$ such that $\|z\| = 1$. (Note that the linearity and continuity of f implies that $\mathcal{N}(f)$ is a closed subspace of H.) Put $u = f(x)z - f(z)x$. Since $f(u) = 0$, we have $u \in \mathcal{N}(f)$. Thus $(u, z) = 0$, which implies

$$f(x) = f(x)(z, z) = f(z)(x, z) = (x, f(z)z) = (x, y),$$

where $y = f(z)z$. To prove uniqueness, suppose $(x, y) = (x, y')$ for all $x \in H$. Then in particular, $(y - y', y - y') = 0$, which implies $y = y'$. From the Cauchy-Schwarz inequality we get $|f(x)| \leq \|x\|\|y\|$, which yields $\|f\| \leq \|y\|$. The reverse inequality follows by choosing $x = y$ in the representation. $\square$

Corollary 1.17. *H is reflexive.*

Let $T : H \to H$ be an operator on the Hilbert space H. We define the **Hilbert space adjoint** $T^* : H \to H$ as follows:

$$(Tx, y) = (x, T^*y) \quad \text{for all} \quad x,\, y \in H.$$

The adjoint operator is easily seen to be linear.

Theorem 1.18. *Let H be a Hilbert space. If $T \in \mathcal{B}(H)$, then $T^* \in \mathcal{B}(H)$ and $\|T\| = \|T^*\|$.*

Proof. For any $y \in H$ and all $x \in H$, set $f(x) = (Tx, y)$. Then it is easily seen that $f \in H^*$. Hence by the Riesz representation theorem, there exists a unique $z \in H$ such that $(Tx, y) = (x, z)$ for all $x \in H$, i.e., $\mathcal{D}(T^*) = H$. Moreover, $\|T^*y\| = \|z\| = \|f\| \leq \|T\|\|y\|$, i.e., $T^* \in \mathcal{B}(H)$ and $\|T^*\| \leq \|T\|$. The reverse inequality follows easily from $\|Tx\|^2 = (Tx, Tx) = (x, T^*Tx) \leq \|Tx\|\|T^*\|\|x\|$. $\square$

1.3. Weak Convergence and Compact Operators

Let X be a normed space. A sequence $x_n \in X$ is said to be **weakly convergent** to an element $x \in X$, written $x_n \rightharpoonup x$, if $\langle f, x_n \rangle \to \langle f, x \rangle$ for all $f \in X^*$.

Theorem 1.19. *Let $\{x_n\}$ be a sequence in X.*

(a) *Weak limits are unique.*

(b) *If $x_n \to x$, then $x_n \rightharpoonup x$.*

(c) *If $x_n \rightharpoonup x$, then $\{x_n\}$ is bounded and $\|x\| \leq \liminf \|x_n\|$.*

Proof. To prove (a), suppose that x and y are both weak limits of the sequence $\{x_n\}$ and set $z = x - y$. Then $\langle f, z \rangle = 0$ for every $f \in X^*$ and by Corollary 1.14, $z = 0$. To prove (b), let $f \in X^*$ and note that $x_n \to x$ implies $\langle f, x_n \rangle \to \langle f, x \rangle$ since f is continuous. To prove (c), assume $x_n \rightharpoonup x$ and consider the sequence $\{Jx_n\}$ of elements of X^{**}, where $J : X \to X^{**}$ is the bounded operator defined above. For each $f \in X^*$, $\sup |Jx_n(f)| = \sup |\langle f, x_n \rangle| < \infty$ (since $\langle f, x_n \rangle$ converges). By the Banach-Steinhaus Theorem, there exists a constant c such that $\|x_n\| = \|Jx_n\| \leq c$ which implies $\{x_n\}$ is bounded. Finally, for $f \in X^*$

$$|\langle f, x \rangle| = \lim |\langle f, x_n \rangle| \leq \liminf \|f\|\|x_n\| = \|f\| \liminf \|x_n\|$$

which implies the desired inequality since $\|x\| = \sup_{\|f\|=1} |\langle f, x \rangle|$. $\square$

We note that in a Hilbert space H, the Riesz representation theorem implies that $x_n \rightharpoonup x$ means $(x_n, y) \to (x, y)$ for all $y \in H$. Moreover, we have

$$(x_n, y_n) \to (x, y) \;\text{ if }\; x_n \rightharpoonup x, y_n \to y.$$

This follows from the estimate

$$|(x, y) - (x_n, y_n)| = |(x - x_n, y) - (x_n, y_n - y)| \leq |(x - x_n, y)| + \|x_n\|\|y - y_n\|$$

and the fact that $\|x_n\|$ is bounded.

The main result of this section is given by:

Theorem 1.20. *If X is a reflexive Banach space, then the closed unit ball is* **weakly compact**, *i.e., the sequence $\{x_n\}$ with $\|x_n\| \le 1$ has a subsequence which converges weakly to an x with $\|x\| \le 1$.*

Let X and Y be normed spaces. An operator $T : X \to Y$ is said to be **compact** if it maps bounded sets in X into relatively compact sets in Y, i.e., if for every bounded sequence $\{x_n\}$ in X, $\{Tx_n\}$ has a subsequence which converges to some element of Y.

Since relatively compact sets are bounded, it follows that a compact operator is bounded. On the other hand, since bounded sets in finite-dimensional spaces are relatively compact, it follows that a bounded operator with finite dimensional range is compact. It can be shown that the identity map $I : X \to X$ ($\|Ix\| = \|x\|$) is compact iff X is finite-dimensional. Finally we note that the operator ST is compact if (a) $T : X \to Y$ is compact and $S : Y \to Z$ is continuous or (b) T is bounded and S is compact.

One of the main methods of proving the compactness of certain operators is based upon the Ascoli theorem.

Let Ω be a subset of the normed space X. A set $S \subset C(\Omega)$ is said to be **equicontinuous** if for each $\varepsilon > 0$ there exists a $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ for all $x, y \in \Omega$ with $\|x - y\| < \delta$ and for all $f \in S$.

Theorem 1.21. (Ascoli) *Let Ω be a relatively compact subset of a normed space X and let $S \subset C(\Omega)$. Then S is relatively compact if it is bounded and equicontinuous.*

Remark. In other words, every bounded equicontinuous sequence of functions has a uniformly convergent subsequence.

Theorem 1.22. *Let X and Y be Banach spaces. If $T_n : X \to Y$ are linear and compact for $n \ge 1$ and if $\lim_{n\to\infty} \|T_n - T\| = 0$, then T is compact. Thus, linear compact operators form a closed subspace of $\mathcal{B}(X, Y)$.*

Proof. Let $\{x_n\}$ be a sequence in X with $M = \sup_n \|x_n\| < \infty$. Let A_1 denote an infinite set of integers such the sequence $\{T_1 x_n\}_{n\in A_1}$ converges. For $k \ge 2$ let $A_k \subset A_{k-1}$ denote an infinite set of integers such that the sequence $\{T_k x_n\}_{n\in A_k}$ converges. Choose $n_1 \in A_1$ and $n_k \in A_k$, $n_k > n_{k-1}$ for $k \ge 2$. Choose $\varepsilon > 0$. Let k be such that $\|T - T_k\|M < \varepsilon/4$ and note that

$$\|Tx_{n_i} - Tx_{n_j}\| \le \|(T - T_k)(x_{n_i} - x_{n_j})\| + \|T_k x_{n_i} - T_k x_{n_j}\| < \varepsilon/2 + \|T_k x_{n_i} - T_k x_{n_j}\|.$$

Since $\{T_k x_{n_i}\}_{i=1}^{\infty}$ converges, $\{Tx_{n_i}\}_{i=1}^{\infty}$ is a Cauchy sequence. □

Theorem 1.23. *Let X and Y be normed spaces.*

(a) *If $T \in \mathcal{B}(X, Y)$, then T is* **weakly continuous**, *i.e.,*

$$x_n \rightharpoonup x \quad \textit{implies} \quad Tx_n \rightharpoonup Tx.$$

(b) *If $T : X \to Y$ is weakly continuous and X is a reflexive Banach space, then T is bounded.*

(c) *If $T \in \mathcal{B}(X, Y)$ is compact, then T is* **strongly continuous**, *i.e.,*

$$x_n \rightharpoonup x \quad \textit{implies} \quad Tx_n \to Tx.$$

(d) *If $T : X \to Y$ is strongly continuous and X is a reflexive Banach space, then T is compact.*

Proof. (a) Let $x_n \rightharpoonup x$. Then for every $g \in Y^*$

$$\langle g, Tx_n \rangle = \langle T^* g, x_n \rangle \to \langle T^* g, x \rangle = \langle g, Tx \rangle.$$

(b) If not, there is a bounded sequence $\{x_n\}$ such that $\|Tx_n\| \to \infty$. Since X is reflexive, $\{x_n\}$ has a weakly convergent subsequence, $\{x_{n'}\}$, and so $\{Tx_{n'}\}$ also converges weakly. But then $\{Tx_{n'}\}$ is bounded, which is a contradiction.

(c) Let $x_n \rightharpoonup x$. Since T is compact and $\{x_n\}$ is bounded, there is a subsequence $\{x_{n'}\}$ such that $Tx_{n'} \to z$, and thus $Tx_{n'} \rightharpoonup z$. By (a), $Tx_n \rightharpoonup Tx$, and so $Tx_{n'} \to Tx$. Now it is easily seen that every subsequence of $\{x_n\}$ has a subsequence, say $\{x_{n'}\}$, such that $Tx_{n'} \to Tx$. But this implies the whole sequence $Tx_n \to Tx$ (See the appendix).

(d) Let $\{x_n\}$ be a bounded sequence. Since X is reflexive, there is a subsequence $\{x_{n'}\}$ such that $x_{n'} \rightharpoonup x$. Hence $Tx_{n'} \to Tx$, which implies T is compact. □

Theorem 1.24. *Let H be a Hilbert space. If $T : H \to H$ is linear and compact, then T^* is compact.*

Proof. Let $\{x_n\}$ be a sequence in H satisfying $\|x_n\| \le m$. The sequence $\{T^* x_n\}$ is therefore bounded, since T^* is bounded. Since T is compact, by passing to a subsequence if necessary, we may assume that the sequence $\{TT^* x_n\}$ converges. But then

$$\begin{aligned} \|T^*(x_n - x_m)\|^2 &= (x_n - x_m, TT^*(x_n - x_m)) \\ &\le 2m\|TT^*(x_n - x_m)\| \to 0 \;\text{ as }\; m, n \to \infty. \end{aligned}$$

Since H is complete, the sequence $\{T^* x_n\}$ is convergent and hence T^* is compact. □

1.4. Spectral Theory for Linear Compact Operators

Theorem 1.25. (Fredholm Alternative) *Let $T : H \to H$ be a linear compact operator on the Hilbert space H. Then equations $(I - T)x = 0$, $(I - T^*)x^* = 0$ have the same finite number of linearly independent solutions. Moreover,*

(a) *For $y \in H$, the equation $(I - T)x = y$ has a solution iff $(y, x^*) = 0$ for every solution x^* of $(I - T^*)x^* = 0$.*

(b) *For $z \in H$, the equation $(I - T^*)x^* = z$ has a solution iff $(z, x) = 0$ for every solution x of $(I - T)x = 0$.*

(c) *The inverse operator $(I - T)^{-1} \in \mathcal{B}(H)$ whenever it exists.*

A subset S of a Hilbert space H is said to be an **orthonormal** set if each element of S has norm 1 and if every pair of distinct elements in S is orthogonal. It easily follows that an orthonormal set is linearly independent. An orthonormal set S is said to be **complete** if $x = \sum_{\phi \in S} (x, \phi)\phi$ for all $x \in H$. It can be shown that $(x, \phi) \neq 0$ for at most countably many $\phi \in S$. This series is called the **Fourier series** for x with respect to the orthonormal set $\{\phi\}$. Let $\{\phi_i\}_{i=1}^{\infty}$ be a countable orthonormal set in H. Upon expanding $\|x - \sum_{n=1}^{N} (x, \phi_n)\phi_n\|^2$, we arrive at **Bessel's inequality**:

$$\sum_{n=1}^{\infty} |(x, \phi_n)|^2 \le \|x\|^2.$$

Let $T : \mathcal{D}(T) \subset H \to H$ be a linear operator on the real Hilbert space H. The set $\rho(T)$ of all scalars $\lambda \in \mathbb{R}$ for which $(T - \lambda I)^{-1} \in \mathcal{B}(H)$ is called the **resolvent set** of T. The operator $R(\lambda) = (T - \lambda I)^{-1}$ is known as the **resolvent** of T. $\sigma(T) = \mathbb{R} \setminus \rho(T)$ is called the

spectrum of T. It can be shown that $\rho(T)$ is an open set and $\sigma(T)$ is a closed set. The set of $\lambda \in \mathbb{R}$ for which there exists a nonzero $x \in \mathcal{N}(T-\lambda I)$ is called the **point spectrum** of T and is denoted by $\sigma_p(T)$. The elements of $\sigma_p(T)$ are called the **eigenvalues** of T and the nonzero members of $\mathcal{N}(T-\lambda I)$ are called the **eigenvectors** (or **eigenfunctions** if X is a function space) of T.

If T is compact and $\lambda \neq 0$, then by the Fredholm alternative, either $\lambda \in \sigma_p(T)$ or $\lambda \in \rho(T)$. Moreover, if H is infinite-dimensional, then $0 \notin \rho(T)$; otherwise, $T^{-1} \in \mathcal{B}(H)$ and $T^{-1}T = I$ is compact. As a consequence, $\sigma(T)$ consists of the nonzero eigenvalues of T together with the point 0. The next result shows that $\sigma_p(T)$ is either finite or a countably infinite sequence tending to zero.

Theorem 1.26. *Let $T : X \to X$ be a linear compact operator on the normed space X. Then for each $r > 0$ there exist at most finitely many $\lambda \in \sigma_p(T)$ for which $|\lambda| > r$.*

Let H be a real Hilbert space. An operator $T \in \mathcal{B}(H)$ is said to be **symmetric** if $(Tx, y) = (x, Ty)$ for all $x, y \in H$. The next result implies that a symmetric compact operator on a Hilbert space has at least one eigenvalue. On the other hand, an arbitrary bounded, linear, symmetric operator need not have any eigenvalues. As an example, let $T : L^2(0,1) \to L^2(0,1)$ be defined by $Tu(x) = xu(x)$.

Theorem 1.27. *Suppose $T \in \mathcal{B}(H)$ is symmetric, i.e., $(Tx, y) = (x, Ty)$ for all $x, y \in H$. Then*

$$\|T\| = \sup_{\|x\|=1} |(Tx, x)|.$$

Moreover, if $H \neq \{0\}$, then there exists a real number $\lambda \in \sigma(T)$ such that $|\lambda| = \|T\|$. If $\lambda \in \sigma_p(T)$, then in absolute value λ is the largest eigenvalue of T.

Proof. Clearly $m \equiv \sup_{\|x\|=1} |(Tx, x)| \leq \|T\|$. To show $\|T\| \leq m$, observe that for all $x, y \in H$

$$\begin{aligned} 2(Tx, y) + 2(Ty, x) &= (T(x+y), x+y) - (T(x-y), x-y) \\ &\leq m(\|x+y\|^2 + \|x-y\|^2) \\ &= 2m(\|x\|^2 + \|y\|^2) \end{aligned}$$

where the last step follows from the paralleogram law. Hence, if $Tx \neq 0$ and $y = (\|x\|/\|Tx\|)Tx$, then

$$2\|x\|\|Tx\| = (Tx, y) + (y, Tx) \leq m(\|x\|^2 + \|y\|^2) = 2m\|x\|^2$$

which implies $\|Tx\| \leq m\|x\|$. Since this is also valid when $Tx = 0$, we have $\|T\| \leq m$. To prove the 'moreover' part, choose $x_n \in H$ such that $\|x_n\| = 1$ and $\|T\| = \lim_{n\to\infty} |(Tx_n, x_n)|$. By renaming a subsequence of $\{x_n\}$, we may assume that (Tx_n, x_n) converge to some real number λ with $|\lambda| = \|T\|$. Observe that

$$\begin{aligned} \|(T-\lambda)x_n\|^2 &= \|Tx_n\|^2 - 2\lambda(Tx_n, x_n) + \lambda^2\|x_n\|^2 \\ &\leq 2\lambda^2 - 2\lambda(Tx_n, x_n) \to 0. \end{aligned}$$

We now claim that $\lambda \in \sigma(T)$. Otherwise, we arrive at the contradiction

$$1 = \|x_n\| = \|(T-\lambda)^{-1}(T-\lambda)x_n\| \leq \|(T-\lambda)^{-1}\| \, \|(T-\lambda)x_n\| \to 0.$$

Finally, we note that if $T\phi = \mu\phi$, with $\|\phi\| = 1$, then $|\mu| = |(T\phi, \phi)| \leq \|T\|$ which implies the last assertion of the theorem. □

Finally we have the following result.

Theorem 1.28. *Let H be a separable Hilbert space and suppose $T : H \to H$ is linear, symmetric and compact. Then there exists a countable complete orthonormal set in H consisting of eigenvectors of T.*

1.5. Some Useful Results in Nonlinear Functional Analysis

In this final preliminary section, we list some useful results in **nonlinear functional analysis**. Proofs and other results can be found in the volumes of ZEIDLER's book.

1.5.1. Contraction Mapping Theorem. Let X be a normed space. A map $T : X \to X$ is called a **contraction** if there exists a number $k < 1$ such that

$$\|Tx - Ty\| \le k\|x - y\| \quad \text{for all} \quad x, y \in X. \tag{1.7}$$

Theorem 1.29. (Contraction Mapping) *Let $T : S \subset X \to S$ be a contraction on the closed nonempty subset S of the Banach space X. Then T has a unique* **fixed point***, i.e., there exists a unique solution $x \in S$ of the equation $Tx = x$. Moreover, $x = \lim_{n\to\infty} T^n x_0$ for any choice of $x_0 \in S$.*

Proof. To prove uniqueness, suppose $Tx = x, Ty = y$. Since $k < 1$, we get $x = y$ from

$$\|x - y\| = \|Tx - Ty\| \le k\|x - y\|.$$

To show that T has a fixed point we set up an iteration procedure. For any $x_0 \in S$ set

$$x_{n+1} = Tx_n,\ n = 0, 1, \ldots$$

Note that $x_{n+1} \in S$ and $x_{n+1} = T^{n+1}x_0$. We now claim that $\{x_n\}$ is a Cauchy sequence. Indeed, for any integers n, p

$$\begin{aligned} \|x_{n+p} - x_n\| &= \|T^{n+p}x_0 - T^n x_0\| \le \sum_{j=n}^{n+p-1} \|T^{j+1}x_0 - T^j x_0\| \\ &\le \sum_{j=n}^{n+p-1} k^j \|Tx_0 - x_0\| \le \frac{k^n}{1-k}\|Tx_0 - x_0\|. \end{aligned}$$

Hence as $n \to \infty$, $\|x_{n+p} - x_n\| \to 0$ independently of p, so that $\{x_n\}$ is a Cauchy sequence with limit $x \in S$. Since T is continuous, we have

$$Tx = \lim_{n\to\infty} Tx_n = \lim_{n\to\infty} x_{n+1} = x$$

and thus x is the unique fixed point. Note that the fixed point x is independent of x_0 since x is a fixed point and fixed points are unique. $\square$

The following result is the so-called **method of continuity** or **continuation method**.

Theorem 1.30. *Let $T_0, T_1 \in \mathcal{B}(X, Y)$, where X is a Banach space and Y is a normed space. For each $t \in [0, 1]$ set*

$$T_t = (1 - t)T_0 + tT_1$$

and suppose there exists a constant $c > 0$ such that for all $t \in [0, 1]$ and $x \in X$

$$\|x\|_X \le c\|T_t x\|_Y. \tag{1.8}$$

Then $R(T_1) = Y$ if $R(T_0) = Y$.

Proof. Set $S = \{t \in [0,1] : \mathcal{R}(T_t) = Y\}$. By hypothesis, $0 \in S$. We need to show that $1 \in S$. In this direction we will show that if $\tau > 0$ and $\tau c(\|T_1\| + \|T_0\|) < 1$, then

$$(1.9) \qquad [0,s] \subset S \quad \text{implies} \quad [0, s+\tau] \subset S.$$

(Note that any smaller τ works.) Since τ can be chosen independently of s, (1.9) applied finitely many times gets us from $0 \in S$ to $1 \in S$.

Let $s \in S$. For $t = s + \tau$, $T_t x = f$ is equivalent to the equation

$$(1.10) \qquad T_s x = f + \tau T_0 x - \tau T_1 x.$$

By (1.8), $T_s^{-1} : Y \to X$ exists and $\|T_s^{-1}\| \leq c$. Hence (1.10) is equivalent to

$$(1.11) \qquad x = T_s^{-1}(f + \tau T_0 x - \tau T_1 x) \equiv Ax$$

and for $A : X \to X$ we have for all $x, y \in X$

$$\|Ax - Ay\| \leq \tau c(\|T_1\| + \|T_0\|)\|x - y\|.$$

By the contraction mapping theorem, (1.11) has a solution and this completes the proof. □

1.5.2. Nemytskii Operators. Let Ω be a nonempty measurable set in $\mathbb{R}^n$ and let $f\colon \Omega \times \mathbb{R}^m \to \mathbb{R}$ be a given function, where $m \geq 1$. Assume

(i) for every $\xi \in \mathbb{R}^m$, $f(x,\xi)$ (as a function of x) is measurable on Ω
(ii) for almost all $x \in \Omega$, $f(x,\xi)$ (as a function of ξ) is continuous on $\mathbb{R}^m$
(iii) for all $(x,\xi) \in \Omega \times \mathbb{R}^m$

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

where b is a fixed nonnegative number, $a \in L^q(\Omega)$ is nonnegative and $1 < p, q < \infty$. Then the **Nemytskii operator** N is defined by

$$Nu(x) = f(x, u(x)),\ x \in \Omega \quad \forall\, u\colon \Omega \to \mathbb{R}^m.$$

We have the following result needed later.

Lemma 1.31. *$N : L^p(\Omega;\mathbb{R}^m) \to L^q(\Omega)$ is continuous and bounded with*

$$(1.12) \qquad \|Nu\|_q \leq \textit{const}\ (\|a\|_q + \|u\|_p^{p/q}) \quad \textit{for all} \quad u \in L^p(\Omega;\mathbb{R}^m).$$

Proof. If $u \in L^p(\Omega;\mathbb{R}^m)$, then each component function of $u(x)$ is measurable on Ω and thus, by (i) and (ii), the function $f(x,u(x))$ is also measurable on Ω. From (iii) we get

$$|f(x,u(x))|^q \leq \text{const}(|a(x)|^q + |u(x)|^p).$$

Integrating over Ω yields (1.12), which shows that N is bounded.

To show that N is continuous, let $u_n \to u$ in $L^p(\Omega;\mathbb{R}^m)$. Then there is a subsequence $\{u_{n'}\}$ and a function $v \in L^p(\Omega)$ such that $u_{n'}(x) \to u(x)$ a.e. and $|u_{n'}(x)| \leq v(x)$ a.e. for all n'. Hence

$$\begin{aligned} \|Nu_{n'} - Nu\|_q^q &= \int_\Omega |f(x,u_{n'}(x)) - f(x,u(x))|^q dx \\ &\leq \text{const} \int_\Omega (|f(x,u_{n'}(x))|^q + |f(x,u(x))|^q) dx \\ &\leq \text{const} \int_\Omega (|a(x)|^q + |v(x)|^p + |u(x)|^p) dx. \end{aligned}$$

By (ii), $f(x, u_{n'}(x)) - f(x, u(x)) \to 0$ as $n \to \infty$ for almost all $x \in \Omega$. The dominating convergence theorem implies that $\|Nu_{n'} - Nu\|_q \to 0$. By repeating this procedure for any subsequence of $u_{n'}$, it follows that $\|Nu_n - Nu\|_q \to 0$ which implies that N is continuous. $\square$

1.5.3. Differentiability. Let S be an open subset of the Banach space X. The functional $f : S \subset X \to \mathbb{R}$ is said to be **Gateaux differentiable** (**G-diff**) at a point $u \in S$ if there exists a functional $g \in X^*$ (often denoted by $f'(u)$) such that

$$\frac{d}{dt} f(u + tv)\Big|_{t=0} = \lim_{t \to 0} \frac{f(u + tv) - f(u)}{t} = [f'(u)]v \quad \text{for all} \quad v \in X.$$

The functional $f'(u)$ is called the **Gateaux derivative** of f at the point $u \in S$. If f is G-diff at each point of S, the map $f' : S \subset X \to X^*$ is called the Gateaux derivative of f on S. In addition, if f' is continuous at u (in the operator norm), then we say that f is C^1 at u. Note that in the case of a real-valued function of several real variables, the Gateaux derivative is nothing more than the directional derivative of the function at u in the direction v.

Let X, Y be Banach spaces and let $A : S \subset X \to Y$ be an arbitrary operator. A is said to be **Frechet differentiable** (**F-diff**) at the point $u \in S$ if there exists an operator $B \in \mathcal{B}(X, Y)$ such that

$$\lim_{\|v\| \to 0} \|A(u + v) - Au - Bv\| / \|v\| = 0.$$

The operator B, often denoted by $A'(u)$, is called the **Frechet derivative** of A at u. Note that if A is Frechet differentiable on S, then $A' : S \to \mathcal{B}(X, Y)$. In addition, if A' is continuous at u (in the operator norm), we say that A is C^1 at u.

Remark. If the functional f is F-diff at $u \in S$, then it is also G-diff at u, and the two derivatives are equal. This follows easily from the definition of the Frechet derivative. The converse is not always true as may be easily verified by simple examples from several variable calculus. However, if the Gateaux derivative exists in a neighborhood of u and if $f \in C^1$ at u, then the Frechet derivative exists at u, and the two derivatives are equal.

Example 1.32. (a) Let $f(\xi) \in C(\mathbb{R})$. Then for $k \geq 0$, the corresponding Nemytskii operator $N : C^k(\bar{\Omega}) \to C(\bar{\Omega})$ is bounded and continuous. If in addition $f(\xi) \in C^1(\mathbb{R})$, then $N \in C^1$ and the Frechet derivative $N'(u)$ is given by

$$[N'(u)v](x) = f'(u(x))v(x).$$

Note that for $u, v \in C^k(\bar{\Omega})$, $|N'(u)v|_0 \leq |f'(u)|_0 |v|_k$ and so $N'(u) \in \mathcal{B}(C^k(\bar{\Omega}), C^(\bar{\Omega}))$ with $\|N'(u)\| \leq |f'(u)|_0$. Clearly $N'(u)$ is continuous at each point $u \in C^k(\bar{\Omega})$. Moreover,

$$\begin{aligned} |N(u + v) - Nu - N'(u)v|_0 &= \sup_x \left| \int_0^1 [\frac{d}{dt} f(u(x) + tv(x)) - f'(u(x))v(x)] dt \right| \\ &\leq |v|_0 \sup_x \int_0^1 |f'(u(x) + tv(x)) - f'(u(x))| dt. \end{aligned}$$

The last integral tends to zero since f' is uniformly continuous on compact subsets of $\mathbb{R}$.

More generally, let $f(\xi) \in C^k(\mathbb{R})$. Then the corresponding Nemytskii operator $N : C^k(\bar{\Omega}) \to C^k(\bar{\Omega})$ is bounded and continuous. If in addition $f(\xi) \in C^{k+1}(\mathbb{R})$, then $N \in C^1$ with Frechet derivative given by $[N'(u)v](x) = f'(u(x))v(x)$. Note that $|uv|_k \leq |u|_k |v|_k$ for $u, v \in C^k(\bar{\Omega})$, and since $C^k(\bar{\Omega}) \subset C(\bar{\Omega})$, the Frechet derivative must be of the stated form.

(b) Let $f(\xi) \in C^{k+1}(\mathbb{R})$, where $k > n/2$. Then we claim that the corresponding Nemytskii operator $N : H^k(\Omega) \to H^k(\Omega)$ is of class C^1 with Frechet derivative given by $[N'(u)v](x) = f'(u(x))v(x)$.

First, suppose $u \in C^k(\bar{\Omega})$. Then $N(u) \in C^k(\bar{\Omega})$ by the usual chain rule. If $u \in H^k(\Omega)$, let $u_m \in C^k(\bar{\Omega})$ with $\|u_m - u\|_{k,2} \to 0$. Since the imbedding $H^k(\Omega) \subset C(\bar{\Omega})$ is continuous, $u_m \to u$ uniformly, and thus $f(u_m) \to f(u)$ and $f'(u_m) \to f'(u)$ uniformly and hence in L^2. Furthermore, $D_i f(u_m) = f'(u_m)D_i u_m \to f'(u)D_i u$ in L^1. Consequently, by Theorem 2.10, we have

$$D_i f(u) = f'(u)D_i u.$$

In a similar fashion we find

$$D_{ij} f(u) = f''(u)D_i u D_j u + f'(u)D_{ij} u$$

with corresponding formulas for higher derivatives.

1.5.4. Implicit Function Theorem. The following lemmas are needed in the proof of the implicit function theorem.

Lemma 1.33. *Let S be a closed nonempty subset of the Banach space X and let $\mathcal{M}$ be a metric space. Suppose $A(x,\lambda) : S \times \mathcal{M} \to S$ is continuous and there is a constant $k < 1$ such that, uniformly for all $\lambda \in \mathcal{M}$*

$$\|A(x,\lambda) - A(y,\lambda)\| \le k\|x - y\| \quad \text{for all} \quad x, y \in S.$$

Then for each $\lambda \in \mathcal{M}$, $A(x,\lambda)$ has a unique fixed point $x(\lambda) \in S$ and moreover, $x(\lambda)$ depends continuously on λ.

Proof. The existence and uniqueness of the fixed point $x(\lambda)$ is of course a consequence of the contraction mapping theorem. To prove continuity, suppose $\lambda_n \to \lambda$. Then

$$\begin{aligned} \|x(\lambda_n) - x(\lambda)\| &= \|A(x(\lambda_n),\lambda_n) - A(x(\lambda),\lambda)\| \\ &\le \|A(x(\lambda_n),\lambda_n) - A(x(\lambda),\lambda_n)\| + \|A(x(\lambda),\lambda_n) - A(x(\lambda),\lambda)\| \\ &\le k\|x(\lambda_n) - x(\lambda)\| + \|A(x(\lambda),\lambda_n) - A(x(\lambda),\lambda)\|. \end{aligned}$$

Therefore

$$\|x(\lambda_n) - x(\lambda)\| \le \frac{1}{1-k}\|A(x(\lambda),\lambda_n) - A(x(\lambda),\lambda)\|.$$

By the assumed continuity of A, the right side tends to zero as $n \to \infty$, and therefore $x(\lambda_n) \to x(\lambda)$. $\square$

Lemma 1.34. *Suppose X, Y are Banach spaces. Let $S \subset X$ be convex and assume $A : S \to Y$ is Frechet differentiable at every point of S. Then*

$$\|Au - Av\| \le \|u - v\| \sup_{w \in S} \|A'(w)\|.$$

In other words, A satisfies a Lipschitz condition with constant $q = \sup_{w\in S} \|A'(w)\|$.

Proof. For fixed $u, v \in S$, set $g(t) = A(u + t(v - u))$, where $t \in [0,1]$. Using the definition of Frechet derivative, we have

$$\begin{aligned} g'(t) &= \lim_{h\to 0} \left(\frac{A(u + (t+h)(v-u)) - A(u + t(v-u))}{h} \right) \\ &= \lim_{h\to 0} \left(\frac{hA'(u + t(v-u))(v-u) + \|h(v-u)\|E}{h} \right) \\ &= A'(u + t(v-u))(v-u). \end{aligned}$$

Hence

$$\|g(0) - g(1)\| = \|Au - Av\| \leq \sup_{t\in[0,1]} \|g'(t)\|$$

which implies the desired result. □

Lemma 1.35. *Let X be a Banach space. Suppose $A : \overline{B(u_0, r)} \subset X \to X$ is a contraction, with Lipschitz constant $q < 1$, where*

$$r \geq (1 - q)^{-1}\|Au_0 - u_0\|.$$

Then A has a unique fixed point $u \in \overline{B(u_0, r)}$.

Proof. For $u \in \overline{B(u_0, r)}$

$$\|Au - u_0\| \leq \|Au - Au_0\| + \|Au_0 - u_0\| \leq q\|u - u_0\| + (1 - q)r.$$

Since $\|u - u_0\| \leq r$, A maps the ball $\overline{B(u_0, r)}$ into itself, and the result follows from the contraction mapping theorem. □

We now consider operator equations of the form $A(u, v) = 0$, where A maps a subset of $X \times Y$ into Z. For a given $[u_0, v_0] \in X \times Y$ we denote the Frechet derivative of A (at $[u_0, v_0]$) with respect to the first (second) argument by $A_u(u_0, v_0)$ $(A_v(u_0, v_0))$.

Theorem 1.36. (Implicit Function) *Let X, Y, Z be Banach spaces. For a given $[u_0, v_0] \in X \times Y$ and $a, b > 0$, let $S = \{[u, v] : \|u - u_0\| \leq a,\ \|v - v_0\| \leq b\}$. Suppose $A : S \to Z$ satisfies the following:*

(i) *A is continuous.*

(ii) *$A_v(\cdot, \cdot)$ exists and is continuous in S (in the operator norm)*

(iii) *$A(u_0, v_0) = 0$.*

(iv) *$[A_v(u_0, v_0)]^{-1}$ exists and belongs to $\mathcal{B}(Z, Y)$.*

Then there are neighborhoods U of u_0 and V of v_0 such that the equation $A(u, v) = 0$ has exactly one solution $v \in V$ for every $u \in U$. The solution v depends continuously on u.

Proof. If in S we define

$$B(u, v) = v - [A_v(u_0, v_0)]^{-1}A(u, v)$$

it is clear that the solutions of $A(u, v) = 0$ and $v = B(u, v)$ are identical. The theorem will be proved by applying the contraction mapping theorem to B. Since

$$B_v(u, v) = I - [A_v(u_0, v_0)]^{-1}A_v(u, v)$$

$B_v(\cdot, \cdot)$ is continuous in the operator norm. Now $B_v(u_0, v_0) = 0$, so for some $\delta > 0$ there is a $q < 1$ such that

$$\|B_v(u, v)\| \leq q$$

for $\|u - u_0\| \leq \delta$, $\|v - v_0\| \leq \delta$. By virtue of Lemma 1.34, $B(u, \cdot)$ is a contraction. Since A is continuous, B is also continuous. Therefore, since $B(u_0, v_0) = v_0$, there is an ε with $0 < \varepsilon \leq \delta$ such that

$$\|B(u, v_0) - v_0\| \leq (1 - q)\delta$$

for $\|u - u_0\| \leq \varepsilon$. The existence of a unique fixed point in the closed ball $\overline{B(v_0, \delta)}$ follows from Lemma 1.35 and the continuity from Lemma 1.33. □

Example 1.37. Let $f(\xi) \in C^{1,\alpha}(\mathbb{R})$, $f(0) = f'(0) = 0$, $g(x) \in C^{\alpha}(\bar{\Omega})$ and consider the boundary value problem

$$\Delta u + f(u) = g(x) \ \text{ in } \ \Omega, \quad u|_{\partial\Omega} = 0. \tag{1.13}$$

Set $X = Z = C^{\alpha}(\bar{\Omega})$, $Y = \{u \in C^{2,\alpha}(\bar{\Omega}) : u|_{\partial\Omega} = 0\}$ and

$$A(g, u) = \Delta u + N(u) - g$$

where N is the Nemytskii operator corresponding to f. The operator A maps $X \times Y$ into the space Z. Clearly $A(0, 0) = 0$ (A is C^1 by earlier examples) and

$$A_u(0, 0)v = \Delta v, \quad v \in Y.$$

It is easily checked that all the conditions of the implicit function theorem are met. In particular, condition (iv) is a consequence of the bounded inverse theorem. Thus, for a function $g \in C^{\alpha}(\bar{\Omega})$ of sufficiently small norm (in the space $C^{\alpha}(\bar{\Omega})$) there exists a unique solution of (1.13) which lies near the zero function. There may, of course, be other solutions which are not close to the zero function. (Note that the condition $f'(0) = 0$ rules out linear functions.)

Remark. Note that the choice of $X = Z = C(\bar{\Omega})$, $Y = \{u \in C^2(\bar{\Omega}) : u|_{\partial\Omega} = 0\}$ would fail above since the corresponding linear problem is not onto. An alternate approach would be to use Sobolev spaces. In fact, if we take $X = Z = W^{k-2}(\Omega)$, $Y = W^k(\Omega) \cap H_0^1(\Omega)$ with k sufficiently large, and if $f(\xi) \in C^{k+1}(\mathbb{R})$, then as above, we can conclude the existence of a unique solution $u \in W^k(\Omega)$ provided $\|g\|_{k-2,2}$ is sufficiently small. Hence, we get existence for more general functions g; however, the solution $u \in W^k(\Omega)$ is not a classical (i.e., C^2) solution in general.

1.5.5. Generalized Weierstrass Theorem. In its simplest form, the classical Weierstrass theorem can be stated as follows: Every continuous function defined on a closed ball in $\mathbb{R}^n$ is bounded and attains both its maximum and minimum on this ball. The proof makes essential use of the fact that the closed ball is compact.

The first difficulty in trying to extend this result to an arbitrary Banach space X is that the closed ball in X is not compact if X is infinite dimensional. However, as we shall show, a generalized Weierstrass theorem is possible if we require a stronger property for the functional.

A set $S \subset X$ is said to be **weakly closed** if $\{u_n\} \subset S$, $u_n \rightharpoonup u$ implies $u \in S$, i.e., S contains all its weak limits. A weakly closed set is clearly closed, but not conversely. Indeed, the set $\{\sin nx\}_1^{\infty}$ in $L^2(0, \pi)$ has no limit point (because it cannot be Cauchy) so it is closed, but zero is a weak limit that does not belong to the set. It can be shown that every convex, closed set in a Banach space is weakly closed.

A functional $f : S \subset X \to \mathbb{R}$ is **weakly continuous** at $u_0 \in S$ if for every sequence $\{u_n\} \subset S$ with $u_n \rightharpoonup u_0$ it follows that $f(u_n) \to f(u_0)$. Clearly, every functional $f \in X^*$ is weakly continuous. A functional $f : S \subset X \to \mathbb{R}$ is **weakly lower semicontinuous (w.l.s.c.)** at $u_0 \in S$ if for every sequence $\{u_n\} \subset S$ for which $u_n \rightharpoonup u_0$ it follows that $f(u_0) \le \liminf_{n\to\infty} f(u_n)$. According to Theorem 1.19, the norm on a Banach space is w.l.s.c.. A functional $f : S \subset X \to \mathbb{R}$ is **weakly coercive** on S if $f(u) \to \infty$ as $\|u\| \to \infty$ on S.

Theorem 1.38. *Let X be a reflexive Banach space and $f : C \subset X \to \mathbb{R}$ be w.l.s.c. and assume*

(i) *C is a nonempty bounded weakly closed set in X or*

(ii) *C is a nonempty weakly closed set in X and f is weakly coercive on C.*

Then

(a) $\inf_{u\in C} f(u) > -\infty$*;*

(b) *there is at least one $u_0 \in C$ such that $f(u_0) = \inf_{u\in C} f(u)$.*

Moreover, if u_0 is an interior point of C and f is G-diff at u_0, then $f'(u_0) = 0$; that is, u_0 is a **critical point** *of f.*

Proof. Assume (i) and let $\{u_n\} \subset C$ be a minimizing sequence, i.e., $\lim_{n\to\infty} f(u_n) = \inf_{u\in C} f(u)$. The existence of such a sequence follows from the definition of inf. Since X is reflexive and C is bounded and weakly closed, there is a subsequence $\{u_{n'}\}$ and a $u_0 \in C$ such that $u_{n'} \rightharpoonup u_0$. But f is w.l.s.c. and so $f(u_0) \le \liminf_{n\to\infty} f(u_{n'}) = \inf_{u\in C} f(u)$, which proves (a). Since by definition, $f(u_0) \ge \inf_{u\in C} f(u)$, we get (b).

Assume (ii) and fix $u_0 \in C$. Since f is weakly coercive, there is a closed ball $B(0,R) \subset X$ such that $u_0 \in B \cap C$ and $f(u) \ge f(u_0)$ outside $B\cap C$. Since $B\cap C$ satisfies the conditions of (i), there is a $u_1 \in B\cap C$ such that $f(u) \ge f(u_1)$ for all $u \in B\cap C$ and in particular for u_0. Thus, $f(u) \ge f(u_1)$ on all of C.

To prove the last statement we set $\varphi_v(t) = f(u_0 + tv)$. For fixed $v \in X$, $\varphi_v(t)$ has a local minimum at $t = 0$, and therefore $\langle f'(u_0), v\rangle = 0$ for all $v \in X$. □

Remark. Even though weakly continuous functionals on closed balls attain both their inf and sup (which follows from the above theorem), the usual functionals that we encounter are not weakly continuous, but are w.l.s.c.. Hence this explains why we seek the inf and not the sup in variational problems.

A set C in a real normed space X is called **convex** if $(1-t)u + tv \in C$ for all $t \in [0,1]$, u, $v \in C$. The following result is needed later.

Theorem 1.39. *A closed convex set in a Banach space is weakly closed.*

1.5.6. Monotone Operators and Convex Functionals. Let $A : X \to X^*$ be an operator, where X is a real Banach space. We say that

(i) A is **monotone** if

$$\langle Au - Av, u - v\rangle \ge 0 \quad \text{for all} \quad u, v \in X.$$

(ii) A is **strongly monotone** if for some $c > 0$ and $p > 1$,

$$\langle Au - Av, u - v\rangle \ge c\|u - v\|_X^p \quad \text{for all} \quad u, v \in X.$$

(iii) A is **coercive** if

$$\lim_{\|u\|\to\infty} \frac{\langle Au, u\rangle}{\|u\|} = +\infty.$$

Remark. A strongly monotone operator is coercive. This follows immediately from $\langle Au, u\rangle = \langle Au - A0, u\rangle + \langle A0, u\rangle \geq c\|u\|_X^p - \|A0\|\|u\|_X$.

Let C be a convex set in the real normed space X. A functional $f : C \subset X \to \mathbb{R}$ is said to be **convex** if

$$f((1-t)u + tv) \leq (1-t)f(u) + tf(v) \quad \text{for all} \quad t \in [0,1], \;\; u, v \in C.$$

In the following we set

$$\varphi(t) = f((1-t)u + tv) = f(u + t(v-u))$$

for fixed u and v.

Lemma 1.40. *Let $C \subset X$ be a convex set in a real normed space X. Then the following statements are equivalent:*

(a) *The real function $\varphi : [0,1] \to \mathbb{R}$ is convex for all $u, v \in C$.*

(b) *The functional $f : C \subset X \to \mathbb{R}$ is convex.*

(c) *$f' : C \subset X \to X^*$ (assuming f is G-diff on C) is monotone.*

Proof. Assume φ is convex. Then

$$\varphi(t) = \varphi((1-t)\cdot 0 + t \cdot 1) \leq (1-t)\varphi(0) + t\varphi(1)$$

for all $t \in [0,1]$, which implies (b).

Similarly, if f is convex, then for $t = (1-\alpha)s_1 + \alpha s_2$, with $\alpha, s_1, s_2 \in [0,1]$, we have

$$\varphi(t) = f(u + t(v-u)) \leq (1-\alpha)f(u + s_1(v-u)) + \alpha f(u + s_2(v-u))$$

for all $u, v \in C$, which implies (a).

Fix $u, v \in C$. Then $\varphi'(t) = \langle f'(u + t(v-u)), v-u\rangle$. If f is convex, then φ is convex and therefore φ' is monotone. From $\varphi'(1) \geq \varphi'(0)$ we obtain

$$\langle f'(v) - f'(u), v - u\rangle \geq 0 \quad \text{for all} \quad u, v \in C$$

which implies (c).

Finally, assume f' is monotone. Then for $s < t$ we have

$$\varphi'(t) - \varphi'(s) = \frac{1}{t-s}\langle f'(u + t(v-u)) - f'(u + s(v-u)), (t-s)(v-u)\rangle \geq 0.$$

Thus φ' is monotone, which implies φ, and thus f is convex. □

Theorem 1.41. *Consider the functional $f : C \subset X \to \mathbb{R}$, where X is a real Banach space. Then f is w.l.s.c. if any one of the following conditions holds:*

(a) *C is closed and convex; f is convex and continuous.*

(b) *C is convex; f is G-diff on C and f' is monotone on C.*

Proof. Set

$$C_r = \{u \in C : \; f(u) \leq r\}.$$

It follows from (a) that C_r is closed and convex for all r, and thus is weakly closed (cf. Theorem 1.39). If f is not w.l.s.c., then there is a sequence $\{u_n\} \subset C$ with $u_n \rightharpoonup u$ and $f(u) > \liminf f(u_n)$. Hence, there is an r and a subsequence $\{u_{n'}\}$ such that $f(u) > r$ and $f(u_{n'}) \leq r$ (i.e., $u_{n'} \in C_r$) for all n' large enough. Since C_r is weakly closed, $u \in C_r$, which is a contradiction.

Assume (b) holds and set $\varphi(t) = f(u + t(v-u))$. Then by Lemma 1.40, $\varphi : [0,1] \to \mathbb{R}$ is convex and φ' is monotone. By the classical mean value theorem,

$$\varphi(1) - \varphi(0) = \varphi'(\theta) \geq \varphi'(0), \quad 0 < \theta < 1$$

i.e.,

$$f(v) \geq f(u) + \langle f'(u), v-u \rangle \quad \text{for all} \quad u, v \in C.$$

If $u_n \rightharpoonup u$, then $\langle f'(u), u_n - u \rangle \to 0$ as $n \to \infty$. Hence, f is w.l.s.c. □

1.5.7. Lagrange Multipliers. Let $f, g : X \to \mathbb{R}$ be two functionals defined on the Banach space X and let

$$M_c = \{u \in X : g(u) = c\}$$

for a given constant c. A point $u_0 \in M_c$ is called a local **extremal point** of f with respect to M_c if there exists a neighborhood of u_0, $U(u_0) \subset X$, such that

$$f(u) \leq f(u_0) \quad \text{for all} \quad u \in U(u_0) \cap M_c$$

or

$$f(u) \geq f(u_0) \quad \text{for all} \quad u \in U(u_0) \cap M_c.$$

In the first case we say that f has (local) **maximum** at u_0 with respect to M_c, while in the second case f has (local) **minimum** at u_0 with respect to M_c.

A point $u_0 \in M_c$ is called an **ordinary point** of the manifold M_c if F-derivative $g'(u_0)$ exists and $g'(u_0) \neq 0$.

Let u_0 be an ordinary point of M_c. Then u_0 is called a **critical point** of f with respect to M_c if there exists a real number λ, called a **Lagrange multiplier**, such that

$$f'(u_0) = \lambda g'(u_0).$$

As we shall see, if u_0 is an extremal point of f with respect to M_c, and if u_0 is an ordinary point, then u_0 is a critical point of f with respect to M_c. Note that if u_0 is an extremal point of f with respect to X, then Lagrange multiplier $\lambda = 0$, which implies the usual result.

Lemma 1.42. *Let X be a Banach space. Suppose the following hold:*

(i) *$f, g : X \to \mathbb{R}$ are of class C^1*

(ii) *For $u_0 \in X$, we can find $v, w \in X$ such that*

$$f'(u_0)v \cdot g'(u_0)w \neq f'(u_0)w \cdot g'(u_0)v. \tag{1.14}$$

Then f cannot have a local extremum with respect to the level set M_c at u_0.

Proof. Fix $v, w \in X$, and for $s, t \in \mathbb{R}$ consider the real-valued functions

$$F(s,t) = f(u_0 + sv + tw), \;\; G(s,t) = g(u_0 + sv + tw) - c.$$

Then

$$\frac{\partial F}{\partial s}(0,0) = f'(u_0)v, \;\; \frac{\partial F}{\partial t}(0,0) = f'(u_0)w$$
$$\frac{\partial G}{\partial s}(0,0) = g'(u_0)v, \;\; \frac{\partial G}{\partial t}(0,0) = g'(u_0)w$$

so that condition (1.14) is simply that the Jacobian $|\partial(F,G)/\partial(s,t)|$ is nonvanishing at $(s,t) = (0,0)$. Since $F, G \in C^1$ on $\mathbb{R}^2$, we may apply the implicit function theorem to conclude that a local extremum cannot occur at u_0. More precisely, assume w.l.o.g. that

$G_t(0,0) \neq 0$. Since $G(0,0) = 0$, the implicit function theorem implies the existence of a C^1 function ϕ such that $\phi(0) = 0$ and $G(s, \phi(s)) = 0$ for sufficiently small s. Moreover,

$$\phi'(0) = -\frac{G_s(0,0)}{G_t(0,0)}.$$

Set $z(s) = F(s, \phi(s)) = f(u_0 + sv + \phi(s)w)$ and note that $g(u_0 + sv + \phi(s)w) = c$. Hence, if to the contrary f has an extremum at u_0, then $z(s)$ has a local extremum at $s = 0$. But, an easy computation shows that $G_t(0,0)z'(0) = f'(u_0)v \cdot g'(u_0)w - f'(u_0)w \cdot g'(u_0)v \neq 0$, which is a contradiction. □

Theorem 1.43. (Lagrange) *Let X be a Banach space. Suppose the following conditions hold:*

(i) *$f, g : X \to \mathbb{R}$ are of class C^1,*

(ii) *$g(u_0) = c$,*

(iii) *u_0 is a local extremal point of f with respect to the constraint M_c.*

Then either

(a) *$g'(u_0)v = 0$ for all $v \in X$, or*

(b) *there exists $\lambda \in \mathbb{R}$ such that $f'(u_0)v = \lambda g'(u_0)v$ for all $v \in X$.*

Proof. If (a) does not hold, then fix $w \in X$ with $g'(u_0)w \neq 0$. By hypothesis and the above lemma, we must have

$$f'(u_0)v \cdot g'(u_0)w = f'(u_0)w \cdot g'(u_0)v \quad \text{for all} \quad v \in X.$$

If we define $\lambda = (f'(u_0)w)/(g'(u_0)w)$, then we obtain (b). □

More generally, one can prove the following:

Theorem 1.44. (Ljusternik) *Let X be a Banach space. Suppose the following hold:*

(i) *$g_0 : X \to \mathbb{R}$ is of class C^1*

(ii) *$g_i : X \to \mathbb{R}$ are of class C^1, $i = 1, \ldots, n$*

(iii) *u_0 is a local extremal point of g_0 with respect to the constraint C:*

$$C = \{u : g_i(u) = c_i \ (i = 1, \ldots, n)\}$$

where the c_i are constants.

Then there are numbers λ_i (not all zero) such that

$$\sum_{i=0}^{n} \lambda_i g_i'(u_0) = 0. \tag{1.15}$$

As an application of Ljusternik's theorem we have

Theorem 1.45. *Let $f, g : X \to \mathbb{R}$ be C^1 functionals on the reflexive Banach space X. Suppose*

(i) *f is w.l.s.c. and weakly coercive on $X \cap \{g(u) \leq c\}$*

(ii) *g is weakly continuous*

(iii) *$g(0) = 0$, $g'(u) = 0$ only at $u = 0$.*

Then the equation $f'(u) = \lambda g'(u)$ has a one parameter family of nontrivial solutions (u_R, λ_R) for all $R \neq 0$ in the range of $g(u)$ and $g(u_R) = R$. Moreover, u_R can be characterized as the function which minimizes $f(u)$ over the set $g(u) = R$.

Proof. Since $g(u)$ is weakly continuous, it follows that $M_R = \{u : g(u) = R\}$ is weakly closed. If M_R is not empty, i.e., if R belongs to the range of g, then by Theorem 1.38, there is a $u_R \in M_R$ such that $f(u_R) = \inf f(u)$ over $u \in M_R$. If $R \neq 0$ then it cannot be that $g'(u_R) = 0$. Otherwise by (iii), $u_R = 0$ and hence $R = g(u_R) = 0$, which is a contradiction. Thus, by Ljusternik's theorem, there exist constants λ_1, λ_2, $\lambda_1^2 + \lambda_2^2 \neq 0$ such that $\lambda_1 f'(u_R) + \lambda_2 g'(u_R) = 0$. Since u_R is an ordinary point, it follows that $\lambda_1 \neq 0$, and therefore $\lambda_R = -\lambda_2/\lambda_1$. □

Remark. In applying this theorem one should be careful and not choose $g(u) = \|u\|$, since this g is not weakly continuous.

The following interpolation inequality, often referred to as **Ehrling's inequality**, will be useful in Sobolev spaces.

Theorem 1.46. *Let X, Y, Z be three Banach spaces such that*

$$X \subset Y \subset Z.$$

Assume that the embedding $X \subset Y$ is compact and the embedding $Y \subset Z$ is continuous. Then for each $\varepsilon > 0$, there is a constant $c(\varepsilon)$ such that

$$\|u\|_Y \leq \varepsilon \|u\|_X + c(\varepsilon)\|u\|_Z \quad \textit{for all} \quad u \in X. \tag{1.16}$$

Proof. If for a fixed $\varepsilon > 0$ the inequality is false, then there exists a sequence $\{u_n\}$ such that

$$\|u_n\|_Y > \varepsilon \|u_n\|_X + n\|u_n\|_Z \quad \text{for all} \quad n. \tag{1.17}$$

As $u_n \neq 0$, without loss of generality, we can assume $\|u_n\|_X = 1$. Since the embedding $X \subset Y$ is compact, there is a subsequence, again denoted by $\{u_n\}$, with $u_n \to u$ in Y. This implies $u_n \to u$ in Z. By (1.17), $\|u_n\|_Y > \varepsilon$ and so $u \neq 0$. Again by (1.17), $u_n \to 0$ in Z, i.e., $u = 0$, which is a contradiction. □

Sobolev Spaces

This chapter is devoted to a discussion of the necessary Sobolev function spaces which permit a modern approach to the study of differential equations.

2.1. Weak Derivatives and Sobolev Spaces

2.1.1. Weak Derivatives. Let Ω be a nonempty open set in $\mathbb{R}^n$. Suppose $u \in C^m(\Omega)$ and $\varphi \in C_0^m(\Omega)$. Then by integration by parts

$$\int_\Omega u D^\alpha \varphi dx = (-1)^{|\alpha|} \int_\Omega v \varphi dx, \quad |\alpha| \leq m \tag{2.1}$$

where $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ is an n-tuple and $v = D^\alpha u = \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \cdots \partial_{x_n}^{\alpha_n} u$.

Motivated by (2.1), we now enlarge the class of functions for which the notion of derivative can be generalized.

Let $u \in L^1_{loc}(\Omega)$. A function $v \in L^1_{loc}(\Omega)$ is called the α^{th} **weak derivative** of u if it satisfies

$$\int_\Omega u D^\alpha \varphi dx = (-1)^{|\alpha|} \int_\Omega v \varphi dx \ \text{ for all } \ \varphi \in C_0^{|\alpha|}(\Omega). \tag{2.2}$$

It can be easily shown that weak derivatives are unique. Thus we write $v = D^\alpha u$ to indicate that v is the α^{th} weak derivative of u. If a function u has an ordinary α^{th} derivative lying in $L^1_{loc}(\Omega)$, then it is clearly the α^{th} weak derivative.

In contrast to the corresponding classical derivative, the weak derivative $D^\alpha u$ is defined globally on all of Ω by (2.2). However, in every subregion $\Omega' \subset \Omega$ the function $D^\alpha u$ will also be the weak derivative of u. It suffices to note that (2.2) holds for every function $\varphi \in C_0^{|\alpha|}(\Omega')$, and extended outside Ω' by assigning to it the value zero. In particular, the weak derivative (if it exists) of a function u having compact support in Ω has itself compact support in Ω and thus belongs to $L^1(\Omega)$.

We also note that in contrast to the classical derivative, the weak derivative $D^\alpha u$ is defined at once for order $|\alpha|$ without assuming the existence of corresponding derivatives of lower orders. In fact, the derivatives of lower orders may not exist even we have a higher order weak derivative exists.

EXAMPLE 2.1. (a) Consider the function $u(x) = \mathrm{sgn} x_1 + \mathrm{sgn} x_2$ in the ball $\Omega = B(0,1) \subset \mathbb{R}^2$. Show that the weak derivative u_{x_1} does not exist, but the weak derivative $u_{x_1x_2}$ does exist and equals 0. In fact, for any $\varphi \in C_0^2(\Omega)$

$$\int_\Omega u\varphi_{x_1x_2}dx = 2\int_{\Omega^+}\varphi_{x_1x_2}dx - 2\int_{\Omega^-}\varphi_{x_1x_2}dx = 2\varphi(0,0) - 2\varphi(0,0) = 0,$$

where $\Omega^+ = \Omega \cap (x_1 > 0, x_2 > 0)$, $\Omega^- = \Omega \cap (x_1 < 0, x_2 < 0)$. However, $\int_\Omega u\varphi_{x_1} = -2\int_{-1}^1 \varphi(0,x_2)dx_2$, which can not be written as $\int_\Omega v\varphi\, dx$ for any function v.

(b) The function $u(x) = |x_1|$ has in the ball $\Omega = B(0,1)$ weak derivatives $u_{x_1} = \mathrm{sgn}\ x_1, u_{x_i} = 0, i = 2,\ldots,n$. In fact, we apply formula (2.2) as follows: For any $\varphi \in C_0^1(\Omega)$

$$\int_\Omega |x_1|\varphi_{x_1}dx = \int_{\Omega^+} x_1\varphi_{x_1}dx - \int_{\Omega^-} x_1\varphi_{x_1}dx$$

where $\Omega^+ = \Omega \cap (x_1 > 0)$, $\Omega^- = \Omega \cap (x_1 < 0)$. Since $x_1\varphi = 0$ on $\partial\Omega$ and also for $x_1 = 0$, an application of the divergence theorem yields

$$\int_\Omega |x_1|\varphi_{x_1}dx = -\int_{\Omega^+}\varphi dx + \int_{\Omega^-}\varphi dx = -\int_\Omega(\mathrm{sgn}\ x_1)\varphi dx.$$

Hence $|x_1|_{x_1} = \mathrm{sgn}\ x_1$. Similarly, since for $i \geq 2$

$$\int_\Omega |x_1|\varphi_{x_i}dx = \int_\Omega(|x_1|\varphi)_{x_i}dx = -\int_\Omega 0\varphi dx$$

$|x_1|_{x_i} = 0$ for $i = 2,\ldots,n$. Note that the function $|x_1|$ has no classical derivative with respect to x_1 in Ω.

(c) Let $\Omega = B(0,1/2) \subset \mathbb{R}^2$ and define $u(x) = \ln(\ln(2/r))$, $x \in \Omega$, where $r = |x| = (x_1^2 + x_2^2)^{1/2}$. Then $u \notin L^\infty(\Omega)$ because of the singularity at the origin. However, we will show that u has weak first partial derivatives; in fact all first weak derivatives are in $L^2(\Omega)$.

First of all $u \in L^2(\Omega)$, for

$$\int_\Omega |u|^2dx = \int_0^{2\pi}\int_0^{1/2} r[\ln(\ln(2/r))]^2dr\ d\theta$$

and a simple application of L'hopitals rule shows that the integrand is bounded and thus the integral is finite. Similarly, it is easy to check that the classical partial derivative

$$u_{x_1} = \frac{-\cos\theta}{r\ln(2/r)}, \quad \text{where } x_1 = r\cos\theta$$

also belongs to $L^2(\Omega)$. Now we show that the defining equation for the weak derivative is met.

Let $\Omega_\varepsilon = \{x : \varepsilon < r < 1/2\}$ and choose $\varphi \in C_0^1(\Omega)$. Then by the divergence theorem and the absolute continuity of integrals

$$\int_\Omega u\varphi_{x_1}dx = \lim_{\varepsilon\to 0}\int_{\Omega_\varepsilon} u\varphi_{x_1}dx = \lim_{\varepsilon\to 0}\left[-\int_{\Omega_\varepsilon} u_{x_1}\varphi dx + \int_{r=\varepsilon} u\varphi n_1 ds\right]$$

where $n = (n_1, n_2)$ is the unit outward normal to Ω_ε on $r = \varepsilon$. But $(ds = \varepsilon d\theta)$

$$|\int_{r=\varepsilon} u\varphi n_1 ds| \leq \int_0^{2\pi}|u|\ |\varphi|\varepsilon d\theta \leq 2\pi\varepsilon c\ln(\ln(2/\varepsilon)) \to 0$$

as $\varepsilon \to 0$. Thus

$$\int_\Omega u\varphi_{x_1}dx = -\int_\Omega u_{x_1}\varphi dx.$$

The same analysis applies to u_{x_2}. Thus u has weak first partial derivatives given by the classical derivatives which are defined on $\Omega\backslash\{0\}$.

2.1.2. Sobolev Spaces. For $p \geq 1$ and k a nonnegative integer, we define

$$W^{k,p}(\Omega) = \{u \in L^p(\Omega)\colon D^\alpha u \in L^p(\Omega),\ \forall\ 0 < |\alpha| \leq k\},$$

where $D^\alpha u$ denotes the α^{th} weak derivative. When $k = 0$, $W^{k,p}(\Omega)$ will mean $L^p(\Omega)$. It is clear that $W^{k,p}(\Omega)$ is a vector space.

A norm is introduced by defining

$$\|u\|_{k,p} = \|u\|_{W^{k,p}(\Omega)} = \left(\int_\Omega \sum_{|\alpha|\leq k} |D^\alpha u|^p dx\right)^{1/p} \tag{2.3}$$

if $1 \leq p < \infty$. For $p = \infty$, we define norm in terms of $\||D^\alpha u|\|_{L^\infty(\Omega)}$. The space $W^{k,p}(\Omega)$ is known as a **Sobolev space of order** k.

We also introduce the space $W_0^{k,p}(\Omega)$ which is defined to be the closure of the space $C_0^k(\Omega)$ with respect to the norm $\|\cdot\|_{k,p}$. As we shall see shortly, $W^{k,p}(\Omega) \neq W_0^{k,p}(\Omega)$ for $k \geq 1$. (Unless $\Omega = \mathbb{R}^n$.)

Remark. The case $p = 2$ is special, since the spaces $W^{k,2}(\Omega)$, $W_0^{k,2}(\Omega)$ will be Hilbert spaces under the inner product

$$(u,v)_{k,2} = (u,v)_{W^{k,2}(\Omega)} = \int_\Omega \sum_{|\alpha|\leq k} D^\alpha u D^\alpha v dx.$$

Since we shall be dealing mostly with these spaces in the sequel, we introduce the special notation:

$$H^k(\Omega) = W^{k,2}(\Omega),\quad H_0^k(\Omega) = W_0^{k,2}(\Omega).$$

Theorem 2.2. *For $1 \leq p \leq \infty$, the space $W^{k,p}(\Omega)$ is a Banach space under the norm defined. If $1 < p < \infty$, it is reflexive; if $1 \leq p < \infty$, it is separable.*

Proof. We only prove the case for $1 \leq p < \infty$; the case when $p = \infty$ is similar.
We first prove that $W^{k,p}(\Omega)$ is complete with respect to the norm (2.3).

Let $\{u_n\}$ be a Cauchy sequence of elements in $W^{k,p}(\Omega)$, i.e.,

$$\|u_n - u_m\|_{k,p}^p = \sum_{|\alpha|\leq k} \int_\Omega |D^\alpha u_n - D^\alpha u_m|^p dx \to 0 \ \text{ as } \ m,n \to \infty.$$

Then for any $\alpha,\ |\alpha| \leq k$, when $m,n \to \infty$

$$\int_\Omega |D^\alpha u_n - D^\alpha u_m|^p dx \to 0$$

and, in particular, when $|\alpha| = 0$

$$\int_\Omega |u_n - u_m|^p dx \to 0.$$

Since $L^p(\Omega)$ is complete, it follows that there are functions $u^\alpha \in L^p(\Omega)$, $|\alpha| \leq k$ such that $D^\alpha u_n \to u^\alpha$ (in $L^p(\Omega)$). Since each $u_n(x)$ has weak derivatives (up to order k) belonging to $L^p(\Omega)$, a simple limit argument shows that u^α is the α^{th} weak derivative of u^0. In fact,

$$\int_\Omega u D^\alpha \varphi dx \leftarrow \int_\Omega u_n D^\alpha \varphi dx = (-1)^{|\alpha|} \int_\Omega \varphi D^\alpha u_n dx \to (-1)^{|\alpha|} \int_\Omega u^\alpha \varphi dx.$$

Hence $u^0 \in W^{k,p}(\Omega)$ and $\|u_n - u^0\|_{k,p} \to 0$ as $n \to \infty$.

Consider the map $T : W^{1,p}(\Omega) \to (L^p(\Omega))^{n+1}$ defined by

$$Tu = (u, D_1 u, \ldots, D_n u).$$

If we endow the latter space with the norm

$$\|v\| = (\sum_{i=1}^{n+1} \|v_i\|_p^p)^{1/p}$$

for $v = (v_1, \ldots, v_{n+1}) \in (L^p(\Omega))^{n+1}$, then T is a (linear) isometry. Now $(L^p(\Omega))^{n+1}$ is reflexive for $1 < p < \infty$ and separable for $1 \le p < \infty$. Since $W^{1,p}(\Omega)$ is complete, its image under the isometry T is a closed subspace of $(L^p(\Omega))^{n+1}$ which inherits the corresponding properties as does $W^{1,p}(\Omega)$. Similarly, we can handle the case $k \ge 2$. $\square$

EXAMPLE 2.3. Let Ω be a bounded open connected set in $\mathbb{R}^n$. Divide Ω into N open disjoint subsets $\Omega_1, \Omega_2, \ldots, \Omega_N$. Suppose the function $u : \Omega \to \mathbb{R}$ has the following properties:

(i) u is continuous on $\bar{\Omega}$.

(ii) For some i, $D_i u$ is continuous on $\Omega_1, \Omega_2, \ldots, \Omega_N$, and can be extended continuously to $\bar{\Omega}_1, \bar{\Omega}_2, \ldots, \bar{\Omega}_N$, respectively.

(iii) The surfaces of discontinuity are such that the divergence theorem applies.

Define $w_i(x) = D_i u(x)$ if $x \in \cup_{i=1}^N \Omega_i$. Otherwise, w_i can be arbitrary. We now claim that $w_i \in L^p(\Omega)$ is a weak partial derivative of u on Ω.

Indeed, for all $\varphi \in C_0^1(\Omega)$, the divergence theorem yields

$$\begin{aligned}
\int_\Omega u D_i \varphi dx &= \sum_j \int_{\Omega_j} u D_i \varphi dx \\
&= \sum_j \left(\int_{\partial\Omega_j} u \varphi n_i dS - \int_{\Omega_j} \varphi D_i u dx \right) \\
&= -\int_\Omega \varphi D_i u dx.
\end{aligned}$$

Note that the boundary terms either vanish, since φ has compact support, or cancel out along the common boundaries, since u is continuous and the outer normals have opposite directions. Similarly, if $u \in C^k(\bar{\Omega})$ and has piecewise continuous derivatives in Ω of order $k+1$, then $u \in W^{k+1,p}(\Omega)$.

Remark. More generally, by using a partition of unity argument, we can show the following: If $\mathcal{O}$ is a collection of nonempty open sets whose union is Ω and if $u \in L^1_{loc}(\Omega)$ is such that for some multi-index α, the α^{th} weak derivative of u exists on each member of $\mathcal{O}$, then the α^{th} weak derivative of u exists on Ω.

2.2. Approximations and Extensions

2.2.1. Approximations. Let $x \in \mathbb{R}^n$ and let $B(x,h)$ denote the open ball with center at x and radius h. For each $h > 0$, let $\omega_h(x) \in C^\infty(\mathbb{R}^n)$ satisfy

$$\omega_h(x) \ge 0; \;\; \omega_h(x) = 0 \;\text{ for }\; |x| \ge h,$$

$$\int_{\mathbb{R}^n} \omega_h(x) dx = \int_{B(0,h)} \omega_h(x) dx = 1.$$

Such functions are called **mollifiers**. For example, let

$$\omega(x) = \begin{cases} k \exp\,[(|x|^2 - 1)^{-1}], & |x| < 1, \\ 0, & |x| \geq 1, \end{cases}$$

where $k > 0$ is chosen so that $\int_{\mathbb{R}^n} \omega(x)\,dx = 1$. Then, a family of mollifiers can be taken as $\omega_h(x) = h^{-n}\omega(x/h)$ for $h > 0$.

Let Ω be a nonempty open set in $\mathbb{R}^n$ and let $u \in L^1(\Omega)$. We set $u = 0$ outside Ω. Define for each $h > 0$ the **mollified function**

$$u_h(x) = \int_\Omega \omega_h(x-y)u(y)dy$$

where ω_h is a mollifier. There are two other forms in which u_h can be represented, namely

$$u_h(x) = \int_{\mathbb{R}^n} \omega_h(x-y)u(y)dy = \int_{B(x,h)} \omega_h(x-y)u(y)dy \tag{2.4}$$

the latter equality being valid since ω_h vanishes outside the (open) ball $B(x,h)$. Thus the values of $u_h(x)$ depend only on the values of u on the ball $B(x,h)$. In particular, if $\operatorname{dist}(x, \operatorname{supp}(u)) \geq h$, then $u_h(x) = 0$.

Theorem 2.4. *Let Ω be a nonempty open set in $\mathbb{R}^n$. Then*

(a) $u_h \in C^\infty(\mathbb{R}^n)$.

(b) *If* $\operatorname{supp}(u)$ *is a compact subset of* Ω*, then* $u_h \in C_0^\infty(\Omega)$ *for all* h *sufficiently small.*

Proof. Since u is integrable and $\omega_h \in C^\infty$, the Lebesgue theorem on differentiating integrals implies that for $|\alpha| < \infty$

$$D^\alpha u_h(x) = \int_\Omega u(y) D^\alpha \omega_h(x-y)dy$$

i.e., $u_h \in C^\infty(\mathbb{R}^n)$. Statement (b) follows from the remark preceding the theorem. $\square$

With respect to a bounded set Ω we construct another set $\Omega^{(h)}$ as follows: with each point $x \in \Omega$ as center, draw a ball of radius h; the union of these balls is then $\Omega^{(h)}$. Clearly $\Omega^{(h)} \supset \Omega$. Moreover, u_h can be different from zero only in $\Omega^{(h)}$.

Corollary 2.5. *Let Ω be a nonempty bounded open set in $\mathbb{R}^n$ and let $h > 0$ be any number. Then there exists a function $\eta \in C^\infty(\mathbb{R}^n)$ such that*

$$0 \leq \eta(x) \leq 1; \;\; \eta(x) = 1, \; x \in \Omega^{(h)}; \;\; \eta(x) = 0, \; x \in (\Omega^{(3h)})^c.$$

Such a function is called a **cut-off function** *for* Ω.

Proof. Let $\chi(x)$ be the characteristic function of the set $\Omega^{(2h)} : \chi(x) = 1$ for $x \in \Omega^{(2h)}$, $\chi(x) = 0$ for $x \notin \Omega^{(2h)}$ and set

$$\eta(x) \equiv \chi_h(x) = \int_{\mathbb{R}^n} \omega_h(x-y)\chi(y)dy.$$

Then

$$\eta(x) = \int_{\Omega^{(2h)}} \omega_h(x-y)dy \in C^\infty(\mathbb{R}^n),$$

$$0 \leq \eta(x) \leq \int_{\mathbb{R}^n} \omega_h(x-y)dy = 1,$$

and

$$\eta(x) = \int_{B(x,h)} \omega_h(x-y)\chi(y)dy = \begin{cases} \int_{B(x,h)} \omega_h(x-y)dy = 1, & x \in \Omega^{(h)}, \\ 0, & x \in (\Omega^{(3h)})^c. \end{cases}$$

In particular, we note that if $\Omega' \subset\subset \Omega$, there is a function $\eta \in C_0^\infty(\Omega)$ such that $\eta(x) = 1$ for $x \in \overline{\Omega'}$, and $0 \le \eta(x) \le 1$ in Ω. □

Remark. Henceforth, the notation $\Omega' \subset\subset \Omega$ means that Ω', Ω are open sets and that $\overline{\Omega'} \subset \Omega$.

We shall have need of the following well-known result.

Theorem 2.6. (Partition of Unity) *Assume $\Omega \subset \mathbb{R}^n$ is bounded and $\Omega \subset\subset \cup_{i=1}^N \Omega_i$, where each Ω_i is open. Then there exist C^∞ functions $\psi_i(x)(i = 1, \ldots, N)$ such that*

(a) $0 \le \psi_i(x) \le 1$,

(b) *ψ_i has its support in Ω_i,*

(c) *$\sum_{i=1}^N \psi_i(x) = 1$ for every $x \in \Omega$.*

Lemma 2.7. *Let Ω be a nonempty bounded open set in $\mathbb{R}^n$. Then every $u \in L^p(\Omega)$ is* **p-mean continuous***, i.e.,*

$$\int_\Omega |u(x+z) - u(x)|^p dx \to 0 \ \text{ as } \ z \to 0.$$

Proof. Choose $a > 0$ large enough so that Ω is strictly contained in the ball $B(0, a)$. Then the function

$$U(x) = \begin{cases} u(x) & \text{if } x \in \Omega, \\ 0 & \text{if } x \in B(0, 2a) \setminus \Omega \end{cases}$$

belongs to $L^p(B(0, 2a))$. For $\varepsilon > 0$, there is a function $\bar{U} \in C(\bar{B}(0, 2a))$ which satisfies the inequality $\|U - \bar{U}\|_{L^p(B(0,2a))} < \varepsilon/3$. By multiplying $\bar{U}$ by an appropriate cut-off function, it can be assumed that $\bar{U}(x) = 0$ for $x \in B(0, 2a)/B(0, a)$.

Therefore for $|z| \le a$, $\|U(x+z) - \bar{U}(x+z)\|_{L^p(B(0,2a))} = \|U(x) - \bar{U}(x)\|_{L^p(B(0,a))} \le \varepsilon/3$. Since the function $\bar{U}$ is uniformly continuous in $B(0, 2a)$, there is a $\delta > 0 (\delta < a)$ such that $\|\bar{U}(x+z) - \bar{U}(x)\|_{L^p(B(0,2a))} \le \varepsilon/3$ whenever $|z| < \delta$. Hence for $|z| < \delta$ we easily see that $\|u(x+z) - u(x)\|_{L^p(\Omega)} = \|U(x+z) - U(x)\|_{L^p(B(0,2a))} \le \varepsilon$. □

Theorem 2.8. *Let Ω be a nonempty open set in $\mathbb{R}^n$. If $u \in L^p(\Omega)$ $(1 \le p < \infty)$, then*

(a) $\|u_h\|_p \le \|u\|_p$

(b) $\|u_h - u\|_p \to 0 \quad \textit{as} \quad h \to 0.$

If $u \in C^k(\bar{\Omega})$, where $\bar{\Omega}$ is compact, then

(c) $\|u_h - u\|_{C^k(\bar{\Omega}')} \to 0 \quad \textit{as} \quad h \to 0,$

where $\Omega' \subset\subset \Omega$.

Proof. If $1 < p < \infty$, let $q = p/(p-1)$. Then $\omega_h = \omega_h^{1/p}\omega_h^{1/q}$ and Hölder's inequality implies

$$\begin{aligned} |u_h(x)|^p &\leq \int_\Omega \omega_h(x-y)|u(y)|^p dy \left(\int_\Omega \omega_h(x-y)dy\right)^{p/q} \\ &\leq \int_\Omega \omega_h(x-y)|u(y)|^p dy \end{aligned}$$

which obviously holds also for $p = 1$. An application of Fubini's Theorem gives

$$\int_\Omega |u_h(x)|^p dx \leq \int_\Omega \left(\int_\Omega \omega_h(x-y)dx\right)|u(y)|^p dy \leq \int_\Omega |u(y)|^p dy$$

which implies (a).

To prove (b), let $\omega(x) = h^n \omega_h(hx)$. Then $\omega(x) \in C^\infty(\mathbb{R}^n)$ and satisfies

$$\omega(x) \geq 0; \;\; \omega(x) = 0 \;\text{ for }\; |x| \geq 1$$

$$\int_{\mathbb{R}^n} \omega(x)dx = \int_{B(0,1)} \omega(x)dx = 1.$$

Using the change of variable $z = (x-y)/h$ we have

$$\begin{aligned} u_h(x) - u(x) &= \int_{B(x,h)} [u(y) - u(x)]\omega_h(x-y)dy \\ &= \int_{B(0,1)} [u(x-hz) - u(x)]\omega(z)dz. \end{aligned}$$

Hence by Hölder's inequality

$$|u_h(x) - u(x)|^p \leq d\int_{B(0,1)} |u(x-hz) - u(x)|^p dz$$

and so by Fubini's Theorem

$$\int_\Omega |u_h(x) - u(x)|^p dx \leq d\int_{B(0,1)} (\int_\Omega |u(x-hz) - u(x)|^p dx)dz.$$

The right-hand side goes to zero as $h \to 0$ since every $u \in L^p(\Omega)$ is p-mean continuous.

We now prove (c) for $k = 0$. Let Ω', Ω'' be such that $\Omega' \subset\subset \Omega'' \subset\subset \Omega$. Let h_0 be the shortest distance between $\partial\Omega'$ and $\partial\Omega''$. Take $h < h_0$. Then

$$u_h(x) - u(x) = \int_{B(x,h)} [u(y) - u(x)]\omega_h(x-y)dy.$$

If $x \in \bar{\Omega}'$, then in the above integral $y \in \bar{\Omega}''$. Now u is uniformly continuous in $\bar{\Omega}''$ and $\omega_h \geq 0$, and therefore for an arbitrary $\varepsilon > 0$ we have

$$|u_h(x) - u(x)| \leq \varepsilon \int_{B(x,h)} \omega_h(x-y)dy = \varepsilon$$

provided h is sufficiently small. The case $k \geq 1$ is handled similarly and is left as an exercise. $\square$

Remark. The following example shows that in (c) we cannot replace Ω' by Ω. Let $u \equiv 1$ for $x \in [0,1]$ and consider $u_h(x) = \int_0^1 \omega_h(x-y)dy$, where $\omega_h(y) = \omega_h(-y)$. Now $\int_{-h}^{h} \omega_h(y)dy = 1$ and so $u_h(0) = 1/2$ for all $h < 1$. Thus $u_h(0) \to 1/2 \neq 1 = u(0)$. Moreover, for $x \in (0,1)$ and h sufficiently small, $(x-h, x+h) \subset (0,1)$ and so $u_h(x) = \int_{x-h}^{x+h} \omega_h(x-y)dy = 1$ which implies $u_h(x) \to 1$ for all $x \in (0,1)$.

Corollary 2.9. *Let Ω be a nonempty open set in $\mathbb{R}^n$. Then $C_0^\infty(\Omega)$ is dense in $L^p(\Omega)$ for all $1 \leq p < \infty$.*

Proof. Suppose first that Ω is bounded and let $\Omega' \subset\subset \Omega$. For a given $u \in L^p(\Omega)$ set

$$v(x) = \begin{cases} u(x), & x \in \Omega' \\ 0, & x \in \Omega \backslash \Omega'. \end{cases}$$

Then

$$\int_\Omega |u-v|^p dx = \int_{\Omega \backslash \Omega'} |u|^p dx.$$

By the absolute continuity of integrals, we can choose Ω' so that the integral on the right is arbitrarily small, i.e., $\|u - v\|_p < \varepsilon/2$.

Since $\text{supp}(v)$ is a compact subset of Ω, Theorems 2.4(b) and 2.8(b) imply that for h sufficiently small, $v_h(x) \in C_0^\infty(\Omega)$ with $\|v - v_h\|_p < \varepsilon/2$, and therefore $\|u - v_h\|_p < \varepsilon$.

If Ω is unbounded, choose a ball B large enough so that

$$\int_{\Omega \backslash \Omega'} |u|^p dx < \varepsilon/2$$

where $\Omega' = \Omega \cap B$, and repeat the proof just given. □

We now consider the following local approximation theorem.

Theorem 2.10. *Let Ω be a nonempty open set in $\mathbb{R}^n$ and suppose $u, v \in L^1_{loc}(\Omega)$. Then $v = D^\alpha u$ iff, for each compact set $S \subset \Omega$, there exists a sequence of $C^\infty(\Omega)$ functions $\{u_h\}$ with $\|u_h - u\|_{L^1(S)} \to 0$, $\|D^\alpha u_h - v\|_{L^1(S)} \to 0$ as $h \to 0$.*

Proof. (Necessity) Suppose $v = D^\alpha u$. Let $S \subset \Omega$ be compact, and choose $d > 0$ small enough so that the sets $\Omega' \equiv S^{(d/2)}, \Omega'' \equiv S^{(d)}$ satisfy $\Omega' \subset\subset \Omega'' \subset\subset \Omega$. For $x \in \mathbb{R}^n$ define

$$u_h(x) = \int_{\Omega''} \omega_h(x-y)u(y)dy, \quad v_h(x) = \int_{\Omega''} \omega_h(x-y)v(y)dy.$$

Clearly, $u_h, v_h \in C^\infty(\mathbb{R}^n)$ for $h > 0$. Moreover, from Theorem 2.8 we have $\|u_h - u\|_{L^1(S)} \leq \|u_h - u\|_{L^1(\Omega'')} \to 0$.

Now we note that if $x \in \Omega'$ and $0 < h < d/2$, then $\omega_h(x-y) \in C_0^\infty(\Omega'')$. Thus by Theorem 2.4 and the definition of weak derivative,

$$\begin{aligned} D^\alpha u_h(x) &= \int_{\Omega''} u(y) D_x^\alpha \omega_h(x-y)dy = (-1)^{|\alpha|} \int_{\Omega''} u(y) D_y^\alpha \omega_h(x-y)dy \\ &= \int_{\Omega''} \omega_h(x-y) \cdot v(y)dy = v_h(x). \end{aligned}$$

Thus, $\|D^\alpha u_h - v\|_{L^1(S)} \to 0$.

(Sufficiency) Choose $\varphi \in C_0^{|\alpha|}(\Omega)$ and consider a compact set $S \supset \text{supp}(\varphi)$. Then as $h \to \infty$

$$\int_S u D^\alpha \varphi dx \leftarrow \int_S u_h D^\alpha \varphi dx = (-1)^{|\alpha|} \int_S \varphi D^\alpha u_h dx \to (-1)^{|\alpha|} \int_S v \varphi dx$$

which is the claim. □

If u is equal to a constant (a.e.) in Ω, then u has the weak derivative $D^\alpha u = 0,\ \ |\alpha| > 0$. An application of Theorem 2.10 yields the converse:

Theorem 2.11. *Let Ω be a bounded open connected set in $\mathbb{R}^n$. If $u \in L^1_{loc}(\Omega)$ has a weak derivative $D^\alpha u = 0$ whenever $|\alpha| = 1$, then $u =$const. a.e. in Ω.*

Proof. Let $\Omega' \subset\subset \Omega$. Then for $x \in \Omega'$ and with u_h as in Theorem 2.10, $D^\alpha u_h(x) = (D^\alpha u)_h(x) = 0$ for all h sufficiently small. Thus $u_h = \text{const} = c(h)$ in Ω' for such h. Since $\|u_h - u\|_{L^1(\Omega')} = \|c(h) - u\|_{L^1(\Omega')} \to 0$ as $h \to 0$, it follows that

$$\|c(h_1) - c(h_2)\|_{L^1(\Omega')} = |c(h_1) - c(h_2)|\text{mes}(\Omega') \to 0 \ \text{ as } \ h_1, h_2 \to 0.$$

Consequently, $c(h) = u_h$ converges uniformly and thus in $L^1(\Omega')$ to some constant. Hence $u = \text{const(a.e.)}$ in Ω' and therefore also in Ω, by virtue of it being connected. □

We now note some properties of $W^{k,p}(\Omega)$ which follow easily from the results of this and the previous section.

(a) If $\Omega' \subset \Omega$ and if $u \in W^{k,p}(\Omega)$, then $u \in W^{k,p}(\Omega')$.

(b) If $u \in W^{k,p}(\Omega)$ and $|a(x)|_{k,\infty} < \infty$, then $au \in W^{k,p}(\Omega)$. In this case any weak derivative $D^\alpha(au)$ is computed according to the usual rule of differentiating the product of functions.

(c) If $u \in W^{k,p}(\Omega)$ and u_h is its mollified function, then for any compact set $S \subset \Omega$, $\|u_h - u\|_{W^{k,p}(S)} \to 0$ as $h \to 0$. If in addition, u has compact support in Ω, then $\|u_h - u\|_{k,p} \to 0$ as $h \to 0$.

More generally, we have the following global approximation theorems. (The proofs make use of a partition of unity argument; see Evans's book.)

Theorem 2.12. (Meyers-Serrin) *Assume Ω is bounded and let $u \in W^{k,p}(\Omega),\ 1 \le p < \infty$. Then there exist functions $u_m \in C^\infty(\Omega) \cap W^{k,p}(\Omega)$ such that*

$$u_m \to u \ \ in \ \ W^{k,p}(\Omega).$$

In other words, $C^\infty(\Omega) \cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$.

Theorem 2.13. *Assume Ω is bounded and $\partial\Omega \in C^1$. Let $u \in W^{k,p}(\Omega),\ 1 \le p < \infty$. Then there exist functions $u_m \in C^\infty(\bar{\Omega})$ such that*

$$u_m \to u \ \ in \ \ W^{k,p}(\Omega).$$

In other words, $C^\infty(\bar{\Omega})$ is dense in $W^{k,p}(\Omega)$.

Example 2.14. (a) Prove the product rule for weak derivatives:

$$D_i(uv) = (D_i u)v + u(D_i v)$$

where $u, D_i u$ are locally $L^p(\Omega)$, $v, D_i v$ are locally $L^q(\Omega)$ $(p > 1, 1/p + 1/q = 1)$.

(b) If $u \in W_0^{k,p}(\Omega)$ and $v \in C^k(\bar{\Omega})$, prove that $uv \in W_0^{k,p}(\Omega)$.

(c) If $u \in W^{k,p}(\Omega)$ and $v \in C_0^k(\Omega)$, prove that $uv \in W_0^{k,p}(\Omega)$.

Theorem 2.15. (Chain Rule) *Let Ω be a bounded open set in $\mathbb{R}^n$. Let $f \in C^1(\mathbb{R}),\ \ |f'(s)| \le M$ for all $s \in \mathbb{R}$ and suppose u has a weak derivative $D^\alpha u$ for $|\alpha| = 1$. Then the composite function $f \circ u$ has a weak derivative $D^\alpha(f \circ u) = f'(u)D^\alpha u$. Moreover, if $f(0) = 0$ and if $u \in W^{1,p}(\Omega)$, then $f \circ u \in W^{1,p}(\Omega)$.*

Proof. According to Theorem 2.10, there exists a sequence $\{u_h\} \subset C^1(\Omega)$ such that $\|u_h - u\|_{L^1(\Omega')} \to 0$, $\|D^\alpha u_h - D^\alpha u\|_{L^1(\Omega')} \to 0$ as $h \to 0$, where $\Omega' \subset\subset \Omega$. Thus

$$\int_{\Omega'} |f(u_h) - f(u)| dx \leq \sup |f'| \int_{\Omega'} |u_h - u| dx \to 0 \text{ as } h \to 0$$

$$\int_{\Omega'} |f'(u_h) D^\alpha u_h - f'(u) D^\alpha u| dx \leq \sup |f'| \int_{\Omega'} |D^\alpha u_h - D^\alpha u| dx$$
$$+ \int_{\Omega'} |f'(u_h) - f'(u)||D^\alpha u| dx.$$

Since $\|u_h - u\|_{L^1(\Omega')} \to 0$, there exists a subsequence of $\{u_h\}$, which we call $\{u_h\}$ again, which converges a.e. in Ω' to u. Moreover, since f' is continuous, $\{f'(u_h)\}$ converges to $f'(u)$ a.e. in Ω'. Hence the last integral tends to zero by the dominated convergence theorem. Consequently, the sequences $\{f(u_h)\}$, $\{f'(u_h) D^\alpha u_h\}$ tend to $f(u)$, $f'(u) D^\alpha u$ respectively, and the first conclusion follows by an application of Theorem 2.10 again.

Since $f(0) = 0$, the mean value theorem implies $|f(s)| \leq M|s|$ for all $s \in \mathbb{R}$. Thus, $|f(u(x))| \leq M|u(x)|$ for all $x \in \Omega$ and so $f \circ u \in L^p(\Omega)$ if $u \in L^p(\Omega)$. Similarly, $f'(u(x)) D^\alpha u \in L^p(\Omega)$ if $u \in W^{1,p}(\Omega)$, which shows that $f \circ u \in W^{1,p}(\Omega)$. □

Corollary 2.16. *Let Ω be a bounded open set in $\mathbb{R}^n$. If u has an α^{th} weak derivative $D^\alpha u$, $|\alpha| = 1$, then so does $|u|$ and*

$$D^\alpha |u| = \begin{cases} D^\alpha u & \text{if } u > 0 \\ 0 & \text{if } u = 0 \\ -D^\alpha u & \text{if } u < 0 \end{cases}$$

i.e., $D^\alpha |u| = (sgn\ u) D^\alpha u$ for $u \neq 0$. In particular, if $u \in W^{1,p}(\Omega)$, then $|u| \in W^{1,p}(\Omega)$.

Proof. The positive and negative parts of u are defined by

$$u^+ = \max\{u, 0\}, \quad u^- = \min\{u, 0\}.$$

If we can show that $D^\alpha u^+$ exists and that

$$D^\alpha u^+ = \begin{cases} D^\alpha u & \text{if } u > 0 \\ 0 & \text{if } u \leq 0 \end{cases}$$

then the result for $|u|$ follows easily from the relations $|u| = u^+ - u^-$ and $u^- = -(-u)^+$. Thus, for $h > 0$ define

$$f_h(u) = \begin{cases} (u^2 + h^2)^{\frac{1}{2}} - h & \text{if } u > 0 \\ 0 & \text{if } u \leq 0. \end{cases}$$

Clearly $f_h \in C^1(\mathbb{R})$ and f_h' is bounded on $\mathbb{R}$. By Theorem 2.15, $f_h(u)$ has a weak derivative, and for any $\varphi \in C_0^1(\Omega)$

$$\int_\Omega f_h(u) D^\alpha \varphi dx = -\int_\Omega D^\alpha(f_h(u)) \varphi dx = -\int_{u>0} \varphi \frac{u D^\alpha u}{(u^2 + h^2)^{\frac{1}{2}}} dx.$$

Upon letting $h \to 0$, it follows that $f_h(u) \to u^+$, and so by the dominating convergence theorem

$$\int_\Omega u^+ D^\alpha \varphi dx = -\int_{u>0} \varphi D^\alpha u dx = -\int_\Omega v \varphi dx$$

where

$$v = \begin{cases} D^\alpha u & \text{if } u > 0 \\ 0 & \text{if } u \leq 0 \end{cases}$$

which establishes the desired result for u^+. □

Remark. Since $u = u^+ + u^-$, we have $\partial u/\partial x_i = \partial u^+/\partial x_i + \partial u^-/\partial x_i$. Consequently, $\partial u/\partial x_i = 0$ a.e. on $\{u = c\} = \{x \in \Omega : u(x) = c\}$.

The following result is of independent importance.

Theorem 2.17. *$u_n \rightharpoonup u$ in $W^{k,p}(\Omega)$, if and only if $D^\alpha u_n \rightharpoonup D^\alpha u$ in $L^p(\Omega)$ for all $|\alpha| \le k$.*

Theorem 2.18. *Let $f : \mathbb{R} \to \mathbb{R}$ be Lipschitz continuous with $f(0) = 0$. Then if Ω is a bounded open set in $\mathbb{R}^n$, $1 < p < \infty$ and $u \in W_0^{1,p}(\Omega)$, we have $f \circ u \in W_0^{1,p}(\Omega)$.*

Proof. Given $u \in W_0^{1,p}(\Omega)$, let $u_n \in C_0^1(\Omega)$ with $\|u_n - u\|_{1,p} \to 0$ and define $v_n = f \circ u_n$. Since u_n has compact support and $f(0) = 0$, v_n has compact support. Also v_n is Lipschitz continuous, for

$$|v_n(x) - v_n(y)| = |f(u_n(x)) - f(u_n(y))| \le c|u_n(x) - u_n(y)| \le c_n|x - y|.$$

Hence $v_n \in L^p(\Omega)$. Since v_n is absolutely continuous on any line segment in Ω, its partial derivatives (which exist almost everywhere) coincide almost everywhere with the weak derivatives. Moreover, we see from above that $|\partial v_n/\partial x_i| \le c_n$ for $1 \le i \le n$, and as Ω is bounded, $\partial v_n/\partial x_i \in L^p(\Omega)$. Thus $v_n \in W^{1,p}(\Omega)$ and has compact support, which implies $v_n \in W_0^{1,p}(\Omega)$. From the relation

$$|v_n(x) - f(u(x))| \le c|u_n(x) - u(x)|$$

it follows that $\|v_n - f \circ u\|_p \to 0$. Furthermore, if e_i is the standard ith basis vector in $\mathbb{R}^n$, we have

$$\frac{|v_n(x + he_i) - v_n(x)|}{|h|} \le c\frac{|u_n(x + he_i) - u_n(x)|}{|h|}$$

and so

$$\limsup_{n\to\infty} \|\frac{\partial v_n}{\partial x_i}\|_p \le c \limsup_{n\to\infty} \|\frac{\partial u_n}{\partial x_i}\|_p.$$

But, $\{\partial u_n/\partial x_i\}$ is a convergent sequence in $L^p(\Omega)$ and therefore $\{\partial v_n/\partial x_i\}$ is bounded in $L^p(\Omega)$ for each $1 \le i \le n$. Since $\|v_n\|_{1,p}$ is bounded and $W_0^{1,p}(\Omega)$ is reflexive, a subsequence of $\{v_n\}$ converges weakly in $W^{1,p}(\Omega)$, and thus weakly in $L^p(\Omega)$ to some element of $W_0^{1,p}(\Omega)$. Thus, $f \circ u \in W_0^{1,p}(\Omega)$. □

Remark. In terms of the trace operator (defined later) we have $\gamma_0(f \circ u) = f \circ \gamma_0(u)$.

Corollary 2.19. *Let $u \in W_0^{1,p}(\Omega)$. Then $|u|$, u^+, $u^- \in W_0^{1,p}(\Omega)$.*

Proof. We apply the preceding theorem with $f(t) = |t|$. Thus $|u| \in W_0^{1,p}(\Omega)$. Now $u^+ = (|u| + u)/2$ and $u^- = (u - |u|)/2$. Thus u^+, $u^- \in W_0^{1,p}(\Omega)$. □

2.2.2. Extensions. If $\Omega \subset \Omega'$, then any function $u(x) \in C_0^k(\Omega)$ has an obvious extension $U(x) \in C_0^k(\Omega')$ by zero outside of Ω. From the definition of $W_0^{k,p}(\Omega)$ it follows that the function $u(x) \in W_0^{k,p}(\Omega)$ and extended as being equal to zero in $\Omega' \backslash \Omega$ belongs to $W_0^{k,p}(\Omega')$. In general, a function $u \in W^{k,p}(\Omega)$ and extended by zero to Ω' will not belong to $W^{k,p}(\Omega')$. (Consider the function $u(x) \equiv 1$ in Ω.) This also shows that in general $W^{k,p}(\Omega) \ne W_0^{k,p}(\Omega)$.

We now consider a more general extension result.

Theorem 2.20. *Let Ω be a bounded open set in $\mathbb{R}^n$ with $\Omega \subset\subset \Omega'$ and assume $k \ge 1$.*

(a) *If* $\partial\Omega \in C^k$, *then any function* $u(x) \in W^{k,p}(\Omega)$ *has an extension* $U(x) \in W^{k,p}(\Omega')$ *into* Ω' *with compact support. Moreover,*

$$\|U\|_{W^{k,p}(\Omega')} \le c\|u\|_{W^{k,p}(\Omega)}$$

where the constant $c > 0$ *does not depend on* u.

(b) *If* $\partial\Omega \in C^k$, *then any function* $u(x) \in C^k(\bar{\Omega})$ *has an extension* $U(x) \in C_0^k(\Omega')$ *into* Ω' *with compact support. Moreover,*

$$\|U\|_{C^k(\bar{\Omega}')} \le c\|u\|_{C^k(\bar{\Omega})}, \quad \|U\|_{W^{k,p}(\Omega')} \le c\|u\|_{W^{k,p}(\Omega)}$$

where the constant $c > 0$ *does not depend on* u.

(c) *If* $\partial\Omega \in C^k$, *then any function* $u(x) \in C^k(\partial\Omega)$ *has an extension* $U(x)$ *into* Ω *which belongs to* $C^k(\bar{\Omega})$. *Moreover*

$$\|U\|_{C^k(\bar{\Omega})} \le c\|u\|_{C^k(\partial\Omega)}$$

where the constant $c > 0$ *does not depend on* u.

Proof. Suppose first that $u \in C^k(\bar{\Omega})$. Let $y = \psi(x)$ define a C^k diffeomorphism that straightens the boundary near $x^0 = (x_1^0, \ldots, x_n^0) \in \partial\Omega$. In particular, we assume there is a ball $B = B(x^0)$ such that $\psi(B \cap \Omega) \subset \mathbb{R}_+^n$ (i.e., $y_n > 0$), $\psi(B \cap \partial\Omega) \subset \partial\mathbb{R}_+^n$. (e.g., we could choose $y_i = x_i - x_i^0$ for $i = 1, \ldots, n-1$ and $y_n = x_n - \varphi(x_1, \ldots, x_{n-1})$, where φ is of class C^k. Moreover, without loss of generality, we can assume $y_n > 0$ if $x \in B \cap \Omega$.)

Let G and $G^+ = G \cap \mathbb{R}_+^n$ be respectively, a ball and half-ball in the image of ψ such that $\psi(x^0) \in G$. Setting $\bar{u}(y) = u \circ \psi^{-1}(y)$ and $y = (y_1, \ldots, y_{n-1}, y_n) = (y', y_n)$, we define an extension $\bar{U}(y)$ of $\bar{u}(y)$ into $y_n < 0$ by

$$\bar{U}(y', y_n) = \sum_{i=1}^{k+1} c_i \bar{u}(y', -y_n/i), \quad y_n < 0$$

where the c_i are constants determined by the system of equations

$$\sum_{i=1}^{k+1} c_i(-1/i)^m = 1, \quad m = 0, 1, \ldots, k. \tag{2.5}$$

Note that the determinant of the system (2.5) is nonzero since it is the Vandemonde determinant. One verifies readily that the extended function $\bar{U}$ is continuous with all derivatives up to order k in G. For example,

$$\lim_{y \to (y',0)} \bar{U}(y) = \sum_{i=1}^{k+1} c_i \bar{u}(y', 0) = \bar{u}(y', 0)$$

by virtue of (2.5) with $m = 0$. A similar computation shows that

$$\lim_{y \to (y',0)} \bar{U}_{y_i}(y) = \bar{u}_{y_i}(y', 0), \quad i = 1, \ldots, n-1.$$

Finally

$$\lim_{y \to (y',0)} \bar{U}_{y_n}(y) = \sum_{i=1}^{k+1} c_i(-1/i)\bar{u}_{y_n}(y', 0) = \bar{u}_{y_n}(y', 0)$$

by virtue of (2.5) with $m = 1$. Similarly we can handle the higher derivatives.

Thus $w = \bar{U} \circ \psi \in C^k(\overline{B'})$ for some ball $B' = B'(x^0)$ and $w = u$ in $B' \cap \Omega$, (If $x \in B' \cap \Omega$, then $\psi(x) \in G^+$ and $w(x) = \bar{U}(\psi(x)) = \bar{u}(\psi(x)) = u(\psi^{-1}\psi(x)) = u(x)$) so that w provides a C^k extension of u into $\Omega \cup B'$. Moreover,

$$\sup_{G^+} |\bar{u}(y)| = \sup_{G^+} |u(\psi^{-1}(y))| \le \sup_{\Omega} |u(x)|$$

and since $x \in B'$ implies $\psi(x) \in G$

$$\sup_{B'} |\bar{U}(\psi(x))| \le c \sup_{G^+} |\bar{u}(y)| \le c \sup_{\Omega} |u(x)|.$$

Since a similar computation for the derivatives holds, it follows that there is a constant $c > 0$, independent of u, such that

$$\|w\|_{C^k(\bar{\Omega} \cup B')} \le c \|u\|_{C^k(\bar{\Omega})}.$$

Now consider a finite covering of $\partial\Omega$ by balls B_i, $i = 1, \ldots, N$, such as B in the preceding, and let $\{w_i\}$ be the corresponding C^k extensions. We may assume the balls B_i are so small that their union with Ω is contained in $\Omega^{'}$. Let $\Omega_0 \subset\subset \Omega$ be such that Ω_0 and the balls B_i provide a finite open covering of Ω. Let $\{\eta_i\}$, $i = 1, \ldots, N$, be a partition of unity subordinate to this covering and set

$$w = u\eta_0 + \sum w_i \eta_i$$

with the understanding that $w_i\eta_i = 0$ if $\eta_i = 0$. Then w is an extension of u into $\Omega^{'}$ and has the required properties. Thus (b) is established.

(a) If $u \in W^{k,p}(\Omega)$, then by Theorem 2.13, there exist functions $u_m \in C^\infty(\bar{\Omega})$ such that $u_m \to u$ in $W^{k,p}(\Omega)$. Let $\Omega \subset \Omega'' \subset \Omega'$, and let U_m be the extension of u_m to Ω'' as given in (b). Then

$$\|U_m - U_l\|_{W^{k,p}(\Omega'')} \le c \|u_m - u_l\|_{W^{k,p}(\Omega)}$$

which implies that $\{U_m\}$ is a Cauchy sequence and so converges to a $U \in W_0^{k,p}(\Omega'')$, since $U_m \in C_0^k(\Omega'')$. Now extend U_m, U by 0 to Ω'. It is easy to see that U is the desired extension.

(c) At any point $x^0 \in \partial\Omega$ let the mapping ψ and the ball G be defined as in (b). By definition, $u \in C^k(\partial\Omega)$ implies that $\bar{u} = u \circ \psi^{-1} \in C^k(G \cap \partial\mathbb{R}^n_+)$. We define $\bar{\Phi}(y', y_n) = \bar{u}(y')$ in G and set $\Phi(x) = \bar{\Phi} \circ \psi(x)$ for $x \in \psi^{-1}(G)$. Clearly, $\Phi \in C^k(\bar{B})$ for some ball $B = B(x^0)$ and $\Phi = u$ on $B \cap \partial\Omega$. Now let $\{B_i\}$ be a finite covering of $\partial\Omega$ by balls such as B and let Φ_i be the corresponding C^k functions defined on B_i. For each i, we define the function $U_i(x)$ as follows: in the ball B_i take it equal to Φ_i, outside B_i take it equal to zero if $x \notin \partial\Omega$ and equal to $u(x)$ if $x \in \partial\Omega$. The proof can now be completed as in (b) by use of an appropriate partition of unity. $\square$

2.2.3. Trace Theorems. Unless otherwise stated, Ω will denote a bounded open connected set in $\mathbb{R}^n$, i.e., a bounded domain.

Let Γ be a surface which lies in $\bar{\Omega}$ and has the representation

$$x_n = \varphi(x'), \;\; x' = (x_1, \ldots, x_{n-1})$$

where $\varphi(x')$ is Lipschitz continuous in $\bar{U}$. Here U is the projection of Γ onto the coordinate plane $x_n = 0$. Let $p \ge 1$. A function u defined on Γ is said to belong to $L^p(\Gamma)$ if

$$\|u\|_{L^p(\Gamma)} \equiv \left(\int_\Gamma |u(x)|^p dS\right)^{\frac{1}{p}} < \infty$$

where

$$\int_\Gamma |u(x)|^p dS = \int_U |u(x', \varphi(x'))|^p [1 + \sum_{i=1}^{n-1} (\frac{\partial \varphi}{\partial x_i}(x'))^2]^{\frac{1}{2}} dx'.$$

Thus $L^p(\Gamma)$ reduces to a space of the type $L^p(U)$ where U is a domain in $\mathbb{R}^{n-1}$.

For every function $u \in C(\bar{\Omega})$, its values $\gamma_0 u \equiv u|_\Gamma$ on Γ are uniquely given. The function $\gamma_0 u$ will be called the **trace** of the function u on Γ. Note that $u \in L^p(\Gamma)$ since $\gamma_0 u \in C(\Gamma)$.

On the other hand, if we consider a function u defined a.e. in Ω (i.e., functions are considered equal if they coincide a.e.), then the values of u on Γ are not uniquely determined since $\text{meas}(\Gamma) = 0$. In particular, since $\partial\Omega$ has measure 0, there exist infinitely many extensions of u to $\bar{\Omega}$ that are equal a.e. We shall therefore introduce the concept of trace for functions in $W^{1,p}(\Omega)$ so that if in addition, $u \in C(\bar{\Omega})$, the new definition of trace reduces to the definition given above.

Lemma 2.21. *Let $\partial\Omega \in C^{0,1}$. Then for $u \in C^1(\bar{\Omega})$,*

$$\|\gamma_0 u\|_{L^p(\partial\Omega)} \le c\|u\|_{1,p} \tag{2.6}$$

where the constant $c > 0$ does not depend on u.

Proof. For simplicity, let $n = 2$. The more general case is handled similarly. In a neighborhood of a boundary point $x \in \partial\Omega$, we choose a local (ξ, η)-coordinate system, where the boundary has the local representation

$$\eta = \varphi(\xi), \quad -\alpha \le \xi \le \alpha$$

with the $C^{0,1}$ function φ. Then there exists a $\beta > 0$ such that all the points (ξ, η) with

$$-\alpha \le \xi \le \alpha, \quad \varphi(\xi) - \beta \le \eta \le \varphi(\xi)$$

belong to $\bar{\Omega}$. Let $u \in C^1(\bar{\Omega})$. Then

$$u(\xi, \varphi(\xi)) = \int_t^{\varphi(\xi)} u_\eta(\xi, \eta) d\eta + u(\xi, t)$$

where $\varphi(\xi) - \beta \le t \le \varphi(\xi)$. Applying the inequality $(a + b)^p \le 2^{p-1}(a^p + b^p)$ together with Hölder's inequality we have

$$|u(\xi, \varphi(\xi))|^p \le 2^{p-1}\beta^{p-1} \int_{\varphi(\xi)-\beta}^{\varphi(\xi)} |u_\eta(\xi, \eta)|^p d\eta + 2^{p-1}|u(\xi, t)|^p.$$

An integration with respect to t yields

$$\beta|u(\xi, \varphi(\xi))|^p \le 2^{p-1} \int_{\varphi(\xi)-\beta}^{\varphi(\xi)} [\beta^p|u_\eta(\xi, \eta)|^p + |u(\xi, \eta)|^p] d\eta.$$

Finally, integration over the interval $[-\alpha, \alpha]$ yields

$$\int_{-\alpha}^{\alpha} \beta|u(\xi, \varphi(\xi))|^p d\xi \le 2^{p-1} \int_S (\beta^p|u_\eta|^p + |u|^p) d\xi d\eta \tag{2.7}$$

where S denotes a local boundary strip.

Suppose $\varphi(\cdot)$ is C^1. Then the differential of arc length is given by $ds = (1 + \varphi'^2)^{1/2} d\xi$. Addition of the local inequalities (2.7) yields the assertion (2.6). Now if $\varphi(\cdot)$ is merely Lipschitz continuous, then the derivative φ' exists a.e. and is bounded. Thus we also obtain (2.6). $\square$

Since $\overline{C^1(\bar\Omega)} = W^{1,p}(\Omega)$, the bounded linear operator $\gamma_0 : C^1(\bar\Omega) \subset W^{1,p}(\Omega) \to L^p(\partial\Omega)$ can be uniquely extended to a bounded linear operator $\gamma_0 : W^{1,p}(\Omega) \to L^p(\partial\Omega)$ such that (2.6) remains true for all $u \in W^{1,p}(\Omega)$. More precisely, we obtain $\gamma_0 u$ in the following way: Let $u \in W^{1,p}(\Omega)$. We choose a sequence $\{u_n\} \subset C^1(\bar\Omega)$ with $\|u_n - u\|_{1,p} \to 0$. Then $\|\gamma_0 u_n - \gamma_0 u\|_{L^p(\partial\Omega)} \to 0$.

The function $\gamma_0 u$ (as an element of $L^p(\partial\Omega)$) will be called the **trace** of the function $u \in W^{1,p}(\Omega)$ on the boundary $\partial\Omega$. ($\|\gamma_0 u\|_{L^p(\partial\Omega)}$ will be denoted by $\|u\|_{L^p(\partial\Omega)}$.) Thus the trace of a function is defined for any element $u \in W^{1,p}(\Omega)$.

The above discussion partly proves the following:

Theorem 2.22. (Trace) *Suppose $\partial\Omega \in C^1$. Then there is a unique bounded linear operator $\gamma_0 : W^{1,p}(\Omega) \to L^p(\partial\Omega)$ such that $\gamma_0 u = u|_{\partial\Omega}$ for $u \in C(\bar\Omega) \cap W^{1,p}(\Omega)$, and $\gamma_0(au) = \gamma_0 a \cdot \gamma_0 u$ for $a(x) \in C^1(\bar\Omega)$, $u \in W^{1,p}(\Omega)$. Moreover, $\mathcal{N}(\gamma_0) = W_0^{1,p}(\Omega)$ and $\overline{\mathcal{R}(\gamma_0)} = L^p(\partial\Omega)$.*

Proof. Suppose $u \in C(\bar\Omega) \cap W^{1,p}(\Omega)$. Then by Theorem 2.20, u can be extended into $\Omega'(\Omega \subset\subset \Omega')$ such that its extension $U \in C(\bar\Omega') \cap W^{1,p}(\Omega')$. Let $U_h(x)$ be the mollified function for U. Since $U_h \to U$ as $h \to 0$ in both the norms $\|\cdot\|_{C(\bar\Omega)}$, $\|\cdot\|_{W^{1,p}(\Omega)}$, we find that as $h \to 0$, $U_h|_{\partial\Omega} \to u|_{\partial\Omega}$ uniformly and $U_h|_{\partial\Omega} \to \gamma_0 u$ in $L^p(\partial\Omega)$. Consequently, $\gamma_0 u = u|_{\partial\Omega}$.

Now $au \in W^{1,p}(\Omega)$ if $a \in C^1(\bar\Omega)$, $u \in W^{1,p}(\Omega)$ and consequently, $\gamma_0(au)$ is defined. Let $\{u_n\} \subset C^1(\bar\Omega)$ with $\|u_n - u\|_{1,p} \to 0$. Then

$$\gamma_0(au_n) = \gamma_0 a \cdot \gamma_0 u_n$$

and the desired product formula follows by virtue of the continuity of γ_0.

If $u \in W_0^{1,p}(\Omega)$, then there is a sequence $\{u_n\} \subset C_0^1(\Omega)$ with $\|u_n - u\|_{1,p} \to 0$. But $u_n|_{\partial\Omega} = 0$ and as $n \to \infty$, $u_n|_{\partial\Omega} \to \gamma_0 u$ in $L^p(\partial\Omega)$ which implies $\gamma_0 u = 0$. Hence $W_0^{1,p}(\Omega) \subset \mathcal{N}(\gamma_0)$. Now suppose $u \in \mathcal{N}(\gamma_0)$. If $u \in W^{1,p}(\Omega)$ has compact support in Ω, then by an earlier remark, $u \in W_0^{1,p}(\Omega)$. If u does not have compact support in Ω, then it can be shown that there exists a sequence of cut-off functions η_k such that $\eta_k u \in W^{1,p}(\Omega)$ has compact support in Ω, and moreover, $\|\eta_k u - u\|_{1,p} \to 0$. By using the corresponding mollified functions, it follows that $u \in W_0^{1,p}(\Omega)$ and $\mathcal{N}(\gamma_0) \subset W_0^{1,p}(\Omega)$. Details can be found in Evans's book.

To see that $\overline{\mathcal{R}(\gamma_0)} = L^p(\partial\Omega)$, let $f \in L^p(\partial\Omega)$ and let $\varepsilon > 0$ be given. Then there is a $u \in C^1(\partial\Omega)$ such that $\|u - f\|_{L^p(\partial\Omega)} < \varepsilon$. If we let $U \in C^1(\bar\Omega)$ be the extension of u into $\bar\Omega$, then clearly $\|\gamma_0 U - f\|_{L^p(\partial\Omega)} < \varepsilon$, which is the desired result since $U \in W^{1,p}(\Omega)$. ☐

Remarks. (i) The range $\mathcal{R}(\gamma_0)$ is the *fractional* Sobolev space $W^{1-\frac{1}{p},p}(\partial\Omega)$; see [**2**].

(ii) Note that the function $u \equiv 1$ belongs to $W^{1,p}(\Omega) \cap C(\bar\Omega)$ and its trace on $\partial\Omega$ is 1. Hence this function does not belong to $W_0^{1,p}(\Omega)$, which establishes the earlier assertion that $W_0^{1,p}(\Omega) \neq W^{1,p}(\Omega)$.

Let $u \in W^{k,p}(\Omega)$, $k > 1$. Since any weak derivative $D^\alpha u$ of order $|\alpha| < k$ belongs to $W^{1,p}(\Omega)$, this derivative has a trace $\gamma_0 D^\alpha u$ belonging to $L^p(\partial\Omega)$. Moreover

$$\|D^\alpha u\|_{L^p(\partial\Omega)} \ \le\ c\|D^\alpha u\|_{1,p} \ \le\ c\|u\|_{k,p}$$

for constant $c > 0$ independent of u.

Assuming the boundary $\partial\Omega \in C^1$, the unit outward normal vector $\mathbf{n}$ to $\partial\Omega$ exists and is bounded. Thus, the concept of traces makes it possible to introduce, for $k \ge 2$, $\partial u/\partial n$

for $u \in W^{k,p}(\Omega)$. More precisely, for $k \geq 2$, there exist traces of the functions u, $D_i u$ so that, if n_i are the direction cosines of the normal, we may define

$$\gamma_1 u = \sum_{i=1}^{n} (\gamma_0(D_i u)) n_i, \quad u \in W^{k,p}(\Omega), \quad k \geq 2.$$

The trace operator $\gamma_1 : W^{k,p}(\Omega) \to L^p(\partial\Omega)$ is continuous and $\gamma_1 u = (\partial u/\partial n)|_{\partial\Omega}$ for $u \in C^1(\bar{\Omega}) \cap W^{k,p}(\Omega)$.

For a function $u \in C^k(\bar{\Omega})$ we define the various traces of normal derivatives given by

$$\gamma_j u = \frac{\partial^j u}{\partial n^j}|_{\partial\Omega}, \quad 0 \leq j \leq k-1.$$

Each γ_j can be extended by continuity to all of $W^{k,p}(\Omega)$ and we obtain the following:

Theorem 2.23. (Higher Trace) *Suppose $\partial\Omega \in C^k$. Then there is a unique continuous linear operator $\gamma = (\gamma_0, \gamma_1, \ldots, \gamma_{k-1}) : W^{k,p}(\Omega) \to \prod_{j=0}^{k-1} W^{k-1-j,p}(\partial\Omega)$ such that for $u \in C^k(\bar{\Omega})$*

$$\gamma_0 u = u|_{\partial\Omega}, \ \gamma_j u = \frac{\partial^j u}{\partial n^j}|_{\partial\Omega}, \ j = 1, \ldots, k-1.$$

Moreover, $\mathcal{N}(\gamma) = W_0^{k,p}(\Omega)$ and $\overline{\mathcal{R}(\gamma)} = \prod_{j=0}^{k-1} W^{k-1-j,p}(\partial\Omega)$.

The Sobolev spaces $W^{k-1-j,p}(\partial\Omega)$, which are defined over $\partial\Omega$, can be defined locally.

Theorem 2.24. (Integration by Parts) *Let $u, v \in H^1(\Omega)$ and let $\partial\Omega \in C^1$. Then for any $i = 1, \ldots, n$*

$$\int_\Omega v D_i u dx = \int_{\partial\Omega} (\gamma_0 u \cdot \gamma_0 v) n_i dS - \int_\Omega u D_i v dx. \tag{2.8}$$

Proof. Let $\{u_n\}$ and $\{v_n\}$ be sequences of functions in $C^1(\bar{\Omega})$ with $\|u_n - u\|_{H^1(\Omega)} \to 0, \|v_n - v\|_{H^1(\Omega)} \to 0$ as $n \to \infty$. Formula (2.8) holds for u_n, v_n

$$\int_\Omega v_n D_i u_n dx = \int_{\partial\Omega} u_n v_n n_i dS - \int_\Omega u_n D_i v_n dx$$

and upon letting $n \to \infty$ relation (2.8) follows. □

Corollary 2.25. *Let $\partial\Omega \in C^1$.*

(a) *If $v \in H^1(\Omega)$ and $u \in H^2(\Omega)$ then*

$$\int_\Omega v \Delta u dx = \int_{\partial\Omega} \gamma_0 v \cdot \gamma_1 u dS - \int_\Omega (\nabla u \cdot \nabla v) dx \quad (\textbf{Green's 1st identity}).$$

(b) *If $u, v \in H^2(\Omega)$ then*

$$\int_\Omega (v \Delta u - u \Delta v) dx = \int_{\partial\Omega} (\gamma_0 v \cdot \gamma_1 u - \gamma_0 u \cdot \gamma_1 v) dS \quad (\textbf{Green's 2nd identity}).$$

In these formulas $\nabla u \equiv (D_1 u, \ldots, D_n u)$ is the gradient vector and $\Delta u \equiv \sum_{i=1}^n D_{ii} u$ is the **Laplace operator**.

Proof. If in (2.8) we replace u by $D_i u$ and sum from 1 to n, then Green's 1st identity is obtained. Interchanging the roles of u, v in Green's 1st identity and subtracting the two identities yields Green's 2nd identity. □

Example 2.26. Establish the following one-dimensional version of the trace theorem: If $u \in W^{1,p}(\Omega)$, where $\Omega = (a, b)$, then

$$\|u\|_{L^p(\partial\Omega)} \equiv (|u(a)|^p + |u(b)|^p)^{1/p} \leq \text{ const } \|u\|_{W^{1,p}(\Omega)}$$

where the constant is independent of u.

2.3. Sobolev Imbedding Theorems

We consider the following question: If a function u belongs to $W^{k,p}(\Omega)$, does u automatically belong to certain other spaces? The answer will be yes, but which other spaces depend upon whether $1 \leq kp < n$, $kp = n$, $n < kp < \infty$.

The result we want to prove is the following:

Theorem 2.27. (Sobolev-Rellich-Kondrachov) *Let $\Omega \subset \mathbb{R}^n$ be bounded and open with $\partial\Omega \in C^1$. Assume $1 \leq p < \infty$ and k is a positive integer.*

(a) *If $kp < n$ and $1 \leq q \leq np/(n - kp)$, then*

$$W^{k,p}(\Omega) \subset L^q(\Omega)$$

is a continuous imbedding; the imbedding is compact if $1 \leq q < np/(n - kp)$. Moreover,

$$\|u\|_{L^q(\Omega)} \leq C\|u\|_{W^{k,p}(\Omega)} \tag{2.9}$$

where the constant C depends only on k,p,n and Ω.

(b) *If $kp = n$ and $1 \leq r < \infty$, then*

$$W^{k,p}(\Omega) \subset L^r(\Omega)$$

is a compact imbedding and

$$\|u\|_{L^r(\Omega)} \leq C\|u\|_{W^{k,p}(\Omega)} \tag{2.10}$$

where the constant depends only on k,p,n and Ω.

(c) *If $kp > n$ and $0 \leq \alpha \leq k - m - n/p$, then*

$$W^{k,p}(\Omega) \subset C^{m,\alpha}(\bar{\Omega})$$

is a continuous imbedding; the imbedding is compact if $0 \leq \alpha < k - m - n/p$. Moreover,

$$\|u\|_{C^{m,\alpha}(\bar{\Omega})} \leq C\|u\|_{W^{k,p}(\Omega)} \tag{2.11}$$

where the constant C depends only on k,p,n,α and Ω.

(d) *Let $0 \leq j < k$, $1 \leq p, q < \infty$. Set $d = 1/p - (k - j)/n$. Then*

$$W^{k,p} \subset W^{j,q}$$

is a continuous imbedding for $d \leq 1/q$; the imbedding is compact for $d < 1/q$.

The above results are valid for $W_0^{k,p}(\Omega)$ spaces on arbitrary bounded domains Ω.

Remark. It is easy to check that the imbedding $W^{1,p}(\Omega) \subset L^p(\Omega)$ is compact for all $p \geq 1$ and all n.

A series of special results will be needed to prove the above theorem. Only selected proofs will be given to illustrate some of the important techniques.

Suppose $1 \leq p < n$. Do there exist constants $C > 0$ and $1 \leq q < \infty$ such that

$$\|u\|_{L^q(\mathbb{R}^n)} \leq C\|\nabla u\|_{L^p(\mathbb{R}^n)} \tag{2.12}$$

for all $u \in C_0^\infty(\mathbb{R}^n)$? The point is that the constants C and q should not depend on u.

We shall show that if such an inequality holds, then q must have a specific form. For this choose any $u \in C_0^\infty(\mathbb{R}^n),\ u \not\equiv 0$, and define for $\lambda > 0$

$$u_\lambda(x) \equiv u(\lambda x) \quad (x \in \mathbb{R}^n).$$

Now

$$\int_{\mathbb{R}^n} |u_\lambda|^q dx = \int_{\mathbb{R}^n} |u(\lambda x)|^q dx = \frac{1}{\lambda^n} \int_{\mathbb{R}^n} |u(y)|^q dy$$

and

$$\int_{\mathbb{R}^n} |\nabla u_\lambda|^p dx = \lambda^p \int_{\mathbb{R}^n} |\nabla u(\lambda x)|^p dx = \frac{\lambda^p}{\lambda^n} \int_{\mathbb{R}^n} |\nabla u(y)|^p dy.$$

Inserting these inequalities into (2.12) we find

$$\frac{1}{\lambda^{n/q}}\|u\|_{L^q(\mathbb{R}^n)} \leq C\frac{\lambda}{\lambda^{n/p}}\|\nabla u\|_{L^p(\mathbb{R}^n)}$$

and so

$$\|u\|_{L^q(\mathbb{R}^n)} \leq C\lambda^{1-n/p+n/q}\|\nabla u\|_{L^p(\mathbb{R}^n)}. \tag{2.13}$$

But then if $1 - n/p + n/q > 0 (< 0)$, we can upon sending λ to 0 (∞) in (2.13) obtain a contradiction $(u = 0)$. Thus we must have $q = p^*$ where

$$p^* = \frac{np}{n-p} \tag{2.14}$$

is called the **Sobolev conjugate** of p. Note that then

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}, \quad p^* > p. \tag{2.15}$$

Next we prove that the inequality (2.12) is in fact correct.

Lemma 2.28. (Gagliardo-Nirenberg-Sobolev Inequality) *Assume $1 \leq p < n$. Then there is a constant C, depending only on p and n, such that*

$$\|u\|_{L^{p^*}(\mathbb{R}^n)} \leq C\|\nabla u\|_{L^p(\mathbb{R}^n)} \tag{2.16}$$

for all $u \in C_0^1(\mathbb{R}^n)$.

Proof. First assume $p = 1$. Since u has compact support, for each $i = 1, \ldots, n$ we have

$$u(x) = \int_{-\infty}^{x_i} u_{x_i}(x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_n) dy_i$$

and so

$$|u(x)| \leq \int_{-\infty}^{\infty} |\nabla u(x_1, \ldots, y_i, \ldots, x_n)| dy_i \quad (i = 1, \ldots, n).$$

Consequently

$$|u(x)|^{\frac{n}{n-1}} \le \prod_{i=1}^{n} \left(\int_{-\infty}^{\infty} |\nabla u(x_1, \ldots, y_i, \ldots, x_n)| dy_i \right)^{\frac{1}{n-1}}. \tag{2.17}$$

Integrate this inequality with respect to x_1:

$$\begin{aligned}
\int_{-\infty}^{\infty} |u(x)|^{\frac{n}{n-1}} dx_1 &\le \int_{-\infty}^{\infty} \prod_{i=1}^{n} \left(\int_{-\infty}^{\infty} |\nabla u| dy_i \right)^{\frac{1}{n-1}} dx_1 \\
&= \left(\int_{-\infty}^{\infty} |\nabla u| dy_1 \right)^{\frac{1}{n-1}} \int_{-\infty}^{\infty} \prod_{i=2}^{n} \left(\int_{-\infty}^{\infty} |\nabla u| dy_i \right)^{\frac{1}{n-1}} dx_1 \\
&\le \left(\int_{-\infty}^{\infty} |\nabla u| dy_1 \right)^{\frac{1}{n-1}} \left(\prod_{i=2}^{n} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\nabla u| dx_1 dy_i \right)^{\frac{1}{n-1}}
\end{aligned}$$

the last inequality resulting from the **generalized Hölder inequality** (1.1), with $r = 1$, $k = n-1$ and $p_i = n-1$.

We continue by integrating with respect to $x_2, \ldots, x_n$ and applying the generalized Hölder inequality to eventually find (pull out an integral at each step)

$$\begin{aligned}
\int_{\mathbb{R}^n} |u(x)|^{\frac{n}{n-1}} dx &\le \prod_{i=1}^{n} \left(\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |\nabla u| dx_1 \ldots dy_i \ldots dx_n \right)^{\frac{1}{n-1}} \\
&= \left(\int_{\mathbb{R}^n} |\nabla u| dx \right)^{\frac{n}{n-1}}
\end{aligned}$$

which is estimate (2.16) for $p = 1$.

Consider now the case that $1 < p < n$. We shall apply the last estimate to $v = |u|^\gamma$, where $\gamma > 1$ is to be selected. Note that at a point x_0 where $u(x_0) \neq 0$

$$(D_i v)(x_0) = \begin{cases} \gamma u^{\gamma-1} D_i u & \text{if } u(x_0) > 0 \\ -\gamma(-u)^{\gamma-1} D_i u & \text{if } u(x_0) < 0 \end{cases}$$

If $u(x_0) = 0$, clearly $(D_i v)(x_0)$ exists at x_0 and equals 0. Thus $v \in C_0^1(\mathbb{R}^n)$, and

$$\begin{aligned}
\left(\int_{\mathbb{R}^n} |u(x)|^{\frac{\gamma n}{n-1}} dx \right)^{\frac{n-1}{n}} &\le \int_{\mathbb{R}^n} |\nabla |u|^\gamma| dx \\
&= \gamma \int_{\mathbb{R}^n} |u|^{\gamma-1} |\nabla u| dx \\
&\le \gamma \left(\int_{\mathbb{R}^n} |u|^{\frac{p(\gamma-1)}{p-1}} dx \right)^{\frac{p-1}{p}} \left(\int_{\mathbb{R}^n} |\nabla u|^p dx \right)^{\frac{1}{p}}.
\end{aligned}$$

We set

$$\gamma = \frac{p(n-1)}{n-p} > 1$$

in which case

$$\frac{\gamma n}{n-1} = \frac{p(\gamma-1)}{p-1} = \frac{np}{n-p} = p^*.$$

Thus, the above estimate becomes

$$\left(\int_{\mathbb{R}^n} |u|^{p^*} dx \right)^{\frac{1}{p^*}} \le C \left(\int_{\mathbb{R}^n} |\nabla u|^p dx \right)^{\frac{1}{p}}.$$

□

Theorem 2.29. *Let $\Omega \subset \mathbb{R}^n$ be bounded and open, with $\partial\Omega \in C^1$. Assume $1 \le p < n$, and $u \in W^{1,p}(\Omega)$. Then $u \in L^{p^*}(\Omega)$ and*

$$\|u\|_{L^{p^*}(\Omega)} \le C\|u\|_{W^{1,p}(\Omega)} \tag{2.18}$$

where the constant C depends only on p,n and Ω.

Proof. Since $\partial\Omega \in C^1$, there exists an extension $U \in W^{1,p}(\mathbb{R}^n)$ such that $U = u$ in Ω, U has compact support and

$$\|U\|_{W^{1,p}(\mathbb{R}^n)} \le C\|u\|_{W^{1,p}(\Omega)}. \tag{2.19}$$

Moreover, since U has compact support, there exist mollified functions $u_m \in C_0^\infty(\mathbb{R}^n)$ such that $u_m \to U$ in $W^{1,p}(\mathbb{R}^n)$. Now according to Lemma 2.28,

$$\|u_m - u_l\|_{L^{p^*}(\mathbb{R}^n)} \le C\|\nabla u_m - \nabla u_l\|_{L^p(\mathbb{R}^n)}$$

for all $l, m \ge 1$; whence $u_m \to U$ in $L^{p^*}(\mathbb{R}^n)$ as well. Since Lemma 2.28 also implies

$$\|u_m\|_{L^{p^*}(\mathbb{R}^n)} \le C\|\nabla u_m\|_{L^p(\mathbb{R}^n)}$$

we get in the limit that

$$\|U\|_{L^{p^*}(\mathbb{R}^n)} \le C\|\nabla U\|_{L^p(\mathbb{R}^n)}.$$

This inequality and (2.19) complete the proof. □

Theorem 2.30. *Let $\Omega \subset \mathbb{R}^n$ be bounded and open. Assume $1 \le p < n$, and $u \in W_0^{1,p}(\Omega)$. Then $u \in L^q(\Omega)$ and*

$$\|u\|_{L^q(\Omega)} \le C\|\nabla u\|_{L^p(\Omega)}$$

for each $q \in [1, p^*]$, the constant C depending only on p,q,n and Ω.

Proof. Since $u \in W_0^{1,p}(\Omega)$, there are functions $u_m \in C_0^\infty(\Omega)$ such that $u_m \to u$ in $W^{1,p}(\Omega)$. We extend each function u_m to be 0 in $\mathbb{R}^n \setminus \bar{\Omega}$ and apply Lemma 2.28 to discover (as above)

$$\|u\|_{L^{p^*}(\Omega)} \le C\|\nabla u\|_{L^p(\Omega)}.$$

Since $|\Omega| < \infty$, we furthermore have

$$\|u\|_{L^q(\Omega)} \le C\|u\|_{L^{p^*}(\Omega)}$$

for every $q \in [1, p^*]$. □

We now turn to the case $n < p < \infty$. The next result shows that if $u \in W^{1,p}(\Omega)$, then u is in fact Hölder continuous, after possibly being redefined on a set of measure zero.

Theorem 2.31. (Morrey's Inequality) *Assume $n < p < \infty$. Then there exists a constant C, depending only on p and n, such that*

$$\|u\|_{C^{0,1-\frac{n}{p}}(\mathbb{R}^n)} \le C\|u\|_{W^{1,p}(\mathbb{R}^n)}, \quad \forall\, u \in C^1(\mathbb{R}^n). \tag{2.20}$$

Proof. We first prove the following inequality: for all $x \in \mathbb{R}^n$, $r > 0$ and all $u \in C^1(\mathbb{R}^n)$,

$$\int_{B(x,r)} |u(y) - u(x)|\, dy \le \frac{r^n}{n} \int_{B(x,r)} \frac{|Du(y)|}{|x-y|^{n-1}}\, dy. \tag{2.21}$$

To prove this, note that, for any w with $|w| = 1$ and $0 < s < r$,

$$\begin{aligned} |u(x+sw) - u(x)| &= \left| \int_0^s \frac{d}{dt} u(x+tw) dt \right| \\ &= \left| \int_0^s \nabla u(x+tw) \cdot w dt \right| \\ &\le \int_0^s |\nabla u(x+sw)| \, dt. \end{aligned}$$

Now we integrate w over $\partial B(0,1)$ to obtain

$$\begin{aligned} \int_{\partial B(0,1)} |u(x+sw) - u(x)| \, dS &\le \int_0^s \int_{\partial B(0,1)} |\nabla u(x+sw)| \, dS dt \\ &= \int_{B(x,s)} \frac{|\nabla u(y)|}{|x-y|^{n-1}} \, dy \\ &\le \int_{B(x,r)} \frac{|\nabla u(y)|}{|x-y|^{n-1}} \, dy. \end{aligned}$$

Multiply both sides by s^{n-1} and integrate over $s \in (0, r)$ and we obtain (2.21).
To establish the bound on $\|u\|_{C^0(\mathbb{R}^n)}$, we observe that, by (2.21), for $x \in \mathbb{R}^n$,

$$\begin{aligned} |u(x)| &\le \frac{1}{|B(x,1)|} \int_{B(x,1)} |u(y) - u(x)| \, dy + \frac{1}{|B(x,1)|} \int_{B(x,1)} |u(y)| \, dy \\ &\le C \left(\int_{\mathbb{R}^n} |\nabla u(y)|^p \, dy \right)^{1/p} \left(\int_{B(x,1)} |y-x|^{\frac{(1-n)p}{p-1}} \, dy \right)^{\frac{p-1}{p}} + C\|u\|_{L^p(\mathbb{R}^n)} \\ &\le C\|u\|_{W^{1,p}(\mathbb{R}^n)}. \end{aligned}$$

To establish the bound on the semi-norm $[u]_\gamma$, $\gamma = 1 - \frac{n}{p}$, take any two points $x, y \in \mathbb{R}^n$. Let $r = |x-y|$ and $W = B(x,r) \cap B(y,r)$. Then

$$|u(x) - u(y)| \le \frac{1}{|W|} \int_W |u(x) - u(z)| \, dz + \frac{1}{|W|} \int_W |u(y) - u(z)| \, dz. \tag{2.22}$$

Note that $|W| = \beta r^n$, $r = |x-y|$ and $\int_W \le \min\{\int_{B(x,r)}, \int_{B(y,r)}\}$. Hence, using (2.21), by Hölder's inequality, we obtain

$$\begin{aligned} \int_W |u(x) - u(z)| \, dz &\le \int_{B(x,r)} |u(x) - u(z)| \, dz \le \frac{r^n}{n} \int_{B(x,r)} |Du(z)| |x-z|^{1-n} \, dz \\ &\le \frac{r^n}{n} \left(\int_{B(x,r)} |\nabla u(z)|^p \, dz \right)^{1/p} \left(\int_{B(x,r)} |z-x|^{\frac{(1-n)p}{p-1}} \, dz \right)^{\frac{p-1}{p}} \\ &\le C\, r^n \, \|\nabla u\|_{L^p(\mathbb{R}^n)} \left(\int_0^r s^{\frac{(1-n)p}{p-1}} \, s^{n-1} \, ds \right)^{\frac{p-1}{p}} \\ &\le C\, r^{n+\gamma} \|\nabla u\|_{L^p(\mathbb{R}^n)}, \end{aligned}$$

where $\gamma = 1 - \frac{n}{p}$; similarly,

$$\int_W |u(y) - u(z)| \, dz \le C\, r^{n+\gamma} \|\nabla u\|_{L^p(\mathbb{R}^n)}.$$

Hence, by (2.22),

$$|u(x) - u(y)| \le C\, |x-y|^\gamma \|\nabla u\|_{L^p(\mathbb{R}^n)}.$$

This inequality and the bound on $\|u\|_{C^0}$ above complete the proof. $\square$

Theorem 2.32. (Estimates for $W^{1,p}$, $n < p \le \infty$) *Let $\Omega \subset \mathbb{R}^n$ be bounded and open, with $\partial\Omega \in C^1$. Assume $n < p < \infty$, and $u \in W^{1,p}(\Omega)$. Then $u \in C^{0,1-\frac{n}{p}}(\bar{\Omega})$ a.e. and*

$$\|u\|_{C^{0,1-\frac{n}{p}}(\bar{\Omega})} \le C\|u\|_{W^{1,p}(\Omega)}$$

where the constant C depends only on p, n and Ω.

Proof. Since $\partial\Omega \in C^1$, there exists an extension $U \in W^{1,p}(\mathbb{R}^n)$ such that $U = u$ in Ω, U has compact support and

$$\|U\|_{W^{1,p}(\mathbb{R}^n)} \le C\|u\|_{W^{1,p}(\Omega)}. \tag{2.23}$$

Moreover, since U has compact support, there exist mollified functions $u_m \in C_0^\infty(\mathbb{R}^n)$ such that $u_m \to U$ in $W^{1,p}(\mathbb{R}^n)$ (and hence on compact subsets). Now according to Morrey's inequality,

$$\|u_m - u_l\|_{C^{0,1-n/p}(\mathbb{R}^n)} \le C\|u_m - u_l\|_{W^{1,p}(\mathbb{R}^n)}$$

for all $l, m \ge 1$; whence there is a function $u^* \in C^{0,1-n/p}(\mathbb{R}^n)$ such that $u_m \to u^*$ in $C^{0,1-n/p}(\mathbb{R}^n)$. Thus $u^* = u$ a.e. in Ω. Since we also have

$$\|u_m\|_{C^{0,1-n/p}(\mathbb{R}^n)} \le C\|u_m\|_{W^{1,p}(\mathbb{R}^n)}$$

we get in the limit that

$$\|u^*\|_{C^{0,1-n/p}(\mathbb{R}^n)} \le C\|U\|_{W^{1,p}(\mathbb{R}^n)}.$$

This inequality and (2.23) complete the proof. $\square$

We can now concatenate the above estimates to obtain more complicated inequalities.

Proof of general Sobolev inequalities. Assume $kp < n$ and $u \in W^{k,p}(\Omega)$. Since $D^\alpha u \in L^p(\Omega)$ for all $|\alpha| \le k$, the Sobolev-Nirenberg-Gagliardo inequality implies

$$\|D^\beta u\|_{L^{p^*}(\Omega)} \le C\|u\|_{W^{k,p}(\Omega)}$$

if $|\beta| \le k-1$, and so $u \in W^{k-1,p^*}(\Omega)$. Moreover, $\|u\|_{k-1,p^*} \le c\|u\|_{k,p}$. Similarly, we find $u \in W^{k-2,p^{**}}(\Omega)$, where

$$\frac{1}{p^{**}} = \frac{1}{p^*} - \frac{1}{n} = \frac{1}{p} - \frac{2}{n}.$$

Moreover, $\|u\|_{k-2,p^{**}} \le c\|u\|_{k-1,p^*}$. Continuing, we find after k steps that $u \in W^{0,q}(\Omega) = L^q(\Omega)$ for

$$\frac{1}{q} = \frac{1}{p} - \frac{k}{n}.$$

The stated estimate (2.9) follows from combining the relevant estimates at each stage of the above argument. In a similar manner the other estimates can be established. $\square$

We now consider the compactness of the imbeddings. Before we present the next result we recall some facts that will be needed. A subset S of a normed space is said to be **totally bounded** if for each $\varepsilon > 0$ there is a finite set of open balls of radius ε which cover S. Clearly, a totally bounded set is bounded, i.e., it is contained in a sufficiently large ball. It is not difficult to see that a relatively compact subset of a normed space is totally bounded, with the converse being true if the normed space is complete. Moreover, a totally bounded subset of a normed space is separable.

Theorem 2.33. (Rellich-Kondrachov) *Let $\Omega \subset \mathbb{R}^n$ be bounded and open. Then for $1 \le p < n$:*

(a) *The imbedding $W_0^{1,p}(\Omega) \subset L^q(\Omega)$ is compact for each $1 \le q < np/(n-p)$.*

(b) *Assuming $\partial\Omega \in C^1$, the imbedding $W^{1,p}(\Omega) \subset L^q(\Omega)$ is compact for each $1 \le q < np/(n-p)$.*

(c) *Assuming $\partial\Omega \in C^1$, $\gamma_0 : W^{1,p}(\Omega) \to L^p(\partial\Omega)$ is compact.*

If $p > n$, then

(d) *Assuming $\partial\Omega \in C^1$, the imbedding $W^{1,p}(\Omega) \subset C^\alpha(\bar{\Omega})$ is compact for each $0 \le \alpha < 1 - (n/p)$.*

Proof. We shall just give the proof for $p = q = 2$. The other cases are proved similarly. (a) Since $C_0^1(\Omega)$ is dense in $H_0^1(\Omega)$, it suffices to show that the imbedding $C_0^1(\Omega) \subset L^2(\Omega)$ is compact. Thus, let $S = \{u \in C_0^1(\Omega) : \|u\|_{1,2} \le 1\}$. We now show that S is totally bounded in $L^2(\Omega)$.

For $h > 0$, let $S_h = \{u_h : u \in S\}$, where u_h is the mollified function for u. We claim that S_h is totally bounded in $L^2(\Omega)$. Indeed, for $u \in S$, we have

$$|u_h(x)| \le \int_{B(0,h)} \omega_h(z)|u(x-z)|dz \le (\sup \omega_h)\|u\|_1 \le c_1(\sup \omega_h)\|u\|_{1,2}$$

and

$$|D_i u_h(x)| \le c_2 \sup |D_i \omega_h| \|u\|_{1,2}, \quad i = 1, \dots, n$$

so that S_h is a bounded and equicontinuous subset of $C(\bar{\Omega})$. Thus by the Ascoli Theorem, S_h is relatively compact (and thus totally bounded) in $C(\bar{\Omega})$ and consequently also in $L^2(\Omega)$.

Now, by earlier estimates, we easily obtain

$$\|u_h - u\|_2^2 \le \int_{B(0,h)} \omega_h(z) \left(\int_\Omega |u(x-z) - u(x)|^2 dx \right) dz$$

and

$$\begin{aligned}
\int_\Omega |u(x-z) - u(x)|^2 dx &= \int_\Omega \left| \int_0^1 \frac{du(x-tz)}{dt} dt \right|^2 dx \\
&= \int_\Omega \left| \int_0^1 (-\nabla u(x-tz) \cdot z) dt \right|^2 dx \\
&\le \int_\Omega |z|^2 \left(\int_0^1 |\nabla u(x-tz)|^2 dt \right) dx \le |z|^2 \|u\|_{1,2}^2.
\end{aligned}$$

Consequently, $\|u_h - u\|_2 \le h$. Since we have shown above that S_h is totally bounded in $L^2(\Omega)$ for all $h > 0$, it follows that S is also totally bounded in $L^2(\Omega)$ and hence relatively compact.

(b) Suppose now that S is a bounded set in $H^1(\Omega)$. Each $u \in S$ has an extension $U \in H_0^1(\Omega')$ where $\Omega \subset\subset \Omega'$. Denote by S' the set of all such extensions of the functions $u \in S$. Since $\|U\|_{H^1(\Omega')} \le c\|u\|_{1,2}$, the set S' is bounded in $H_0^1(\Omega')$. By (a) S' is relatively compact in $L^2(\Omega')$ and therefore S is relatively compact in $L^2(\Omega)$.

(c) Let S be a bounded set in $H^1(\Omega)$. For any $u(x) \in C^1(\bar{\Omega})$, the inequality (2.7) with $p = 2$ yields

$$\|u\|_{L^2(\partial\Omega)}^2 \le \frac{c_1}{\beta}\|u\|_2^2 + c_2\beta\|u\|_{1,2}^2 \tag{2.24}$$

where the constants c_1, c_2 do not depend on u or β. By completion, this inequality is valid for any $u \in H^1(\Omega)$. By (b), any infinite sequence of elements of the set S has a subsequence $\{u_n\}$ which is Cauchy in $L^2(\Omega)$: given $\varepsilon > 0$, an N can be found such that for all $m, n \geq N$, $\|u_m - u_n\|_2 < \varepsilon$. Now we choose $\beta = \varepsilon$. Applying the inequality (2.24) to $u_m - u_n$, it follows that the sequence of traces $\{\gamma_0 u_n\}$ converges in $L^2(\partial\Omega)$.

(d) By Morrey's inequality, the imbedding is continuous if $\alpha = 1 - (n/p)$. Now use the fact that C^β is compact in C^α if $\alpha < \beta$. □

Remarks. (a) When $p = n$, we can easily show that the imbedding in (a) is compact for all $1 \leq q < \infty$. Hence, it follows that the imbedding $W_0^{1,p}(\Omega) \subset L^p(\Omega)$ is compact for all $p \geq 1$.

(b) The boundedness of Ω is essential in the above theorem. For example, let $I = (0, 1)$ and $I_j = (j, j+1)$. Let $f \in C_0^1(I)$ and define f_j to be the same function defined on I_j by translation. We can normalize f so that $\|f\|_{W^{1,p}(I)} = 1$. The same is then true for each f_j and thus $\{f_j\}$ is a bounded sequence in $W^{1,p}(\mathbb{R})$. Clearly $f \in L^q(\mathbb{R})$ for every $1 \leq q \leq \infty$. Further, if

$$\|f\|_{L^q(\mathbb{R})} = \|f\|_{L^q(I)} = a > 0$$

then for any $j \neq k$ we have

$$\|f_j - f_k\|_{L^q(\mathbb{R})}^q = \int_j^{j+1} |f_j|^q + \int_k^{k+1} |f_k|^q = 2a^q$$

and so f_i cannot have a convergent subsequence in $L^q(\mathbb{R})$. Thus none of the imbeddings $W^{1,p}(\mathbb{R}) \subset L^q(\mathbb{R})$ can be compact. This example generalizes to n dimensional space and to open sets like a half-space.

2.4. Additional Properties

2.4.1. Equivalent Norms. Two norms $\|\cdot\|$ and $|\cdot|$ on a vector space X are **equivalent** if there exist constants $c_1, c_2 \in (0, \infty)$ such that

$$\|x\| \leq c_1|x| \leq c_2\|x\| \quad \text{for all } x \in X.$$

Note that the property of a set to be open, closed, compact, or complete in a normed space is not affected if the norm is replaced by an equivalent norm. A **seminorm** q on a vector space has all the properties of a norm except that $q(u) = 0$ need not imply $u = 0$.

Theorem 2.34. *Let $\partial\Omega \in C^1$ and let $1 \leq p < \infty$. Set*

$$\|u\| = \left(\int_\Omega \sum_{i=1}^n |D_i u|^p dx + (q(u))^p \right)^{1/p}$$

where $q : W^{1,p}(\Omega) \to \mathbb{R}$ is a seminorm with the following two properties:

(i) *There is a positive constant d such that for all $u \in W^{1,p}(\Omega)$*

$$q(u) \leq d\|u\|_{1,p}.$$

(ii) *If $u =$ constant, then $q(u) = 0$ implies $u = 0$.*

Then $\|\cdot\|$ is an equivalent norm on $W^{1,p}(\Omega)$.

Proof. First of all, it is easy to check that $\|\cdot\|$ defines a norm. Now by (i), it suffices to prove that there is a positive constant c such that

$$\|u\|_{1,p} \leq c\|u\| \quad \text{for all} \quad u \in W^{1,p}(\Omega). \tag{2.25}$$

Suppose (2.25) is false. Then there exist $u_n \in W^{1,p}(\Omega)$ such that $\|u_n\|_{1,p} > n\|u_n\|$. Set $v_n = u_n/\|u_n\|_{1,p}$. Then

$$\|u_n\|_{1,p} = 1 \quad \text{and} \quad 1 > n\|u_n\|. \tag{2.26}$$

According to Theorem 2.33, there is a subsequence, call it again $\{u_n\}$, which converges to u in $L^p(\Omega)$. From (2.26) we have $\|u_n\| \to 0$ and therefore $\nabla u_n \to 0$ in $L^p(\Omega)$ and $q(u_n) \to 0$. This implies $\nabla u = 0$ and hence $u = \text{const}$ and $u_n \to u$ in $W^{1,p}(\Omega)$. Hence, $\|u\|_{1,p} = 1$ and by triangle inequality, $q(u) = 0$. Since u is constant, this implies $u = 0$, which contradicts $\|u\|_{1,p} = 1$. □

EXAMPLE 2.35. Let $\partial\Omega \in C^1$. Assume $a(x) \in C(\overline{\Omega}), \sigma(x) \in C(\partial\Omega)$ with $a \geq 0$ ($\not\equiv 0$), $\sigma \geq 0$ ($\not\equiv 0$). Then the following norms are equivalent to $\|\cdot\|_{1,p}$ on $W^{1,p}(\Omega)$:

$$\|u\| = \left(\int_\Omega \sum_{i=1}^n |D_i u|^p dx + \left|\int_\Omega u dx\right|^p\right)^{1/p} \quad \text{with } q(u) = \left|\textstyle\int_\Omega u dx\right|. \tag{2.27}$$

$$\|u\| = \left(\int_\Omega \sum_{i=1}^n |D_i u|^p dx + \left|\int_{\partial\Omega} \gamma_0 u dS\right|^p\right)^{1/p} \quad \text{with } q(u) = \left|\textstyle\int_{\partial\Omega} \gamma_0 u dS\right|. \tag{2.28}$$

$$\|u\| = \left(\int_\Omega \sum_{i=1}^n |D_i u|^p dx + \int_{\partial\Omega} \sigma|\gamma_0 u|^p dS\right)^{1/p} \quad \text{with } q(u) = \left(\textstyle\int \sigma|\gamma_0 u|^p dS\right)^{1/p}. \tag{2.29}$$

$$\|u\| = \left(\int_\Omega \sum_{i=1}^n |D_i u|^p dx + \int_\Omega a|u|^p dx\right)^{1/p} \quad \text{with } q(u) = \left(\textstyle\int_\Omega a|u|^p dx\right)^{1/p}. \tag{2.30}$$

Clearly, for all these q's, condition (ii) of Theorem 2.34 holds. To verify condition (i) of Theorem 2.34, one uses the trace theorem in (2.28) and (2.29). Also, the $q(u)$ in (2.27) is bounded by $q(u)$ in (2.30). Condition (i) for $q(u)$ in (2.30) follows from the imbedding $W^{1,p}(\Omega) \to L^p(\Omega)$.

EXAMPLE 2.36. (**Poincaré's inequalities**.) Note that $\|u\|_{L^p(\Omega)} \leq c\|u\|$ if $\|u\|$ is any equivalent norm of $W^{1,p}(\Omega)$. Using such an inequality with $\|u\|$ defined by (2.27) on function $u - (u)_\Omega$, where $(u)_\Omega = \frac{1}{|\Omega|}\int_\Omega u dx$ is the mean of u over Ω, we have

$$\int_\Omega |u(x) - (u)_\Omega|^p dx \leq c \int_\Omega \sum_{i=1}^n |D_i u|^p dx, \quad \forall\, u \in W^{1,p}(\Omega), \tag{2.31}$$

where the constant $c > 0$ is independent of u. This inequality is often referred to as a **Poincaré's inequality**. Also if $u \in W_0^{1,p}(\Omega)$, by (2.28), it follows that

$$\int_\Omega |u(x)|^p dx \leq c \int_\Omega \sum_{i=1}^n |D_i u|^p dx, \quad \forall\, u \in W_0^{1,p}(\Omega), \tag{2.32}$$

where the constant $c > 0$ is independent of u. This inequality is also referred to as a **Poincaré's inequality**. In particular, (2.32) implies that quantity

$$\|u\|_{1,p,0} = \left(\int_\Omega \sum_{i=1}^n |D_i u|^p dx\right)^{1/p}$$

defines an equivalent norm on $W_0^{1,p}(\Omega)$.

2.4.2. Difference Quotients. Assume $u \in L^1_{loc}(\Omega)$ and $\Omega' \subset\subset \Omega$. Let $e_1, \cdots, e_n$ be the standard basis vectors of $\mathbb{R}^n$. For all $h \in \mathbf{R}$ with $0 < |h| < \operatorname{dist}(\Omega', \partial\Omega)$ and $i = 1, 2, \cdots, n$, we define the ***i*-th difference quotient** of u of size h on Ω' by

$$\delta_i^h u(x) = \frac{u(x + he_i) - u(x)}{h}, \quad x \in \Omega'.$$

Denote $\delta^h u = (\delta_1^h u, \delta_2^h u, \cdots, \delta_n^h u)$.

Theorem 2.37. (Difference Quotients and Weak Derivatives)

(i) *Suppose* $1 \le p < \infty$ *and* $u \in W^{1,p}(\Omega)$. *Then for each* $\Omega' \subset\subset \Omega$ *there exists a constant* C *such that*

$$\|\delta^h u\|_{L^p(\Omega')} \le C \|\nabla u\|_{L^p(\Omega)} \quad \forall\, 0 < |h| < \operatorname{dist}(\Omega', \partial\Omega).$$

(ii) *Let* $1 < p < \infty$, $u \in L^p_{loc}(\Omega)$ *and* $\Omega' \subset\subset \Omega$. *Assume* $i \in \{1, 2, \cdots, n\}$. *Suppose that there exist constant* C *and sequence* $h_k \to 0$ *such that*

$$\|\delta_i^{h_k} u\|_{L^p(\Omega')} \le C \quad \forall\, k = 1, 2, \cdots. \tag{2.33}$$

Then weak derivative u_{x_i} *exists in* $L^p(\Omega')$ *and* $\|u_{x_i}\|_{L^p(\Omega')} \le C$.

Proof. 1. We first assume u is smooth. Let $x \in \Omega'$ and $0 < |h| < \operatorname{dist}(\Omega', \partial\Omega)$. Then

$$u(x + he_i) - u(x) = \int_0^1 u_{x_i}(x + the_i) h \, dt$$

and hence

$$|\delta_i^h u(x)| = \left|\frac{u(x + he_i) - u(x)}{h}\right| \le \int_0^1 |\nabla u(x + the_i)| dt.$$

Raising to the p-th power, using Hölder's inequality and integrating over $x \in \Omega'$ will yield

$$\int_{\Omega'} |\delta_i^h u(x)|^p \, dx \le \int_{\Omega'} \int_0^1 |\nabla u(x + the_i)|^p \, dt dx$$

$$\le \int_0^1 \int_{\Omega'} |\nabla u(x + the_i)|^p \, dx dt \le \int_0^1 \int_{\Omega} |\nabla u(x)|^p \, dx dt = \int_{\Omega} |\nabla u(x)|^p \, dx.$$

Thus

$$\int_{\Omega'} |\delta^h u|^p \, dx \le C \int_{\Omega} |\nabla u(x)|^p \, dx$$

with C actually depending only on n, p. This estimate holds for all smooth u and hence is also valid by approximation for arbitrary $u \in W^{1,p}(\Omega)$.

2. By (2.33), the sequence $\{\delta_i^{h_k} u\}$ is bounded in $L^p(\Omega')$. Since $p > 1$, there exists a subsequence $\{\delta_i^{h_k} u\}$, still denoted by k, that converges weakly to $v \in L^p(\Omega')$. Clearly $\|v\|_{L^p(\Omega')} \le C$. We show that v is the weak derivative u_{x_i} on Ω'. Let $\phi \in C_0^1(\Omega')$. For each $0 < |h| < \frac{1}{2} \operatorname{dist}(\Omega', \partial\Omega)$, it is easy to check that

$$\int_{\Omega'} u(x) \delta_i^{-h} \phi(x) \, dx = -\int_{\Omega'} \phi(x) \delta_i^h u(x) \, dx.$$

(This is the *integration by parts* formula for difference quotients.) Take $h = h_k \to 0$ in this formula, by the weak convergence of $\delta_i^{h_k} u$ to v and the fact $\delta_i^{-h} \phi(x) \to \phi_{x_i}(x)$ uniformly on Ω' as $h \to 0$, and one proves that

$$\int_{\Omega'} u(x) \phi_{x_i}(x) \, dx = -\int_{\Omega'} \phi(x) v(x) \, dx \quad \forall\, \phi \in C_0^1(\Omega').$$

By definition of weak derivatives, $v = u_{x_i}$. ☐

2.4.3. Fourier Transform Methods. In defining Sobolev spaces H^k on the whole $\mathbb{R}^n$, it is often useful to use the **Fourier transform**.

For a function $u \in L^1(\mathbb{R}^n)$, we define the **Fourier transform** of u by

$$\hat{u}(y) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot y} u(x)\, dx, \quad \forall\, y \in \mathbb{R}^n,$$

and the **inverse Fourier tranform** by

$$\check{u}(y) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ix\cdot y} u(x)\, dx, \quad \forall\, y \in \mathbb{R}^n.$$

Theorem 2.38. (Plancherel's Theorem) *Assume $u \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Then $\hat{u}$, $\check{u} \in L^2(\mathbb{R}^n)$ and*

$$\|\hat{u}\|_{L^2(\mathbb{R}^n)} = \|\check{u}\|_{L^2(\mathbb{R}^n)} = \|u\|_{L^2(\mathbb{R}^n)}.$$

Since $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$, we can use this result to extend the Fourier transforms on to $L^2(\mathbb{R}^n)$. We still use the same notations for them. Then we have

Theorem 2.39. (Property of Fourier Tranforms) *Assume $u, v \in L^2(\mathbb{R}^n)$. Then*

(i) $\int_{\mathbb{R}^n} u\bar{v}\, dx = \int_{\mathbb{R}^n} \hat{u}\bar{\hat{v}}\, dy$,

(ii) $\widehat{D^\alpha u}(y) = (iy)^\alpha \hat{u}(y)$ *for each multiindex α such that $D^\alpha u \in L^2(\mathbb{R}^n)$,*

(iii) $u = \check{\hat{u}}$.

Next we use the Fourier transform to characterize the spaces $H^k(\mathbb{R}^n)$.

Theorem 2.40. *Let k be a nonnegative integer. Then, a function $u \in L^2(\mathbb{R}^n)$ belongs to $H^k(\mathbb{R}^n)$ if and only if*

$$(1 + |y|^k)\hat{u}(y) \in L^2(\mathbb{R}^n).$$

In addition, there exists a constant C such that

$$C^{-1}\, \|u\|_{H^k(\mathbb{R}^n)} \le \|(1 + |y|^k)\, \hat{u}\|_{L^2(\mathbb{R}^n)} \le C\, \|u\|_{H^k(\mathbb{R}^n)}$$

for all $u \in H^k(\mathbb{R}^n)$.

Using the Fourier transform, we can also define *fractional* Sobolev spaces $H^s(\mathbb{R}^n)$ for any $0 < s < \infty$ as follows

$$H^s(\mathbb{R}^n) = \{u \in L^2(\mathbb{R}^n) \,|\, (1 + |y|^s)\, \hat{u} \in L^2(\mathbb{R}^n)\},$$

and define the norm by

$$\|u\|_{H^s(\mathbb{R}^n)} = \|(1 + |y|^s)\, \hat{u}\|_{L^2(\mathbb{R}^n)}.$$

From this we easily get the estimate

$$\begin{aligned} \|u\|_{L^\infty(\mathbb{R}^n)} &\le \|\hat{u}\|_{L^1(\mathbb{R}^n)} \\ &= \|(1 + |y|^s)\hat{u}\, (1 + |y|^s)^{-1}\|_{L^1(\mathbb{R}^n)} \\ &\le \|(1 + |y|^s)\hat{u}\|_{L^2(\mathbb{R}^n)} \|(1 + |y|^s)^{-2}\|^2_{L^1(\mathbb{R}^n)} \\ &\le C\, \|u\|_{H^s(\mathbb{R}^n)}, \end{aligned}$$

where $C = \|(1 + |y|^s)^{-2}\|^2_{L^1(\mathbb{R}^n)} < \infty$ if and only if $s > \frac{n}{2}$. Therefore we have an easy imbedding, which is known valid for integers s by the previous Sobolev imbedding theorem,

$$H^s(\mathbb{R}^n) \subset L^\infty(\mathbb{R}^n) \quad \text{if } s > \tfrac{n}{2}.$$

Chapter 3

Second-Order Linear Elliptic PDEs in Divergence Form

3.1. Second-Order PDEs in Divergence Form

3.1.1. Single Equation Case. Henceforth, $\Omega \subset \mathbb{R}^n$ denotes a bounded domain with boundary $\partial\Omega \in C^1$. Consider the (Dirichlet) boundary value problem (BVP)

$$Lu = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega. \tag{3.1}$$

Here f is a given function in $L^2(\Omega)$ (or more generally, an element in the dual space of $H_0^1(\Omega)$) and L is a formal **second-order differential operator in divergence form** given by

$$Lu \equiv -\sum_{i,j=1}^{n} D_i\left(a_{ij}(x)D_j u\right) + \sum_{i=1}^{n} b_i(x)D_i u + c(x)u$$

with real coefficients $a_{ij}(x)$, $b_i(x)$ and $c(x)$ in $L^\infty(\Omega)$. Moreover, L is assumed to be **uniformly elliptic** in Ω, i.e., there exists a number $\theta > 0$ such that for every $x \in \Omega$ and every real vector $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq \theta \sum_{i=1}^{n} |\xi_i|^2. \tag{3.2}$$

A function $u \in H_0^1(\Omega)$ is called a **weak solution** of (3.1) if the following holds

$$B_1(u,v) \equiv \int_\Omega \Big[\sum_{i,j=1}^{n} a_{ij}D_j u D_i v + (\sum_{i=1}^{n} b_i D_i u + cu)v\Big]\,dx = \int_\Omega fv\,dx, \quad \forall\ v \in H_0^1(\Omega). \tag{3.3}$$

Other problems can also be formulated in the weak sense as above in different Hilbert spaces.

EXAMPLE 3.1. Consider the following weak formulation: Given $f \in L^2(\Omega)$, find $u \in H^1(\Omega)$ satisfying

$$\int_\Omega \nabla u \cdot \nabla v dx = \int_\Omega f v dx \quad \text{for all} \quad v \in H^1(\Omega).$$

Find the boundary value problem solved by u. What is the necessary condition for the existence of such a u?

3.1.2. General System of PDEs in Divergence Form. For N unknown functions, $u^1, \cdots, u^N$, we write $u = (u^1, \cdots, u^N)$ as a vector field. We say $u \in X(\Omega; \mathbb{R}^N)$ if each $u^k \in X(\Omega)$, where X is a symbol of any function spaces we have learned. For example, if $u \in W^{1,p}(\Omega; \mathbb{R}^N)$ then we use Du to denote the $N \times n$ Jacobi matrix $(\partial u^k / \partial x_i)_{1 \le k \le N, 1 \le i \le n}$.

The (Dirichlet) BVP for general system of nonlinear partial differential equations in divergence form can be written as follows:

$$-\operatorname{div} A(x, u, Du) + b(x, u, Du) = F \ \text{ in } \ \Omega, \quad u = 0 \ \text{ on } \ \partial\Omega, \tag{3.4}$$

where $A(x, s, \xi) = (A_i^k(x, s, \xi))$, $1 \le i \le n$, $1 \le k \le N$, and $b(x, s, \xi) = (b^k(x, s, \xi))$, $1 \le k \le N$, are given functions of $(x, s, \xi) \in \Omega \times \mathbb{R}^N \times \mathbb{M}^{N \times n}$ satisfying some *structural conditions*, and $F = (f^k)$, $1 \le k \le N$, with each f^k being a given functional in the dual space of $W_0^{1,p}(\Omega)$.

The suitable structural conditions will generally assure that both $|A(x, u, Du)|$ and $|b(x, u, Du)|$ belong to $L^{p'}(\Omega)$ for all $u \in W^{1,p}(\Omega; \mathbb{R}^N)$, where $p' = \frac{p}{p-1}$.

A function $u \in W_0^{1,p}(\Omega; \mathbb{R}^N)$ is called a **weak solution** to problem (3.4) if

$$\int_\Omega \left[\sum_{i=1}^n A_i^k(x, u, Du) D_i \varphi + b^k(x, u, Du) \varphi \right] dx = \langle f^k, \varphi \rangle \tag{3.5}$$

for all $\varphi \in W_0^{1,p}(\Omega)$ and each $k = 1, 2, \cdots, N$.

System (3.4) is said to be **linear** if both $A(x, s, \xi)$ and $b(x, s, \xi)$ are linear in the variables (s, ξ), so that

$$\begin{aligned} A_i^k(x, u, Du) &= \sum_{1 \le l \le N, 1 \le j \le n} a_{ij}^{kl}(x)\, D_j u^l + \sum_{l=1}^N d_i^{kl}(x)\, u^l, \\ b^k(x, u, Du) &= \sum_{1 \le j \le n, 1 \le l \le N} b_j^{kl}(x)\, D_j u^l + \sum_{l=1}^N c^{kl}(x)\, u^l. \end{aligned} \tag{3.6}$$

In the linear case, as in the single equation case, we work in the Hilbert space $H_0^1(\Omega; \mathbb{R}^N)$, which has the inner product defined by

$$(u, v) \equiv \sum_{1 \le i \le n, 1 \le k \le N} \int_\Omega D_i u^k \, D_i v^k \, dx.$$

The pairing between $H_0^1(\Omega; \mathbb{R}^N)$ and its dual is given by

$$\langle F, u \rangle = \sum_{k=1}^N \langle f^k, u^k \rangle \quad \text{if } F = (f^k) \text{ and } u = (u^k).$$

Then a weak solution of linear system (3.4) is a function $u \in H_0^1(\Omega; \mathbb{R}^N)$ such that

$$B_2(u, v) \equiv \int_\Omega \left(a_{ij}^{kl} D_j u^l D_i v^k + d_i^{kl} u^l D_i v^k + b_j^{kl} D_j u^l v^k + c^{kl} u^l v^k \right) dx = \langle F, v \rangle \tag{3.7}$$

holds for all $v \in H_0^1(\Omega; \mathbb{R}^N)$. Here the conventional summation notation is used.

3.1.3. Ellipticity Conditions for Systems. There are several **ellipticity conditions** for the system (3.5) in terms of the leading coefficients $A(x, s, \xi)$. Assume A is smooth on ξ and define

$$A_{ij}^{kl}(x, s, \xi) = \frac{\partial A_i^k(x, s, \xi)}{\partial \xi_j^l}, \quad \xi = (\xi_j^l).$$

The system (3.5) is said to satisfy the (uniform, strict) **Legendre ellipticity condition** if there exists a $\nu > 0$ such that, for all (x, s, ξ), it holds

$$\sum_{i,j=1}^{n} \sum_{k,l=1}^{N} A_{ij}^{kl}(x, s, \xi)\, \eta_i^k\, \eta_j^l \geq \nu\, |\eta|^2 \quad \text{for all } N \times n \text{ matrix } \eta = (\eta_i^k). \tag{3.8}$$

A weaker condition, obtained by setting $\eta = q \otimes p = (q^k p_i)$ with $p \in \mathbb{R}^n$, $q \in \mathbb{R}^N$, is the following so-called (strong) **Legendre-Hadamard condition**:

$$\sum_{i,j=1}^{n} \sum_{k,l=1}^{N} A_{ij}^{kl}(x, s, \xi)\, q^k q^l p_i p_j \geq \nu\, |p|^2\, |q|^2 \quad \forall\, p \in \mathbb{R}^n,\ q \in \mathbb{R}^N. \tag{3.9}$$

Note that for systems with *linear leading terms* A given by (3.6), the Legendre condition and Legendre-Hadamard condition become, respectively,

$$\sum_{i,j=1}^{n} \sum_{k,l=1}^{N} a_{ij}^{kl}(x)\, \eta_i^k\, \eta_j^l \geq \nu\, |\eta|^2 \quad \forall\, \eta = (\eta_i^k)_{1 \leq k \leq N,\, 1 \leq i \leq n}; \tag{3.10}$$

$$\sum_{i,j=1}^{n} \sum_{k,l=1}^{N} a_{ij}^{kl}(x)\, q^k q^l p_i p_j \geq \nu\, |p|^2\, |q|^2 \quad \forall\, p \in \mathbb{R}^n,\ q \in \mathbb{R}^N. \tag{3.11}$$

EXAMPLE 3.2. The Legendre-Hadamard condition does not imply the Legendre ellipticity condition. For example, let $n = N = 2$ and $\varepsilon > 0$; define *constants* a_{ij}^{kl} by

$$\sum_{i,j,k,l=1}^{2} a_{ij}^{kl}\, \xi_i^k\, \xi_j^l \equiv \det \xi + \epsilon\, |\xi|^2.$$

Since

$$\sum_{i,j,k,=1}^{2} a_{ij}^{kl} p_i p_j\, q^k q^l = \det(q \otimes p) + \epsilon\, |q \otimes p|^2 = \epsilon\, |p|^2 |q|^2,$$

the Legendre-Hadamard condition holds for all $\epsilon > 0$. However, show that the Legendre condition holds for this system if and only if $\epsilon > 1/2$.

EXAMPLE 3.3. Let $u = (v, w)$ and $(x_1, x_2) = (x, y)$. Then the system of differential equations defined by a_{ij}^{kl} given above is

$$\epsilon \Delta v + w_{xy} = 0, \quad \epsilon \Delta w - v_{xy} = 0.$$

This system reduces to two fourth-order equations for v, w (where $\Delta f = f_{xx} + f_{yy}$):

$$\epsilon^2 \Delta^2 v - v_{xxyy} = 0, \quad \epsilon^2 \Delta^2 w + w_{xxyy} = 0.$$

Show that both equations are **elliptic** (in the sense of linear operators) if and only if $\epsilon > 1/2$.

3.2. The Lax-Milgram Theorem

Let H denote a real Hilbert space with inner product $(\cdot,\cdot)$ and norm $\|\cdot\|$. A function $B : H \times H \to \mathbb{R}$ is called a **bilinear form** if

$$B(\alpha u + \beta v, w) = \alpha B(u,w) + \beta B(v,w)$$

$$B(w, \alpha u + \beta v) = \alpha B(w,u) + \beta B(w,v)$$

for all $u, v, w \in H$ and all $\alpha, \beta \in \mathbb{R}$.

Our first existence result is frequently referred to as the **Lax-Milgram Theorem**.

Theorem 3.4. (Lax-Milgram Theorem) *Let $B\colon H \to H$ be a bilinear form. Assume*

(i) *B is* **bounded***; i.e., $|B(u,v)| \le \alpha \|u\|\|v\| \ \ \forall\, u, v \in H$, for some $\alpha > 0$; and*

(ii) *B is* **strongly positive** *(also called* **coercive***); i.e., $B(u,u) \ge \beta \|u\|^2 \ \ \forall\, u \in H$, for some $\beta > 0$.*

Let $f \in H^*$. Then there exists a unique element $u \in H$ such that

$$B(u,v) = \langle f, v\rangle, \quad \forall\, v \in H. \tag{3.12}$$

Moreover, the solution u satisfies $\|u\| \le \frac{1}{\beta}\,\|f\|$.

Proof. For each fixed $u \in H$, the functional $v \mapsto B(u,v)$ is in H^*, and hence by the Riesz Representation Theorem, there exists a unique element $w = Au \in H$ such that

$$B(u,v) = (w,v) = (Au,v) \quad \forall\, v \in H.$$

It is easy to see that $A : H \to H$ is linear. From (i), $\|Au\|^2 = B(u, Au) \le \alpha \|u\|\|Au\|$, and hence $\|Au\| \le \alpha\|u\|$ for all $u \in H$; that is, A is bounded. Furthermore, by (ii), $\beta\|u\|^2 \le B(u,u) = (Au,u) \le \|Au\|\|u\|$ and hence $\|Au\| \ge \beta\|u\|$ for all $u \in H$. By the Riesz Representation Theorem again, we have a unique $w_0 \in H$ such that $\langle f, v\rangle = (w_0, v)$ for all $v \in H$ and $\|f\| = \|w_0\|$. We will show that the equation $Au = w_0$ has a unique solution $u \in H$. The uniqueness of u follows easily from the property $\|Au - Av\| \ge \beta\|u - v\|$ for all $u, v \in H$. There are many different proofs for existence; here we use the Contraction Mapping Theorem.

Note that the solution u to equation $Au = w_0$ is equivalent to the fixed-point of the map $T\colon H \to H$ defined by $T(v) = v - tAv + tw_0$ $(v \in H)$ for any fixed $t > 0$. We show that for $t > 0$ small enough T is a contraction. Note that for all $v, w \in H$ we have $\|T(v) - T(w)\| = \|(I - tA)(v - w)\|$. We compute that for all $u \in H$

$$\begin{aligned} \|(I - tA)u\|^2 &= \|u\|^2 + t^2\|Au\|^2 - 2t(Au,u) \\ &\le \|u\|^2(1 + t^2\alpha^2 - 2\beta t) \\ &\le \gamma\|u\|^2, \end{aligned}$$

for some $0 < \gamma < 1$ if we choose t such that $0 < t < \frac{2\beta}{\alpha^2}$. Therefore, map $T\colon H \to H$ is a contraction (with constant $\sqrt{\gamma}$) on H and thus has a fixed point. This fixed point u solves $Au = w_0$ and is thus the (unique) solution of (3.12). Moreover, we have $\|f\| = \|w_0\| = \|Au\| \ge \beta\|u\|$ and hence $\|u\| \le \frac{1}{\beta}\,\|f\|$. The proof is complete. $\square$

3.3. Gårding's Estimates and Existence Theory

In the following, we assume all coefficients involved in the bilinear forms B_1 and B_2 defined above are in $L^\infty(\Omega)$. Then one easily shows the boundedness:

$$|B_j(u,v)| \le \alpha\|u\|\|v\|$$

for all u, v in the respective Hilbert spaces $H = H_0^1(\Omega)$ or $H = H_0^1(\Omega;\mathbb{R}^N)$ for $j = 1, 2$. Here and below, the norm $\|u\|$ on H is the equivalent norm $\|Du\|_{L^2(\Omega)}$ induced by the inner product $(u,v)_H = \int_\Omega (Du, Dv)\,dx$ in both scaler or vectorial cases.

The strong positivity (also called **coercivity**) for B_1 or B_2 is not always guaranteed and involves estimating the quadratic form $B_j(u,u)$, which usually involves the so-called **Gårding's estimates**. We will derive these estimates below and then state the corresponding existence theorem.

3.3.1. Estimate for $B_1(u,u)$.

Theorem 3.5. *Assume the ellipticity condition* (3.2) *holds. Then, there are constants* $\beta > 0$ *and* $\gamma \in \mathbb{R}$ *such that*

$$B_1(u,u) \ge \beta\|u\|^2 - \gamma\|u\|^2_{L^2(\Omega)} \tag{3.13}$$

for all $u \in H = H_0^1(\Omega)$.

Proof. Note that, by the ellipticity,

$$B_1(u,u) - \int_\Omega (\sum_{i=1}^n b_i D_i u + cu)u\,dx \ge \theta \int_\Omega \sum_{i=1}^n |D_i u|^2 dx.$$

Let $m = \max\{\|b_i\|_{L^\infty(\Omega)} \,|\, 1 \le i \le n\}$. Then

$$\begin{aligned} |(b_i D_i u, u)_2| &\le m\|D_i u\|_2\|u\|_2 \\ &\le (m/2)(\varepsilon\|D_i u\|_2^2 + (1/\varepsilon)\|u\|_2^2), \end{aligned}$$

where in the last step we have used **Cauchy's inequality with** ϵ:

$$|\alpha\beta| \le (\varepsilon/2)\alpha^2 + (1/2\varepsilon)\beta^2 \quad \forall\ \alpha,\, \beta \in \mathbb{R}.$$

Also, $(cu,u)_2 \ge k_0\|u\|^2_{L^2(\Omega)}$, where $k_0 = \operatorname{ess\,inf}_{x\in\Omega} c(x)$. Combining these estimates we find

$$B_1(u,u) \ge (\theta - m\varepsilon/2)\|Du\|^2_{L^2(\Omega)} - (mn/2\varepsilon - k_0)\|u\|^2_{L^2(\Omega)}.$$

By choosing $\varepsilon > 0$ so that $\beta = \theta - m\varepsilon/2 > 0$ we arrive at the desired estimate, with $\gamma = mn/2\varepsilon - k_0$. □

Theorem 3.6. *Let Lu be defined as above. There is a number $\gamma \in \mathbb{R}$ such that for each $\lambda \ge \gamma$ and for each function $f \in L^2(\Omega)$, the boundary value problem*

$$Lu + \lambda u = f(x) \ \text{ in } \Omega, \quad u = 0 \ \text{ on } \partial\Omega$$

has a unique weak solution $u \in H = H_0^1(\Omega)$ which satisfies

$$\|u\|_H \le c\|f\|_2$$

where the positive constant c is independent of f.

Proof. Let γ be defined from (3.13) and let $\lambda \geq \gamma$. Define the bilinear form

$$B^\lambda(u,v) \equiv B_1(u,v) + \lambda(u,v)_2 \quad \text{for all} \quad u,v \in H$$

which corresponds to the operator $Lu + \lambda u$. Then $B^\lambda(u,v)$ satisfies the hypotheses of the Lax-Milgram Theorem. Hence equation $Lu + \lambda u = f$ has a unique weak solution in H; moreover, $\|u\| \leq \frac{1}{\beta}\|f\|_2$, where $\beta > 0$ is the constant from (3.13). □

EXAMPLE 3.7. Consider the **Neumann boundary** value problem

$$-\Delta u(x) = f(x) \ \text{ in } \Omega, \quad \frac{\partial u}{\partial \nu} = 0 \ \text{ on } \partial\Omega. \tag{3.14}$$

A function $u \in H^1(\Omega)$ is said to be a **weak solution** to (3.14 if

$$\int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega f v \, dx, \quad \forall \ v \in H^1(\Omega). \tag{3.15}$$

Obviously, taking $v \equiv 1 \in H^1(\Omega)$, a necessary condition to have a weak solution is $\int_\Omega f(x)\, dx = 0$. We show this is also a sufficient condition for existence of the weak solutions. Note that, if u is a weak solution, then $u + c$, for all constants c, is also a weak solution. Therefore, to fix the constants, we consider the vector space

$$H = \Big\{ u \in H^1(\Omega) \,\Big|\, \int_\Omega u(x)\, dx = 0 \Big\}$$

equipped with inner product

$$(u,v)_H = \int_\Omega \nabla u \cdot \nabla v \, dx.$$

By the theorem on equivalent norms, it follows that H with this inner product, is indeed a Hilbert space, and $(f,u)_{L^2(\Omega)}$ is a bounded linear functional on H:

$$|(f,u)_{L^2(\Omega)}| \leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \leq \|f\|_{L^2(\Omega)} \|v\|_H.$$

Hence the Riesz Representation Theorem implies that there exists a unique $u \in H$ such that

$$(u,w)_H = (f,w)_{L^2(\Omega)}, \quad \forall \ w \in H. \tag{3.16}$$

It follows that u is a weak solution to the Neumann problem since for any $v \in H^1(\Omega)$ we take $w = v - c \in H$, where $c = \frac{1}{|\Omega|}\int_\Omega v dx$, in (3.16) and obtain (3.15) using $\int_\Omega f dx = 0$.

EXAMPLE 3.8. Denote by H^1_c the space

$$H^1_c = \{u \in H^1(\Omega) : \gamma_0 u = \text{const}\}.$$

Note that the constant may be different for different u's.

(a) Prove that H^1_c is complete.

(b) Prove the existence and uniqueness of a function $u \in H^1_c$ satisfying

$$\int_\Omega (\nabla u \cdot \nabla v + uv) dx = \int_\Omega f v dx \quad \text{for all} \quad v \in H^1_c$$

where $f \in C(\bar{\Omega})$.

(c) If $u \in C^2(\bar{\Omega})$ satisfies the equation in (b), find the underlying BVP.

EXAMPLE 3.9. Let us consider the nonhomogeneous Dirichlet boundary value problem

$$-\Delta u = f \text{ in } \Omega, \quad u|_{\partial\Omega} = \varphi \tag{3.17}$$

where $f \in L^2(\Omega)$ and φ is the trace of a function $w \in H^1(\Omega)$. Note that it is not sufficient to just require that $\varphi \in L^2(\partial\Omega)$ since the trace operator is not onto. If, for example, $\varphi \in C^1(\partial\Omega)$, then φ has a C^1 extension to $\bar{\Omega}$, which is the desired w.

The function $u \in H^1(\Omega)$ is called a weak solution of (3.17) if $u - w \in H_0^1(\Omega)$ and if

$$\int_\Omega \nabla u \cdot \nabla v dx = \int_\Omega f v dx \quad \text{for all} \quad v \in H_0^1(\Omega).$$

Let u be a weak solution of (3.17) and set $u = z + w$. Then $z \in H_0^1(\Omega)$ satisfies

$$\int_\Omega \nabla z \cdot \nabla v dx = \int_\Omega (fv - \nabla v \cdot \nabla w) dx \quad \text{for all} \quad v \in H_0^1(\Omega). \tag{3.18}$$

Since the right-hand side belongs to the dual space $H_0^1(\Omega)^*$, the Lax-Milgram theorem yields the existence of a unique $z \in H_0^1(\Omega)$ which satisfies (3.18). Hence (3.17) has a unique weak solution u.

EXAMPLE 3.10. Let $\Omega \subset \mathbb{R}^n$. Show that if $u, v \in H_0^2(\Omega)$, then

$$\int_\Omega \Delta u \Delta v dx = \sum_{i,j=1}^n \int_\Omega D_{ij} u D_{ij} v dx.$$

Hence, $\|\Delta u\|_2$ defines a norm on $H_0^2(\Omega)$ which is equivalent to the usual norm of $H_0^2(\Omega)$.

EXAMPLE 3.11. Now let us consider the boundary value (also called Dirichlet) problem for the fourth order biharmonic operator:

$$\Delta^2 u = f \text{ in } \Omega, \quad u|_{\partial\Omega} = \frac{\partial u}{\partial n}|_{\partial\Omega} = 0.$$

We take $H = H_0^2(\Omega)$. By the general trace theorem, $H = H_0^2(\Omega) = \{v \in H^2(\Omega) : \gamma_0 v = \gamma_1 v = 0\}$. Therefore, this space H is the right space for the boundary conditions.

Accordingly, for $f \in L^2(\Omega)$, a function $u \in H = H_0^2(\Omega)$ is a weak solution of the Dirichlet problem for the **biharmonic** operator provided

$$\int_\Omega \Delta u \Delta v dx = \int_\Omega f v dx \quad \forall\, v \in H.$$

Consider the bilinear form

$$B(u, v) = \int_\Omega \Delta u \Delta v dx.$$

Its boundedness follows from the Cauchy-Schwarz inequality

$$|B(u, v)| \le \|\Delta u\|_2 \|\Delta v\|_2 \le d\|u\|_{2,2}\|v\|_{2,2}.$$

Furthermore,

$$B(u, u) = \|\Delta u\|_2^2 \ge c\|u\|_{2,2}^2.$$

So, by the Lax-Milgram theorem (in fact, just the Riesz Representation Theorem), there exists a unique weak solution $u \in H$.

3.3.2. Estimate for $B_2(u,u)$. Let $H = H_0^1(\Omega;\mathbb{R}^N)$ and let $(u,v)_H$ and $\|u\|_H$ be the inner product and norm defined above on H.

We consider the system (3.4) with A and b given by (3.6). Let $B_2(u,v)$ be the bilinear form on $H\times H$ associated with this problem, defined by (3.7). We will derive the Gårding estimate for $B_2(u,u)$.

Define the bilinear form of the leading terms by

$$A(u,v) = \sum_{i,j=1}^{n}\sum_{k,l=1}^{N}\int_\Omega a_{ij}^{kl}(x)\, D_j u^l\, D_i v^k\, dx \quad (u,\, v\in H).$$

Lemma 3.12. *Assume that* **either** *coefficients a_{ij}^{kl} satisfy the* Legendre condition **or** a_{ij}^{kl} *are all constants and satisfy the* Legendre-Hadamard condition. *Then*

$$A(u,u) \ge \nu\, \|u\|_H^2, \quad \forall\, u\in H.$$

Proof. The conclusion follows easily if the coefficeints a_{ij}^{kl} satisfy the Legendre condition. We prove the second case when a_{ij}^{kl} are *constants* satisfying the Legendre-Hadamard condition

$$\sum_{i,j=1}^{n}\sum_{k,l=1}^{N} a_{ij}^{kl}\, q^k\, q^l\, p_i p_j \ge \nu\, |p|^2|q|^2, \quad \forall p\in\mathbb{R}^n,\ q\in\mathbb{R}^N.$$

It suffices to prove

$$A(u,u) = \sum_{i,j=1}^{n}\sum_{k,l=1}^{N}\int_\Omega a_{ij}^{kl}\, D_j u^l\, D_i u^k\, dx \ge \nu \int_\Omega |Du|^2\, dx$$

for all $u\in C_0^\infty(\Omega;\mathbb{R}^N)$. Given such a function u, we extend it onto $\mathbb{R}^n$ by zero outside Ω so that $u\in C_0^\infty(\mathbb{R}^n;\mathbb{R}^N)$. Define the **Fourier transform** of u by

$$\hat u(y) = (2\pi)^{-n/2}\int_{\mathbb{R}^n} e^{-i\, y\cdot x}\, u(x)\, dx; \quad y\in\mathbb{R}^n.$$

Then, for any $u,\, v\in C_0^\infty(\mathbb{R}^n;\mathbb{R}^N)$, we have

$$\int_{\mathbb{R}^n} u(x)\cdot v(x)\, dx = \int_{\mathbb{R}^n} \hat u(y)\cdot \overline{\hat v(y)}\, dy, \quad \widehat{D_j u^k}(y) = i\, y_j \widehat{u^k}(y);$$

the second identity can also be written as $\widehat{Du}(y) = i\,\hat u(y)\otimes y$. Now, using these identities, we have

$$\int_{\mathbb{R}^n} a_{ij}^{kl}\, D_i u^k(x)\, D_j u^l(x)\, dx = \int_{\mathbb{R}^n} a_{ij}^{kl}\, \widehat{D_i u^k}(y)\,\overline{\widehat{D_j u^l}(y)}\, dy$$

$$= \int_{\mathbb{R}^n} a_{ij}^{kl}\, y_i\, y_j\, \widehat{u^k}(y)\,\overline{\widehat{u^l}(y)}\, dy = \mathrm{Re}\left(\int_{\mathbb{R}^n} a_{ij}^{kl}\, y_i\, y_j\, \widehat{u^k}(y)\,\overline{\widehat{u^l}(y)}\, dy\right).$$

Write $\hat u(y) = \eta + i\xi$ with $\eta,\,\xi\in\mathbb{R}^N$. Then

$$\mathrm{Re}\left(\widehat{u^k}(y)\,\overline{\widehat{u^l}(y)}\right) = \eta^k\,\eta^l + \xi^k\,\xi^l.$$

Therefore, by the Legendre-Hadamard condition,

$$\mathrm{Re}\sum_{i,j=1}^{n}\sum_{k,l=1}^{N}\left(a_{ij}^{kl}\, y_i\, y_j\, \widehat{u^k}(y)\,\overline{\widehat{u^l}(y)}\right) \ge \nu\, |y|^2\,(|\eta|^2+|\xi|^2) = \nu\, |y|^2\, |\hat u(y)|^2.$$

Hence,

$$
\begin{aligned}
A(u,u) &= \sum_{i,j=1}^{n}\sum_{k,l=1}^{N}\int_{\mathbb{R}^n} a_{ij}^{kl}\, D_i u^k(x)\, D_j u^l(x)\, dx \\
&= \mathrm{Re}\sum_{i,j=1}^{n}\sum_{k,l=1}^{N}\left(\int_{\mathbb{R}^n} a_{ij}^{kl}\, y_i\, y_j\, \widehat{u^k}(y)\, \overline{\widehat{u^l}(y)}\, dy\right) \\
&\geq \nu \int_{\mathbb{R}^n} |y|^2\, |\hat{u}(y)|^2\, dy = \nu \int_{\mathbb{R}^n} |i\hat{u}(y)\otimes y|^2\, dy \\
&= \nu \int_{\mathbb{R}^n} |\widehat{Du}(y)|^2\, dy = \nu \int_{\mathbb{R}^n} |Du(x)|^2\, dx.
\end{aligned}
$$

The proof is complete. □

Theorem 3.13. *Let* $B_2(u,v)$ *be defined by* (3.7). *Assume*

1) $a_{ij}^{kl} \in C(\bar{\Omega})$,

2) the Legendre-Hadamard condition *holds for all* $x \in \Omega$; *that is,*

$$a_{ij}^{kl}(x)\, q^k q^l\, p_i\, p_j \geq \nu\, |p|^2\, |q|^2, \quad \forall p \in \mathbb{R}^n,\ q \in \mathbb{R}^N.$$

3) $b_i^{kl},\ c^{kl},\ d_i^{kl} \in L^\infty(\Omega)$.

Then, there exist constants $\lambda_0 > 0$ *and* $\lambda_1 \geq 0$ *such that*

$$B_2(u,u) \geq \lambda_0\, \|u\|_H^2 - \lambda_1\, \|u\|_{L^2}^2, \quad \forall\, u \in H_0^1(\Omega;\mathbb{R}^N).$$

Proof. By uniform continuity, we can choose a small $\epsilon > 0$ such that

$$|a_{ij}^{kl}(x) - a_{ij}^{kl}(y)| \leq \nu/2, \quad \forall\, x,\, y \in \bar{\Omega},\ |x-y| \leq \epsilon.$$

We claim

$$\int_\Omega a_{ij}^{kl}(x)\, D_i u^k\, D_j u^l\, dx \geq \frac{\nu}{2}\int_\Omega |Du(x)|^2\, dx \tag{3.19}$$

for all test functions $u \in C_0^\infty(\Omega;\mathbb{R}^N)$ with diameter of the support $\mathrm{diam}(\mathrm{supp}\, u) \leq \epsilon$. To see this, we choose any point $x_0 \in \mathrm{supp}\, u$. Then, by Lemma 3.12,

$$
\begin{aligned}
\int_\Omega a_{ij}^{kl}(x)\, D_i u^k\, D_j u^l\, dx &= \int_\Omega a_{ij}^{kl}(x_0)\, D_i u^k\, D_j u^l\, dx \\
&\quad + \int_{\mathrm{supp}\, u} \big(a_{ij}^{kl}(x) - a_{ij}^{kl}(x_0)\big)\, D_i u^k\, D_j u^l\, dx \\
&\geq \nu \int_\Omega |Du(x)|^2\, dx - \frac{\nu}{2}\int_\Omega |Du(x)|^2\, dx,
\end{aligned}
$$

which proves (3.19). We now cover $\bar{\Omega}$ with finitely many open balls $\{B_{\epsilon/4}(x^m)\}$ with $x^m \in \Omega$ and $m = 1, 2, ..., M$. For each m, let $\zeta_m \in C_0^\infty(B_{\epsilon/2}(x^m))$ with $\zeta_m(x) = 1$ for $x \in B_{\epsilon/4}(x^m)$. Since for any $x \in \bar{\Omega}$ we have at least one m such that $x \in B_{\epsilon/4}(x^m)$ and thus $\zeta_m(x) = 1$, we may therefore define

$$\varphi_m(x) = \frac{\zeta_m(x)}{\big(\sum_{j=1}^M \zeta_j^2(x)\big)^{1/2}}, \quad m = 1, 2, ..., M.$$

Then $\sum_{m=1}^{M} \varphi_m^2(x) = 1$ for all $x \in \Omega$. (This is a special case of *partition of unity.*) We have thus

$$a_{ij}^{kl}(x)\, D_i u^k\, D_j u^l = \sum_{m=1}^{M} \left(a_{ij}^{kl}(x)\, \varphi_m^2\, D_i u^k\, D_j u^l \right) \tag{3.20}$$

and each term (no summation on m)

$$a_{ij}^{kl}(x)\, \varphi_m^2\, D_i u^k\, D_j u^l = a_{ij}^{kl}(x)\, D_i(\varphi_m\, u^k)\, D_j(\varphi_m\, u^l)$$
$$-a_{ij}^{kl}(x) \left(\varphi_m\, D_i\varphi_m\, u^l\, D_i u^k + \varphi_m\, D_i\varphi_m\, u^k\, D_j u^l + D_i\varphi_m\, D_j\varphi_m\, u^k\, u^l \right).$$

Since $\varphi_m u \in C_0^\infty(\Omega \cap B_{\epsilon/2}(x^m); \mathbb{R}^N)$ and $\operatorname{diam}(\Omega \cap B_{\epsilon/2}(x^m)) \le \epsilon$, we have by (3.19)

$$\int_\Omega a_{ij}^{kl}(x)\, D_i(\varphi_m\, u^k)\, D_j(\varphi_m\, u^l)\, dx \ge \frac{\nu}{2} \int_\Omega |D(\varphi_m\, u)|^2\, dx.$$

Note also that

$$|D(\varphi_m\, u)|^2 = \varphi_m^2\, |Du|^2 + |D\varphi_m|^2\, |u|^2 + 2\, \varphi_m\, D_i\varphi_m\, u^k\, D_i u^k.$$

Therefore, we have by (3.20) and the fact that $\sum_{m=1}^{M} \varphi_m^2 = 1$ on Ω,

$$\int_\Omega a_{ij}^{kl}(x)\, D_i u^k\, D_j u^l\, dx$$
$$\ge \frac{\nu}{2} \int_\Omega |Du|^2\, dx - C_1\, \|u\|_{L^2}\, \|Du\|_{L^2} - C_2\, \|u\|_{L^2}^2.$$

The terms in $B_2(u,u)$ involving b, c and d can be estimated by $\|u\|_{L^2}\, \|Du\|_{L^2}$ and $\|u\|_{L^2}^2$. Finally, by all of these estimates and the inequality

$$ab \le \epsilon a^2 + \frac{1}{4\epsilon}\, b^2$$

we have $B_2(u,u) \ge \lambda_0\, \|u\|_H^2 - \lambda_1\, \|u\|_{L^2}^2$ for all $u \in H_0^1(\Omega; \mathbb{R}^N)$. This completes the proof. □

Note that the bilinear form $B^\lambda(u,v) = B_2(u,v) + \lambda\, (u,v)_{L^2}$ satisfies the condition of the Lax-Milgram theorem on $H = H_0^1(\Omega; \mathbb{R}^N)$ for all $\lambda \ge \lambda_1$; thus, by the Lax-Milgram theorem, we easily obtain the following existence result.

Theorem 3.14. *Under the hypotheses of the previous theorem, for $\lambda \ge \lambda_1$, the Dirichlet problem for the linear system*

$$-\operatorname{div}(A(x,u,Du)) + b(x,u,Du) + \lambda u = F, \quad u|_{\partial\Omega} = 0 \tag{3.21}$$

has a unique weak solution u in $H_0^1(\Omega; \mathbb{R}^N)$ for any bounded linear functional F on H. Moreover, the solution u satisfies $\|u\|_H \le C\, \|F\|$ with a constant C depending on λ and the L^∞-norms of the coefficients of linear terms $A(x,s,\xi)$ and $b(x,s,\xi)$ given above.

Corollary 3.15. *Given $\lambda \ge \lambda_1$ as in the theorem, the operator $\mathcal{K}\colon L^2(\Omega; \mathbb{R}^N) \to L^2(\Omega; \mathbb{R}^N)$, where, for each $F \in L^2(\Omega; \mathbb{R}^N)$, $u = \mathcal{K}F$ is the unique weak solution to the BVP (3.21) above, is a compact linear operator.*

Proof. By the theorem, $\|u\|_{H_0^1(\Omega;\mathbb{R}^N)} \le C\|F\|_{L^2(\Omega;\mathbb{R}^N)}$. Hence $\mathcal{K}$ is a bounded linear operator from $L^2(\Omega; \mathbb{R}^N)$ to $H_0^1(\Omega; \mathbb{R}^N)$. So, by the compact embedding $H_0^1(\Omega; \mathbb{R}^N) \to L^2(\Omega; \mathbb{R}^N)$, the linear operator $\mathcal{K}\colon L^2(\Omega; \mathbb{R}^N) \to L^2(\Omega; \mathbb{R}^N)$ is compact. □

3.4. Regularity of Weak Solutions

We now discuss the question as to whether a weak solution is smooth. We restrict ourselves to the case (3.1) of single equation for one unknown scalar u. We also emphasize the interior regularity; see EVAN's book for more.

Consider the equation

$$Lu = f \text{ in } \Omega, \tag{3.22}$$

where L is given by

$$Lu \equiv -\sum_{i,j=1}^{n} D_i\left(a_{ij}(x)D_j u\right) + \sum_{i=1}^{n} b_i(x)D_i u + c(x)u$$

with uniform ellipticity condition: for some $\theta > 0$,

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq \theta \sum_{i=1}^{n} |\xi_i|^2 \quad \forall\ x \in \Omega,\ \xi \in \mathbb{R}^n. \tag{3.23}$$

Assume $f \in L^2(\Omega)$. A function $u \in H^1(\Omega)$ is called a **weak solution** to $Lu = f$ if $B_1(u, v) = \int_\Omega fv\, dx$ holds for all $v \in H_0^1(\Omega)$, where $B_1(u, v)$ is the bilinear form defined before. Note that we do not assume $u \in H_0^1(\Omega)$.

We first prove the following interior regularity result.

Theorem 3.16. (Interior H^2-Regularity) *In addition to ellipticity condition, assume*

$$a_{ij} \in C^1(\Omega), \quad b_i,\, c \in L^\infty(\Omega).$$

Suppose $f \in L^2(\Omega)$ and $u \in H^1(\Omega)$ is a weak solution to $Lu = f$. Then $u \in H^2_{loc}(\Omega)$. Moreover, for each $\Omega' \subset\subset \Omega$, we have the estimate

$$\|u\|_{H^2(\Omega')} \leq C(\|u\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}), \tag{3.24}$$

where constant C depends only on Ω', Ω, and the coefficient of L.

Proof. Set $q = f - \sum_{i=1}^n b_i D_i u - cu$. Since u is a weak solution of $Lu = f$, it satisfies

$$\int_\Omega \sum_{i,j=1}^{n} a_{ij} D_i u D_j \varphi\, dx = \int_\Omega q\varphi\, dx \quad \text{for all} \quad \varphi \in H_0^1(\Omega). \tag{3.25}$$

Choose an open set Ω'' such that $\Omega' \subset\subset \Omega'' \subset\subset \Omega$. Then select a smooth function η satisfying

$$\eta = 1 \text{ on } \Omega', \quad \operatorname{supp}\eta \subset \Omega'', \quad 0 \leq \eta \leq 1.$$

First, set $\varphi = \eta^2 u$ in (3.25) and use $D_j\varphi = \eta^2 D_j u + 2\eta u D_j\eta$ to discover

$$\int_\Omega \sum_{i,j=1}^{n} a_{ij}(\eta D_i u)(\eta D_j u)\, dx = -\int_\Omega \sum_{i,j=1}^{n} 2a_{ij}u(D_j\eta)(\eta D_i u)\, dx + \int_\Omega q\eta^2 u\, dx.$$

By the ellipticity condition, the left-hand side of this identity is $\geq \theta \int_\Omega \eta^2|\nabla u|^2\, dx$, while the right-hand side, by the definition of q and Cauchy's inequality with ϵ, is no greater than

$$\epsilon \int_\Omega \eta^2|\nabla u|^2\, dx + C_\epsilon \int_\Omega (u^2 + f^2)\, dx.$$

Hence choosing $\epsilon = \theta/2$ we have shown that

$$\int_\Omega \eta^2|\nabla u|^2\, dx \leq C \int_\Omega (u^2 + f^2)\, dx.$$

Thus

$$\|u\|_{H^1(\Omega'')} \le C(\|u\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}). \tag{3.26}$$

We now prove a similar H^2-estimate. Since we cannot differentiate the equation $Lu = f$, we need to use difference quotient operator δ^h defined earlier.

For $0 < |h| < \frac{1}{2}\operatorname{dist}(\Omega'', \partial\Omega)$ and $k = 1, 2, \cdots, n$, we choose the test function $\varphi = \delta_k^{-h} v$ with $v = \eta^2 \delta_k^h u$, in (3.25) and use $D_j\varphi = \delta_k^{-h}(D_j v)$ and integration by parts for difference quotients to obtain

$$\int_\Omega [\delta_k^h(\sum_{i,j=1}^n a_{ij} D_i u)]\, D_j v dx = -\int_\Omega q\, \delta_k^{-h} v dx.$$

Henceforth, we omit the $\sum$ sign. Using the definition of q and the equality

$$\delta_k^h(a_{ij}D_i u) = a_{ij}^h(x)(\delta_k^h D_i u) + (D_i u)(\delta_k^h a_{ij}),$$

where $a_{ij}^h(x) = a_{ij}(x + he_k)$, and noting that supp $v \subset \Omega''$, we get

$$\begin{aligned} \int_\Omega a_{ij}^h (D_i \delta_k^h u)(D_j v) dx &= -\int_\Omega \left((\delta_k^h a_{ij}) D_i u D_j v + q\,(\delta_k^{-h} v) \right) dx \\ &\le c\left(\|u\|_{H^1(\Omega'')} + \|f\|_{L^2(\Omega)}\right) \|\nabla v\|_{L^2(\Omega)}. \end{aligned}$$

Since $D_j v = \eta^2 D_j(\delta_k^h u) + 2\eta(D_j\eta)\delta_k^h u$, it follows that

$$\begin{aligned} \int_\Omega \eta^2 a_{ij}^h (D_i \delta_1^h u)(D_j \delta_1^h u) \le -2 \int_\Omega \eta a_{ij}^h (D_i \delta_k^h u)(D_j \eta) \delta_k^h u \\ + c\left(\|u\|_{H^1(\Omega'')} + \|f\|_{L^2(\Omega)}\right)(\|\eta \nabla \delta_k^h u\|_{L^2(\Omega)} + \|\delta_k^h u \nabla\eta\|_{L^2(\Omega)}). \end{aligned}$$

Using the ellipticity condition and Cauchy's inequality with ϵ, we obtain

$$\frac{\theta}{2}\int_\Omega |\eta \delta_k^h \nabla u|^2 dx \le c\int_\Omega |\nabla\eta|^2 |\delta_k^h u|^2 dx + c\left(\|u\|^2_{H^1(\Omega'')} + \|f\|^2_{L^2(\Omega)}\right).$$

Hence

$$\|\eta\delta_k^h \nabla u\|^2_{L^2(\Omega)} \le C\left(\|u\|^2_{H^1(\Omega'')} + \|f\|^2_{L^2(\Omega)}\right).$$

Since $\eta = 1$ on Ω', we derive that $(\nabla u)_{x_k} \in L^2(\Omega')$ for all $k = 1, 2, \cdots, n$. Hence $u \in H^2(\Omega')$. Moreover,

$$\|D^2 u\|_{L^2(\Omega')} \le C\left(\|u\|_{H^1(\Omega'')} + \|f\|_{L^2(\Omega)}\right).$$

By (3.26), we obtain the estimate (3.24). $\square$

Theorem 3.17. (Global H^2-Regularity) *Assume in addition to the assumptions of Theorem 3.16 that $a_{ij} \in C^1(\bar\Omega)$ and $\partial\Omega \in C^2$. Then a weak solution u of $Lu = f$ satisfying $u \in H_0^1(\Omega)$ belongs to $H^2(\Omega)$, and*

$$\|u\|_{H^2(\Omega)} \le C(\|u\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}) \tag{3.27}$$

where the constant C depends only on $n, \|a_{ij}\|_{W^{1,\infty}(\Omega)}, \|b_i\|_{L^\infty(\Omega)}, \|c\|_{L^\infty(\Omega)}$ and $\partial\Omega$.

Proof. Cover Ω by a finite number of balls. In each ball containing a part of $\partial\Omega$, we apply the usual transformation to flatten the boundary, and then derive the corresponding boundary estimate in a half ball. In each ball not containing a part of $\partial\Omega$, we apply the above interior estimate. $\square$

By using an induction argument, we can also get higher regularity for the solution.

Theorem 3.18. (Interior Higher Regularity) *Assume L is uniformly elliptic, and let $a_{ij}, b_i, c \in C^{k+1}(\Omega)$, $f \in H^k(\Omega)$. If $u \in H^1(\Omega)$ is a weak solution of $Lu = f$, then for any $\Omega' \subset\subset \Omega$ we have $u \in H^{k+2}(\Omega')$.*

Theorem 3.19. (Global Higher Regularity) *Assume in addition to the assumptions of Theorem 3.18 that $a_{ij}, b_i, c \in C^{k+1}(\bar{\Omega})$, $\partial\Omega \in C^{k+2}$ and $f \in H^k(\Omega)$. Then a weak solution u of $Lu = f$ satisfying $u \in H_0^1(\Omega)$ belongs to $H^{k+2}(\Omega)$, and*

$$\|u\|_{H^{k+2}(\Omega)} \le C(\|u\|_{L^2(\Omega)} + \|f\|_{H^k(\Omega)}) \tag{3.28}$$

where the constant C is independent of u and f. Furthermore, if the only weak solution $u \in H_0^1(\Omega)$ of $Lu = 0$ is $u \equiv 0$, then

$$\|u\|_{H^{k+2}(\Omega)} \le C\|f\|_{H^k(\Omega)} \tag{3.29}$$

where C is independent of u and f.

Proof. We just prove the last statement for $k = 0$; the more general case is similar. In view of (3.28), it suffices to show that

$$\|u\|_{L^2(\Omega)} \le c\|f\|_{L^2(\Omega)}$$

If to the contrary this inequality is false, there would exist sequences $u_n \in H^2(\Omega) \cap H_0^1(\Omega)$ and $f_n \in L^2(\Omega)$ for which $\|u_n\|_2 = 1$ and $\|f_n\|_2 \to 0$. By (3.28) we have $\|u_n\|_{2,2} \le c$. Thus we can assume that u_n converges weakly to u in $H^2(\Omega)$ and strongly in $L^2(\Omega)$. For fixed $v \in H_0^1(\Omega)$, $B_1(u, v) \in (H^2(\Omega))^*$, and so by passing to the limit in

$$B_1(u_n, v) = \int_\Omega f_n v dx \quad \text{for all} \quad v \in H_0^1(\Omega)$$

we see that u is a weak solution of $Lu = 0$. Hence, $u \equiv 0$ by weak uniqueness. This contradicts $\|u\|_2 = \lim_{n\to\infty} \|u_n\|_2 = 1$. □

Remark. Similar results are valid for the more general boundary condition $B_\varepsilon u|_{\partial\Omega} = 0$. Moreover, from the above regularity results and the Sobolev imbedding theorem, we easily deduce the following: If $u \in H^1(\Omega)$ is a weak solution of $Lu = f \in H^k(\Omega)$, $B_\varepsilon u|_{\partial\Omega} = 0$, where $k > n/2$, then $u \in C^2(\bar{\Omega})$. In particular, if $u \in H^1(\Omega)$ is a weak solution of $Lu = f$, $B_\varepsilon u|_{\partial\Omega} = 0$, and if $a_{ij}, a_i, a, f \in C^\infty(\bar{\Omega})$, then $u \in C^\infty(\bar{\Omega})$.

The following global results on L^p estimates will play an important role in studying nonlinear problems.

Theorem 3.20. *Suppose L is uniformly elliptic with $a_{ij} \in C^1(\bar{\Omega}), b_i, c \in C(\bar{\Omega})$ and let $\partial\Omega \in C^2$. If $u \in W^{1,p}(\Omega)$, $1 < p < \infty$, satisfies*

$$B_1(u, \phi) = \int_\Omega f\phi dx \quad \textit{for all} \quad \phi \in C_0^\infty(\Omega),$$

then $u \in W^{2,p}(\Omega)$. Moreover, for all $u \in W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega)$

$$\|u\|_{W^{2,p}(\Omega)} \le c(\|Lu\|_{L^p(\Omega)} + \|u\|_{L^p(\Omega)}) \tag{3.30}$$

where the constant c is independent of u.

Theorem 3.21. *Let $\partial\Omega \in C^{k+2,\alpha}(k \ge 0)$. Suppose $u \in C(\bar{\Omega}) \cap C^2(\Omega)$ satisfies the uniformly elliptic problem $Lu = f$, $u|_{\partial\Omega} = 0$, where f and the coefficients belong to $C^{k,\alpha}(\bar{\Omega})$. Then $u \in C^{k+2,\alpha}(\bar{\Omega})$.*

Part of the next result improves Theorem 3.21.

Theorem 3.22. *Let $u \in W^{2,p}_{loc}(\Omega)$ be a solution of the uniformly elliptic equation $Lu = f$. Suppose the coefficients of L belong to $C^{k-1,1}(\bar{\Omega})(C^{k-1,\alpha}(\bar{\Omega}))$, $f \in W^{k,q}(\Omega)(C^{k-1,\alpha}(\bar{\Omega}))$, and $\partial\Omega \in C^{k+1,1}(C^{k+1,\alpha})$, with $1 < p, q < \infty,\ k \geq 1,\ 0 < \alpha < 1$. Then $u \in W^{k+2,q}(\Omega)(C^{k+1,\alpha}(\bar{\Omega}))$.*

3.5. Symmetric Elliptic Operators and Eigenvalue Problems

3.5.1. Symmetric Elliptic Operators. In what follows, we assume $\Omega \subset \mathbb{R}^n$ is a bounded domain and $N \geq 1$ is integer. We consider the operator

$$Lu = -\operatorname{div} A(x, Du) + c(x)u, \quad u \in H^1_0(\Omega; \mathbb{R}^N),$$

where $A(x, Du)$ is a linear system defined with $A(x, \xi)$, $\xi \in \mathbb{M}^{N\times n}$, given by

$$A^k_i(x,\xi) = \sum_{1\leq l\leq N,\, 1\leq j\leq n} a^{kl}_{ij}(x)\xi^l_j.$$

Here $a^{kl}_{ij}(x)$ and $c(x)$ are given functions in $L^\infty(\Omega)$. The bilinear form associated to L is

$$B(u,v) = \int_\Omega (\sum_{k,l=1}^{N}\sum_{i,j=1}^{n} a^{kl}_{ij}(x) D_j u^l D_i v^k + c(x) u\cdot v)\, dx, \quad u,\, v \in H^1_0(\Omega;\mathbb{R}^N). \tag{3.31}$$

In order for B to be *symmetric* on $H^1_0(\Omega;\mathbb{R}^N)$, that is, $B(u,v) = B(v,u)$ for all $u, v \in H^1_0(\Omega;\mathbb{R}^N)$, we need the following symmetry condition:

$$a^{kl}_{ij}(x) = a^{lk}_{ji}(x), \quad \forall\ i,\ j = 1,2,\cdots,n;\ \ k,\ l = 1,2,\cdots,N. \tag{3.32}$$

We also assume the Gårding inequality holds (see Theorem 3.13 for sufficient conditions):

$$B(u,u) \geq \sigma\|u\|^2_{H^1_0} - \mu\|u\|^2_{L^2}, \quad \forall\ u \in H^1_0(\Omega;\mathbb{R}^N), \tag{3.33}$$

where $\sigma > 0$ and $\mu \in \mathbb{R}$ are constants.

3.5.2. The Compact Inverse. For each $F \in L^2(\Omega;\mathbb{R}^N)$, define $u = \mathcal{K}F$ to be the unique weak solution in $H^1_0(\Omega;\mathbb{R}^N)$ of the BVP

$$Lu + \mu u = F \quad \text{in } \Omega, \quad u|_{\partial\Omega} = 0.$$

By Theorem 3.14 and Corollary 3.15, this $\mathcal{K}$ is well defined and is a compact linear operator on $L^2(\Omega;\mathbb{R}^N)$. Sometime, we write $\mathcal{K} = (L+\mu I)^{-1}$. Here I denotes the identity on $L^2(\Omega;\mathbb{R}^N)$ and also the identity embedding of $H^1_0(\Omega;\mathbb{R}^N)$ into $L^2(\Omega;\mathbb{R}^N)$.

Theorem 3.23. $\mathcal{K}\colon L^2(\Omega;\mathbb{R}^N) \to L^2(\Omega;\mathbb{R}^N)$ *is symmetric and positive; that is,*

$$(\mathcal{K}F, G)_{L^2} = (\mathcal{K}G, F)_{L^2}, \quad (\mathcal{K}F, F)_{L^2} \geq 0, \quad \forall\ F,\ G \in L^2(\Omega;\mathbb{R}^N).$$

Furthermore, given $\lambda \in \mathbb{R}$ and $F \in L^2(\Omega;\mathbb{R}^N)$, $u \in H^1_0(\Omega;\mathbb{R}^N)$ is a weak solution of $Lu - \lambda u = F$ if and only if $[I - (\lambda+\mu)\mathcal{K}]u = \mathcal{K}F$.

Proof. Let $u = \mathcal{K}F$ and $v = \mathcal{K}G$. Then

$$(u, G)_{L^2} = B(v,u) + \mu(v,u)_{L^2} = B(u,v) + \mu(v,u)_{L^2} = (v, F)_{L^2},$$

proving the symmetry. Also, by (3.33),

$$(\mathcal{K}F, F)_{L^2} = (u, F)_{L^2} = B(u,u) + \mu\|u\|^2_{L^2} \geq \sigma\|u\|^2_{H^1_0} = \sigma\|\mathcal{K}F\|^2_{H^1_0} \geq 0. \tag{3.34}$$

Finally, $u \in H^1_0(\Omega;\mathbb{R}^N)$ is a weak solution of $Lu - \lambda u = F$ if and only if $Lu + \mu u = F + (\lambda+\mu)u$, which is equivalent to the equation $u = \mathcal{K}[F + (\lambda+\mu)u] = \mathcal{K}F + (\lambda+\mu)\mathcal{K}u$; that is, $[I - (\lambda+\mu)\mathcal{K}]u = \mathcal{K}F$. $\square$

3.5.3. Orthogonality Conditions. From Theorem 3.23 above, the BVP

$$Lu = F, \quad u|_{\partial\Omega} = 0 \tag{3.35}$$

has a solution if and only if $\mathcal{K}F \in \mathcal{R}(I - \mu\mathcal{K}) = [\mathcal{N}(I - \mu\mathcal{K})]^\perp$.

If $\mu = 0$, the Lax-Milgram theorem already implies that the problem (3.35) has a unique weak solution $u = \mathcal{K}F$. If $\mu \neq 0$, then it is easy to see that

$$\mathcal{N} = \mathcal{N}(I - \mu\mathcal{K}) = \{v \in H_0^1(\Omega; \mathbb{R}^N) \,|\, Lv = 0\}.$$

Note that, by the Fredholm Alternative Theorem, this null space $\mathcal{N}$ is of finite dimension k, and let $\{v_1, v_2, \cdots, v_k\}$ be a basis of this null space. Then (3.35) is solvable if and only if the following **orthogonality condition** holds:

$$\int_\Omega F \cdot G_i \, dx = 0, \quad i = 1, 2, \cdots, k,$$

where $G_i = \mathcal{K}v_i$; that is, G_i is the unique weak solution to the BVP: $LG_i + \mu G_i = v_i$, in $H_0^1(\Omega; \mathbb{R}^N)$.

3.5.4. Eigenvalue Problems. A number $\lambda \in \mathbb{R}$ is called a (Dirichlet) **eigenvalue** of operator L if the BVP problem

$$Lu - \lambda u = 0, \quad u|_{\partial\Omega} = 0$$

has *nontrivial* weak solutions in $H_0^1(\Omega; \mathbb{R}^N)$; these nontrivial solutions are called the **eigenfunctions** corresponding to eigenvalue λ.

From Theorem 3.23, we see that λ is an eigenvalue of L if and only if equation $(I - (\lambda + \mu)\mathcal{K})u = 0$ has nontrivial solutions $u \in L^2(\Omega; \mathbb{R}^N)$; this exactly says that $\lambda \neq -\mu$ and $\frac{1}{\lambda+\mu}$ is an eigenvalue of operator $\mathcal{K}$. Since, by (3.34), $\mathcal{K}$ is strictly positive, all eigenvalues of $\mathcal{K}$ consist of a countable set of positive numbers tending to zero and hence the eigenvalues of L consist of a set of numbers $\{\lambda_j\}_{j=1}^\infty$ with $-\mu < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_j \to \infty$.

Theorem 3.24. (Eigenvalue Theorem) *Assume* (3.32) *and* (3.33) *with* $\mu = 0$. *Then the eigenvalues of* L *consist of a countable set* $\Sigma = \{\lambda_k\}_{k=1}^\infty$, *where*

$$0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots$$

and

$$\lim_{k\to\infty} \lambda_k = \infty.$$

Note that λ_1 *is called the (Dirichlet)* **principal eigenvalue** *of* L. *Let* w_k *be an eigenfunction corresponding to* λ_k *satisfying* $\|w_k\|_{L^2(\Omega;\mathbb{R}^N)} = 1$. *Then* $\{w_k\}_{k=1}^\infty$ *forms an* **orthonormal basis** *of* $L^2(\Omega; \mathbb{R}^N)$.

Note that

$$\lambda_1 = \min_{u \in H_0^1(\Omega;\mathbb{R}^N),\, \|u\|_{L^2}=1} B(u, u). \tag{3.36}$$

Theorem 3.25. *Let* $N = 1$ *and let* w_1 *be an eigenfunction corresponding to the principal eigenvalue* λ_1. *Then, either* $w_1(x) > 0$ *for all* $x \in \Omega$ *or* $w_1(x) < 0$ *for all* $x \in \Omega$. *Moreover, the* **eigenspace** *corresponding to* λ_1 *is one-dimensional.*

Proof. The first part of the theorem relies on the characterization (3.36) and a maximum principle which we do not study here. We prove only the second part. Let w be another eigenfunction. Then, either $w(x) > 0$ for all $x \in \Omega$ or $w(x) < 0$ for all $x \in \Omega$. Let $t \in \mathbb{R}$ be such that

$$\int_\Omega w(x)\, dx = t \int_\Omega w_1(x) dx.$$

Note that $u = w - tw_1$ is also a solution to $Lu = \lambda_1 u$. We claim $u \equiv 0$ and hence $w = tw_1$, proving the eigenspace is one-dimensional. Suppose $u \neq 0$. Then u is another eigenfunction corresponding to λ_1. Then, by the theorem, we would have either $u(x) > 0$ for all $x \in \Omega$ or $u(x) < 0$ for all $x \in \Omega$ and hence $\int_\Omega u(x)dx \neq 0$, which is a contradiction. □

Remark. Why is the eigenvalue problem important? In one specific case, one can use eigenvalues and eigenfunctions to study some evolution (i.e. time-dependent) problems. For example, to solve the **initial boundary value problem** (IBVP) for the time-dependent parabolic equation:

$$u_t + Lu = 0, \quad u|_{\partial\Omega} = 0, \quad u(x,0) = \varphi(x), \tag{3.37}$$

one can try to find the special solutions of the form: $u(x,t) = e^{-\lambda t} w(x)$; this reduces to the eigen-problem: $Lw = \lambda w$. Therefore, for each pair (λ_i, w_i) of eigenvalue and eigenfunction, one obtains a special solution to the evolution equation given by $u_i(x,t) = e^{-\lambda_i t} w_i(x)$. Then one proceeds to solve the IBVP (3.37) by finding the solution of the form

$$u(x,t) = \sum_i a_i e^{-\lambda_i t} w_i(x),$$

where a_i is determined by the eigen-expansion of the initial data φ:

$$\varphi(x) = \sum_i a_i w_i(x).$$

However, this course does not study the parabolic or other time-dependent problems.

Chapter 4

Variational Methods for Nonlinear PDEs

4.1. Variational Problems

This chapter and the next will discuss some methods for solving the boundary value problem for certain nonlinear partial differential equations. All these problems can be written in the abstract form:

$$\mathcal{A}[u] = 0 \quad \text{in } \Omega, \quad \mathcal{B}[u] = 0 \quad \text{on } \partial\Omega, \tag{4.1}$$

where $\mathcal{A}[u]$ denotes a given PDE for unknown u and $\mathcal{B}[u]$ is a given boundary value condition. There is, of course, no general theory for solving such problems.

The *Calculus of Variations* identifies an important class of problems which can be solved using relatively simple techniques from nonlinear functional analysis. This is the class of **variational problems**, where the operator $\mathcal{A}[u]$ can be formulated as the **first variation** ("derivative") of an appropriate "energy" functional $I(u)$ on a Banach space X; that is, $\mathcal{A}[u] = I'(u)$. In this way, $\mathcal{A}\colon X \to X^*$ and the equation $\mathcal{A}[u] = 0$ can be formulated weakly as

$$\langle I'(u), v\rangle = 0, \quad \forall\ v \in X.$$

The advantage of this new formulation is that solving problem (4.1) (at least *weakly*) is equivalent to finding the **critical points** of I on X. The **minimization method** for a variational problem is to solve the problem by finding the minimizers of the related energy functional. In this chapter and the next, we shall only study the variational problems on the Sobolev space $X = W^{1,p}(\Omega; \mathbb{R}^N)$.

We should also mention that many of the physical laws in applications arise directly as variational principles. However, although powerful, not *all* PDE problems can be formulated as variational problems; there are lots of other (nonvariational) important methods for studying PDEs.

4.2. Multiple Integrals in the Calculus of Variations

Consider the multiple integral functional

$$I(u) = \int_\Omega F(x, u(x), Du(x))\, dx, \tag{4.2}$$

where $F(x, s, \xi)$ is a given function on $\Omega \times \mathbb{R}^N \times \mathbb{M}^{N\times n}$.

4.2.1. First Variation and Euler-Lagrange Equations. Suppose $F(x, s, \xi)$ is continuous and is also smooth in s and ξ. Assume u is a nice (say, $u \in C^1(\bar\Omega; \mathbb{R}^N)$) minimizer of $I(u)$ with its own boundary data; that is, u is a map such that

$$I(u) \le I(u + t\varphi)$$

for all $t \in \mathbf{R}^1$ and $\varphi \in C_0^\infty(\Omega; \mathbb{R}^N)$. Then by taking derivative of $I(u + t\,\varphi)$ at $t = 0$ we see that u satisfies

$$\int_\Omega \big(F_{\xi_i^k}(x, u, Du)\, D_i\varphi^k(x) + F_{s^k}(x, u, Du)\, \varphi^k(x)\big)\, dx = 0 \tag{4.3}$$

for all $\varphi \in C_0^\infty(\Omega; \mathbb{R}^N)$. (Summation notation is used here.) The left-hand side is called the **first variation** of I at u. Since this holds for all test functions, we conclude after integration by parts:

$$-\operatorname{div} A(x, u, Du) + b(x, u, Du) = 0, \tag{4.4}$$

where A, b are defined by

$$A_i^k(x, s, \xi) = F_{\xi_i^k}(x, s, \xi), \quad b^k(x, s, \xi) = F_{s^k}(x, s, \xi). \tag{4.5}$$

This coupled system of nonlinear partial differential equations in divergence form is called the system of **Euler-Lagrange equations** for the functional $I(u)$.

4.2.2. Second Variation and Legendre-Hadamard Conditions. If F, u are sufficiently smooth (e.g. of class C^2) then, at the minimizer u, we have

$$\frac{d^2}{dt^2} I(u + t\varphi)\Big|_{t=0} \ge 0.$$

This implies

$$\int_\Omega \big[F_{\xi_i^k\,\xi_j^l}(x, u, Du)\, D_i\varphi^k D_j\varphi^l\ + 2F_{\xi_i^k\,s^l}(x, u, Du)\, \varphi^l D_i\varphi^k + F_{s^k s^l}(x, u, Du)\, \varphi^k\varphi^l\big] \ge 0 \tag{4.6}$$

for all $\varphi \in C_0^\infty(\Omega; \mathbb{R}^N)$. The left-hand side of this inequality is called the **second variation** of I at u.

We can extract some useful information from (4.6). Note that routine approximation argument shows that (4.6) is also valid for all functions $\varphi \in W_0^{1,\infty}(\Omega; \mathbb{R}^N)$ (that is, all Lipschitz functions vanishing on $\partial\Omega$). Let $\rho\colon \mathbb{R} \to \mathbb{R}$ be the periodic *zig-zag* function of period 1 defined by

$$\rho(t + 1) = \rho(t) \quad (t \in \mathbb{R}), \tag{4.7}$$

$$\rho(t) = t \quad \text{if } 0 \le t \le \tfrac{1}{2};\ \rho(t) = 1 - t \quad \text{if } \tfrac{1}{2} \le t \le 1. \tag{4.8}$$

Given any vectors $p \in \mathbb{R}^n$, $q \in \mathbb{R}^N$ and $\epsilon > 0$, define $\varphi \in W_0^{1,\infty}(\Omega; \mathbb{R}^N)$ by

$$\varphi(x) = \epsilon\, \rho(\frac{x \cdot p}{\epsilon})\zeta(x)q, \quad \forall\ x \in \Omega,$$

where $\zeta \in C_0^\infty(\Omega)$ is a given scalar test function. Note that $D_i\varphi^k(x) = \rho'(\frac{x\cdot p}{\epsilon})p_i q^k \zeta + O(\epsilon)$ as $\epsilon \to 0^+$. Substitute this φ into (4.6) and let $\epsilon \to 0^+$ and we obtain

$$\int_\Omega \big[\sum_{i,j=1}^n \sum_{k,l=1}^N F_{\xi_i^k\,\xi_j^l}(x,u,Du)\,p_i p_j q^k q^l\big]\zeta^2\,dx \geq 0.$$

Since this holds for all $\zeta \in C_0^\infty(\Omega)$, we deduce

$$\sum_{i,j=1}^n \sum_{k,l=1}^N F_{\xi_i^k\,\xi_j^l}(x,u,Du)\,p_i p_j q^k q^l \geq 0, \quad \forall\; x\in\Omega,\; p\in\mathbb{R}^n,\; q\in\mathbb{R}^N. \tag{4.9}$$

This is the weak **Legendre-Hadamard condition** for F at the minimum point u.

4.2.3. Ellipticity Conditions and Convexities. We now consider the ellipticity of the *Euler-Lagrange* equation (4.4), where $A(x,s,\xi)$, $b(x,s,\xi)$ are given by (4.5) and $F(x,s,\xi)$ is C^2 in ξ. In this case, the Legendre ellipticity condition (3.8) and the Legendre-Hadamard condition (3.9) defined in the previous chapter reduce to, respectively:

$$F_{\xi_i^k\,\xi_j^l}(x,s,\xi)\,\eta_i^k\,\eta_j^l \geq \nu\,|\eta|^2 \quad \forall \eta\in\mathbb{M}^{N\times n}; \tag{4.10}$$

$$F_{\xi_i^k\,\xi_j^l}(x,s,\xi)\,q^k q^l\,p_i p_j \geq \nu|p|^2|q|^2 \quad \forall q\in\mathbb{R}^N,\; p\in\mathbb{R}^n. \tag{4.11}$$

Obviously, (4.10) implies (4.11). We also have the following equivalent conditions.

Lemma 4.1. *Let F be C^2 in ξ. Then condition* (4.10) *or* (4.11) *is equivalent to the following condition respectively:*

$$F(x,s,\eta) \geq F(x,s,\xi) + F_{\xi_i^k}(x,s,\xi)\,(\eta_i^k - \xi_i^k) + \frac{\nu}{2}\,|\eta-\xi|^2, \tag{4.12}$$

$$F(x,s,\xi+q\otimes p) \geq F(x,s,\xi) + F_{\xi_i^k}(x,s,\xi)\,p_i\,q^k + \frac{\nu}{2}\,|p|^2|q|^2 \tag{4.13}$$

for all $x\in\Omega$, s, $q\in\mathbb{R}^N$, ξ, $\eta\in\mathbb{M}^{N\times n}$ and $p\in\mathbb{R}^n$.

Proof. Let $\zeta = \eta-\xi$ and $f(t) = F(x,s,\xi+t\zeta)$. Then, by Taylor's formula,

$$f(1) = f(0) + f'(0) + \int_0^1 (1-t)\,f''(t)\,dt.$$

Note that

$$f'(t) = F_{\xi_i^k}(x,s,\xi+t\zeta)\,\zeta_i^k, \quad f''(t) = F_{\xi_i^k\xi_j^l}(x,s,\xi+t\zeta)\,\zeta_i^k\zeta_j^l.$$

From this and the Taylor formula, inequality (4.12) or (4.13) is equivalent to (4.10) or (4.11), respectively. $\square$

Remark. Interchanging η,ξ in (4.12), we also see that condition (4.10) implies

$$(F_{\eta_i^k}(x,s,\xi) - F_{\xi_i^k}(x,s,\eta))\,(\eta_i^k - \xi_i^k) \geq \nu\,|\eta-\xi|^2; \tag{4.14}$$

that is, $(DF(x,s,\eta) - DF(x,s,\xi))\cdot(\eta-\xi) \geq \nu|\eta-\xi|^2$ for all $x\in\Omega$, $s\in\mathbb{R}^N$, ξ, $\eta\in\mathbb{M}^{N\times n}$.

A function $F(x,s,\xi)$ is said to be **convex** in $\xi\in\mathbb{M}^{N\times n}$ if

$$F(x,s,t\xi+(1-t)\eta) \leq tF(x,s,\xi) + (1-t)F(x,s,\eta)$$

for all x, s, ξ, η and $0 \le t \le 1$. While $F(x,s,\xi)$ is said to be **rank-one convex** in ξ if the function $f(t) = F(x,s,\xi + t\, q \otimes p)$ is convex in $t \in \mathbb{R}^1$ for all x, s, ξ and $q \in \mathbb{R}^N$, $p \in \mathbb{R}^n$. Obviously, a convex function is always rank-one convex.

We easily have the following result.

Lemma 4.2. *Let $F(x,s,\xi)$ be C^2 in ξ. Then the* convexity *of $F(x,s,\xi)$ in ξ is equivalent to* (4.10) *with $\nu = 0$, while the* rank-one convexity *of $F(x,s,\xi)$ in ξ is equivalent to* (4.11) *with $\nu = 0$.*

Remark. Rank-one convexity does not imply convexity. For example, take $n = N \ge 2$, and $F(\xi) = \det \xi$. Then $F(\xi)$ is rank-one convex but not convex in ξ. (**Exercise!**) Later on, we will study other convexity conditions related to the energy functionals given by (4.2).

4.2.4. Structural Conditions. Distributional solutions to the Euler-Lagrange equations (4.4) can be defined as long as $A(x,u,Du)$ and $b(x,u,Du)$ are in $L^1_{loc}(\Omega;\mathbb{R}^N)$, but we need some **structural conditions** on $F(x,s,\xi)$ so that the weak solutions to the BVP

$$-\operatorname{div} A(x,u,Du) + b(x,u,Du) = 0 \quad \text{in } \Omega, \quad u = \varphi \quad \text{on } \partial\Omega, \tag{4.15}$$

can be defined and studied in $W^{1,p}(\Omega;\mathbb{R}^N)$. These conditions are also sufficient for the functional I to be **Gateaux-differentiable** on $W^{1,p}(\Omega;\mathbb{R}^N)$.

Standard Growth Conditions. We assume $F(x,s,\xi)$ is C^1 in (s,ξ) and

$$|F(x,s,\xi)| \le c_1(|\xi|^p + |s|^p) + c_2(x), \qquad c_2 \in L^1(\Omega); \tag{4.16}$$

$$|D_s F(x,s,\xi)| \le c_3(|\xi|^{p-1} + |s|^{p-1}) + c_4(x), \quad c_4 \in L^{\frac{p}{p-1}}(\Omega); \tag{4.17}$$

$$|D_\xi F(x,s,\xi)| \le c_5(|\xi|^{p-1} + |s|^{p-1}) + c_6(x), \quad c_6 \in L^{\frac{p}{p-1}}(\Omega), \tag{4.18}$$

where c_1, c_3, c_5 are constants.

Theorem 4.3. *Under the standard conditions above, the functional $I\colon W^{1,p}(\Omega;\mathbb{R}^N) \to \mathbb{R}$ is Gateaux-differentiable and, for u, $v \in W^{1,p}(\Omega;\mathbb{R}^N)$, the directional derivative $\langle I'(u), v\rangle$ is exactly given by*

$$\langle I'(u), v\rangle = \int_\Omega \big(F_{\xi^k_i}(x,u,Du)\, D_i v^k(x) + F_{s^k}(x,u,Du)\, v^k(x)\big)\, dx \tag{4.19}$$

(as usual, summation notation is used here).

Proof. Given u, $v \in X = W^{1,p}(\Omega;\mathbb{R}^N)$, let $h(t) = I(u+tv)$. Then, by (4.16), h is finite valued, and for $t \ne 0$ we have

$$\frac{h(t)-h(0)}{t} = \int_\Omega \frac{F(x,u+tv,Du+tDv) - F(x,u,Du)}{t}\, dx \equiv \int_\Omega F^t(x)\, dx,$$

where for almost every $x \in \Omega$,

$$F^t(x) = \frac{1}{t}[F(x,u+tv,Du+tDv) - F(x,u,Du)] = \frac{1}{t}\int_0^t \frac{d}{ds} F(x,u+sv,Du+sDv)\, ds.$$

Clearly,

$$\lim_{t\to 0} F^t(x) = F_{\xi^k_i}(x,u,Du)\, D_i v^k(x) + F_{s^k}(x,u,Du)\, v^k(x) \quad a.e.$$

We also write

$$F^t(x) = \frac{1}{t}\int_0^t \Big[F_{\xi^k_i}(x,u+sv,Du+sDv)\, D_i v^k(x) + F_{s^k}(x,u+sv,Du+sDv)\, v^k(x)\Big]\, ds.$$

Using conditions (4.17), (4.18), and Young's inequality: $ab \le \frac{a^p}{p} + \frac{b^q}{q}$, where $\frac{1}{p} + \frac{1}{q} = 1$, we obtain that, for all $0 < |t| \le 1$,

$$|F^t(x)| \le C_1(|Du|^p + |Dv|^p + |u|^p + |v|^p) + C_2(x), \quad C_2 \in L^1(\Omega).$$

Hence, by the Lebesgue dominated convergence theorem,

$$h'(0) = \lim_{t\to 0} \int_\Omega F^t(x)\,dx = \int_\Omega \Big[F_{\xi_i^k}(x,u,Du)\,D_i v^k(x) + F_{s^k}(x,u,Du)\,v^k(x)\Big]\,dx.$$

This proves the theorem. □

Dirichlet Classes. Given $\varphi \in W^{1,p}(\Omega;\mathbb{R}^N)$, we define the **Dirichlet class** of φ to be the set

$$\mathcal{D}_\varphi = \{u \in W^{1,p}(\Omega;\mathbb{R}^N)\,|\,u - \varphi \in W_0^{1,p}(\Omega;\mathbb{R}^N)\}.$$

4.2.5. Weak Solutions of Euler-Lagrange Equations. Under the standard growth conditions, we say $u \in W^{1,p}(\Omega;\mathbb{R}^N)$ is a **weak solution** to the BVP of Euler-Lagrange equations (4.15) if $u \in \mathcal{D}_\varphi$ and

$$\int_\Omega \big(F_{\xi_i^k}(x,u,Du)\,D_i v^k(x) + F_{s^k}(x,u,Du)\,v^k(x)\big)\,dx = 0 \tag{4.20}$$

for all $v \in W_0^{1,p}(\Omega;\mathbb{R}^N)$.

Theorem 4.4. *Under the standard growth conditions above, any minimizer $u \in \mathcal{D}_\varphi$ of*

$$I(u) = \min_{v\in\mathcal{D}_\varphi} I(v)$$

is a weak solution of the BVP for the Euler-Lagrange equation (4.15).

Proof. This follows from Theorem 4.3. □

4.2.6. Minimality and Uniqueness of Weak Solutions. We study the weak solutions of Euler-Lagrange equations under the hypotheses of *convexity* and certain *growth* conditions. For this purpose, we consider a simple case where $F(x,s,\xi) = F(x,\xi)$ satisfies, for some $1 < p < \infty$,

$$|F_{\xi_i^k}(x,\xi)| \le \mu\,(\chi(x) + |\xi|^{p-1}) \quad \forall\, x \in \Omega,\ \xi \in \mathbb{M}^{N\times n}, \tag{4.21}$$

where $\mu > 0$ is a constant and $\chi \in L^{\frac{p}{p-1}}(\Omega)$ is some function. Let

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

Theorem 4.5. *Let $F(x,\xi)$ be C^2 and convex in ξ. Let $u \in W^{1,p}(\Omega;\mathbb{R}^N)$ be a weak solution of the* Euler-Lagrange *equation of I and $I(u) < \infty$. Then u must be a minimizer of I in the Dirichlet class $\mathcal{D}_u$ of $W^{1,p}(\Omega;\mathbb{R}^N)$. Furthermore, if F satisfies the* Legendre condition (4.10) *(with $\nu > 0$), then u is the unique minimizer of I in $\mathcal{D}_u$.*

Proof. Since u is a weak solution of the *Euler-Lagrange* equation of I, it follows that

$$\int_\Omega F_{\xi_i^k}(x, Du(x))\,D_i v^k\,dx = 0 \tag{4.22}$$

for all $v \in W_0^{1,p}(\Omega;\mathbb{R}^N)$. The growth condition (4.21) implies $F_{\xi_i^k}(x,Du) \in L^{\frac{p}{p-1}}(\Omega)$. Now let $v \in W^{1,p}(\Omega;\mathbb{R}^N)$ with $v-u \in W_0^{1,p}(\Omega;\mathbb{R}^N)$. By the convexity condition, we have

$$F(x,\eta) \geq F(x,\xi) + F_{\xi_i^k}(x,\xi)\,(\eta_i^k - \xi_i^k) + \frac{\nu}{2}|\eta-\xi|^2, \quad \forall\,\xi,\,\eta, \tag{4.23}$$

where $\nu = 0$ if F is only convex in ξ and $\nu > 0$ if F satisfies the Legendre condition. This implies

$$\begin{aligned}\int_\Omega F(x,Dv)\,dx \;\; &\geq \;\; \int_\Omega F(x,Du)\,dx + \int_\Omega F_{\xi_i^k}(x,Du)\,D_i(v^k-u^k)\,dx \\ &+ \;\; \frac{\nu}{2}\int_\Omega |Du-Dv|^2\,dx.\end{aligned}$$

Since u is a weak solution, we have

$$\int_\Omega F_{\xi_i^k}(x,Du)\,D_i(v^k-u^k)\,dx = 0.$$

Therefore, it follows that

$$I(v) \geq I(u) + \frac{\nu}{2}\int_\Omega |Du-Dv|^2\,dx \geq I(u) \tag{4.24}$$

for all $v \in W^{1,p}(\Omega;\mathbb{R}^N)$ with $v-u \in W_0^{1,p}(\Omega;\mathbb{R}^N)$. This shows that u is a minimizer of I in the Dirichlet class $\mathcal{D}_u$.
Now assume $\nu > 0$ and $v \in \mathcal{D}_u$ is another such minimizer of I. Then from (4.24) we easily obtain $Du = Dv$ in Ω and hence $v \equiv u$ since $u - v \in W_0^{1,p}(\Omega;\mathbb{R}^N)$. The proof is now completed. □

4.2.7. Regularity. We now study functional $I(u)$ scalar functions u, where

$$I(u) = \int_\Omega F(Du(x))\,dx$$

for a smooth convex function $F(\xi)$ satisfying the following condition: for some constants C and $\theta > 0$,

$$|D^2F(\xi)| \leq C, \quad F_{\xi_i\xi_j}(\xi)p_ip_j \geq \theta|p|^2, \quad \forall\;\xi,\,p \in \mathbb{R}^n. \tag{4.25}$$

Let $f \in L^2(\Omega)$. Recall that a function $u \in H^1(\Omega)$ is called a **weak solution** to the nonlinear PDE (the Euler-Lagrange equation for $I(u)$):

$$-\sum_{i=1}^n (F_{\xi_i}(Du))_{x_i} = f \quad \text{in } \Omega \tag{4.26}$$

if

$$\int_\Omega \sum_{i=1}^n F_{\xi_i}(Du)v_{x_i}\,dx = \int_\Omega fv\,dx \quad \forall\; v \in H_0^1(\Omega).$$

Theorem 4.6. **(H^2-Regularity)** *Let $u \in H^1(\Omega)$ be a weak solution to* (4.26). *Then $u \in H^2_{loc}(\Omega)$. If in addition $u \in H^1_0(\Omega)$ and $\partial\Omega$ is C^2, then $u \in H^2(\Omega)$ and*

$$\|u\|_{H^2(\Omega)} \leq C\|f\|_{L^2(\Omega)}.$$

Proof. The proof of for the interior regularity will follow largely from the same proof for linear equations. However, we only obtain an estimate like

$$\|u\|_{H^2(\Omega')} \le C(\Omega')(\|u\|_{H^1(\Omega)} + \|f\|_{L^2(\Omega)}) \quad \forall\, \Omega' \subset\subset \Omega. \tag{4.27}$$

If $u \in H_0^1(\Omega)$, we can estimate $\|u\|_{H_0^1(\Omega)}$ by $\|f\|_{L^2(\Omega)}$. In fact, by (4.14) and (4.25), it follows that $(DF(\xi) - DF(0)) \cdot \xi \ge \theta|\xi|^2$ for all $\xi \in \mathbb{R}^n$. So

$$DF(\xi) \cdot \xi \ge DF(0)) \cdot \xi + \theta|\xi|^2, \quad \forall\, \xi \in \mathbb{R}^n.$$

Using $v = u \in H_0^1(\Omega)$ as test function, we have

$$\int_\Omega fu = \int_\Omega DF(Du) \cdot Du \ge \int_\Omega DF(0) \cdot Du + \int_\Omega \theta|Du|^2 = \theta \int_\Omega |Du|^2.$$

By Poincaré's inequality, $\|u\|_{L^2(\Omega)} \le c\|Du\|_{L^2(\Omega)}$ and hence, using Cauchy's inequality with ϵ, we obtain

$$\|Du\|_{L^2(\Omega)} \le C\|f\|_{L^2(\Omega)}.$$

Then the global estimate can also be proved largely following the idea for global regularity of linear equations. □

Remarks. (1) Assume f and F are C^∞. Assume $u \in H^1(\Omega)$ is a weak solution to (4.26). Then $u \in H^2_{loc}(\Omega)$ and, using $v = w_{x_k}$ as test function, where $w \in C_0^\infty(\Omega)$, we obtain

$$\int_\Omega \sum_{i,j=1}^n F_{\xi_i\xi_j}(Du) u_{x_j x_k} w_{x_i}\, dx = \int_\Omega f_{x_k} w\, dx \quad \forall\, w \in C_0^\infty(\Omega).$$

Let $a_{ij}(x) = F_{\xi_i\xi_j}(Du(x))$, $\tilde{u} = u_{x_k}$ and $\tilde{f} = f_{x_k}$. This identity implies that, for any $\Omega' \subset\subset \Omega$, $\tilde{u} \in H^1(\Omega')$ is a weak solution of the equation

$$-D_i(a_{ij}(x) D_j \tilde{u}) = \tilde{f} \quad \text{in } \Omega'. \tag{4.28}$$

However, the coefficient $(a_{ij}(x))$, which depends on solution u and satisfies the strict ellipticity condition, is only in $L^\infty(\Omega')$; therefore, we cannot use the regularity theorem (e.g., Theorem 3.16) of linear equations (which needs $a_{ij} \in C^1$).

(2) There is a well-known **DeGiorgi-Nash-Moser theorem** for equations of the form (4.28) which asserts that if $\tilde{f} \in C^\infty$ then any weak solution $\tilde{u} \in H^1(\Omega')$ must be in $C^\alpha(\Omega')$ for some $0 < \alpha < 1$. Hence $Du \in C^\alpha(\Omega')$, which implies $a_{ij} = F_{\xi_i\xi_j}(Du) \in C^\alpha$. Then, a classical **Schauder's estimate** for equation (4.28) will imply $\tilde{u} \in C^{1,\alpha}$; such an estimate can also be established by using the technique of **Cappanato spaces**. Therefore, $a_{ij}(x)$ defined is in fact in $C^{1,\alpha}$. This bootstrap argument will then show that if f and F are C^∞ then any weak solution $u \in H^1(\Omega)$ must be C^∞. For details, see Giaquinta [**13**] and Gilbarg & Trudinger [**14**].

4.3. Direct Method for Minimization

4.3.1. Weak Lower Semicontinuity. Assume $1 < p < \infty$. Let $X = W^{1,p}(\Omega; \mathbb{R}^N)$. A set $\mathcal{C} \subset X$ is called **weakly closed** if $\{u_\nu\} \subset \mathcal{C}$, $u_\nu \rightharpoonup u$ implies $u \in \mathcal{C}$. For instance, all Dirichlet classes $\mathcal{D}_\varphi$ are weakly closed.

A functional $I \colon X \to \bar{\mathbb{R}}$ is called **weakly lower semicontinuous**(**w.l.s.c.**) on X if for every $u_0 \in X$ and every sequence $\{u_\nu\}$ weakly convergent to u_0 in X it follows that

$$I(u_0) \le \liminf_{\nu\to\infty} I(u_\nu).$$

I is called **weakly coercive** on a (unbounded) set C in X if $I(u) \to \infty$ as $\|u\| \to \infty$ on C.

4.3.2. Direct Method in the Calculus of Variations. The following theorem is a special case of the generalized Weierstrass theorem (see Theorem 1.38).

Theorem 4.7. (Existence of Minimizers) *Assume* $1 < p < \infty$. *Let* $I: \mathcal{C} \subseteq X = W^{1,p}(\Omega; \mathbb{R}^N) \to \mathbb{R}$ *be w.l.s.c. and weakly coercive on a nonempty weakly closed set* $\mathcal{C}$ *in* X. *Then there is at least one* $u_0 \in \mathcal{C}$ *such that* $I(u_0) = \inf_{u\in\mathcal{C}} I(u)$. *In this case we say* $u_0 \in \mathcal{C}$ *is a* **minimizer** *of* I *on* $\mathcal{C}$.

Proof. This theorem is a special case of Theorem 1.38 under assumption (ii) there; the proof involves the basic ideas of what is known as the **direct method** of calculus of variations. We explain this method by giving a proof of this theorem. First of all, by the weak coercivity, one can easily show that $\inf_{u\in\mathcal{C}} I(u)$ is finite; so take a sequence $\{u_\nu\}$, called a **minimizing sequence**, such that

$$\lim_{\nu\to\infty} I(u_\nu) = \inf_{u\in\mathcal{C}} I(u).$$

Then the weak coercivity condition implies that $\{u_\nu\}$ must be bounded in X. Since $1 < p < \infty$ and thus $X = W^{1,p}(\Omega; \mathbb{R}^N)$ is reflexive, there exists a subsequence of $\{u_\nu\}$, denoted by $\{u_{\nu_j}\}$, and $u_0 \in X$ such that $u_{\nu_j} \rightharpoonup u_0$ *weakly* in X. The weak closedness of $\mathcal{C}$ implies $u_0 \in \mathcal{C}$. Now the w.l.s.c. of I implies

$$I(u_0) \le \liminf_{j\to\infty} I(u_{\nu_j}) = \inf_{u\in\mathcal{C}} I(u).$$

This implies u_0 is a minimizer of I over $\mathcal{C}$. □

4.3.3. An Example: p-Laplace Equations. As an example, we consider the BVP for nonlinear **p-Laplace equations** ($p \ge 2$):

$$-\sum_{i=1}^{n} D_i(|\nabla u|^{p-2} D_i u) + f(x,u) = 0 \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega, \tag{4.29}$$

where $f : \Omega \times \mathbb{R} \to \mathbb{R}$ be a given function satisfying the **Carathéodory property,** i.e., for every $s \in \mathbb{R}$, $f(x,s)$ (as a function of x) is measurable on Ω, and for almost all $x \in \Omega$, $f(x,s)$ (as a function of s) is continuous on $\mathbb{R}$. Note that when $p = 2$, (4.29) becomes the **semilinear** elliptic problem.

Define

$$F(x,s) = \int_0^s f(x,t)\,dt$$

and the functional I on $X = W_0^{1,p}(\Omega)$ by

$$I(u) = \int_\Omega \left[\frac{1}{p}|\nabla u|^p + F(x,u)\right] dx.$$

Assume the following growth conditions are satisfied:

$$F(x,s) \ge -c_1|s| - c_2(x), \tag{4.30}$$

$$|F(x,s)| \le c_3(x) + c_4|s|^p, \tag{4.31}$$

$$|f(x,s)| \le c_5(x) + c_6|s|^{p-1} \tag{4.32}$$

for all $x \in \Omega$ and $s \in \mathbb{R}$, where c_1, c_4, c_6 are nonnegative constants, and c_2, $c_3 \in L^1(\Omega)$, and $c_5 \in L^{\frac{p}{p-1}}(\Omega)$ are given functions. Note that (4.32) implies (4.31).

Theorem 4.8. *Under these conditions, the functional* I *has a minimizer on* $X = W_0^{1,p}(\Omega)$ *and hence* (4.29) *has a weak solution.*

Proof. We only need to check I is w.l.s.c and weakly coercive on X. We write

$$I(u) = I_1(u) + I_2(u) = \frac{1}{p}\int_\Omega |\nabla u|^p\,dx + \int_\Omega F(x,u)\,dx.$$

Note that $pI_1(u) = \|u\|^p_{1,p,0}$ on X which is w.l.s.c. And since the embedding $X \subset L^p(\Omega)$ is always *compact*, by (4.31), we can show that I_2 is in fact continuous under the weak convergence. Hence I is w.l.s.c. on X. By (4.30), we have

$$I(u) \geq \frac{1}{p}\|u\|^p_{1,p,0} - c_1\|u\|_{L^1(\Omega)} - C \geq \frac{1}{p}\|u\|^p_{1,p,0} - c\|u\|_{1,p,0} - C,$$

and hence $I(u) \to \infty$ if $\|u\|_X = \|u\|_{1,p,0} \to \infty$. This proves the weak coercivity of I. The result follows then from Theorem 4.7. $\square$

EXAMPLE 4.9. Let $n \geq 3$ and $2 \leq p < n$. Show the theorem is valid if (4.32) above is replaced by

$$|f(x,s)| \leq a(x) + b\,|s|^q,$$

where $b \geq 0$ is a constant, $a \in L^{\frac{q+1}{q}}(\Omega)$ and $1 \leq q < p^* - 1$.

4.4. Minimization with Constraints

In some cases, we need to minimize functional I under certain constraints. If I is a multiple integral functional on $X = W^{1,p}(\Omega;\mathbb{R}^N)$ defined before, there may be constraints given in terms of one of the following:

$$J(u) = 0, \quad \text{where } J(u) = \textstyle\int_\Omega G(x,u)\,dx. \tag{4.33}$$

$$h(u(x)) = 0, \quad \forall\ a.e.\ x \in \Omega. \tag{4.34}$$

$$M(Du(x)) = 0, \quad \forall\ a.e.\ x \in \Omega. \tag{4.35}$$

All these lead to some PDEs involving the **Lagrange multipliers**. For different types of constraints, the Lagrange multiplier comes in significantly different ways.

4.4.1. Nonlinear Eigenvalue Problems.

The following theorem was proved in the Preliminaries (see Theorem 1.43).

Theorem 4.10. *Let X be a Banach space. Let $f, g : X \to \mathbb{R}$ be of class C^1 and $g(u_0) = c$. Assume u_0 is a local extremum of f with respect to the constraint $g(u) = c$. Then either $g'(u_0)v = 0$ for all $v \in X$, or there exists $\lambda \in \mathbb{R}$ such that $f'(u_0)v = \lambda g'(u_0)v$ for all $v \in X$; that is, u_0 is a critical point of $f - \lambda g$.*

If $u_0 \neq 0$ then the corresponding λ is called an eigenvalue for the **nonlinear eigenvalue problem**: $f'(u) = \lambda g'(u)$, and u_0 is the corresponding eigenfunction.

We have following applications.

Theorem 4.11. *Let $k(x), l(x) \in C(\bar\Omega)$ with $l(x) \geq \alpha > 0$ on $\bar\Omega$. Then, for each $R \in (0,\infty)$, the problem*

$$\Delta u + k(x)u + \lambda l(x)|u|^{\tau-1}u = 0 \ \text{ in } \ \Omega, \ \ u|_{\partial\Omega} = 0 \qquad (0 < \tau < \frac{n+2}{n-2}) \tag{4.36}$$

has a (weak) solution pair (u_R, λ_R) with $\frac{1}{\tau+1}\int_\Omega l(x)|u_R|^{\tau+1}dx = R$. Furthermore, one can have $u_R \geq 0$ on Ω.

Proof. Define the functionals
$$f(u) = \frac{1}{2}\int_\Omega (|\nabla u|^2 - k(x)u^2)\,dx, \quad g(u) = \frac{1}{\tau+1}\int_\Omega l(x)|u|^{\tau+1}\,dx.$$
As before, we easily see that f is w.l.s.c., and g is weakly continuous on $H_0^1(\Omega)$ (it is here that the assumption $\tau < \frac{n+2}{n-2}$ is used to guarantee that the embedding $H_0^1(\Omega) \subset L^{\tau+1}(\Omega)$ be compact). Note that for each $R > 0$ the set $\mathcal{C} = \{u \in H_0^1(\Omega) : g(u) = R\}$ is nonempty (Why?) and weakly closed. Now we show that f is weakly coercive on $\mathcal{C}$. First of all, by Hölder's inequality we have
$$\begin{aligned}\left(\int_\Omega |u|\,dx\right)^{\tau+1} &\le |\Omega|^\tau \int_\Omega |u|^{\tau+1}\,dx \\ &\le \frac{|\Omega|^\tau}{\alpha}\int_\Omega l(x)|u|^{\tau+1}\,dx = Cg(u) = CR.\end{aligned}$$
Hence $\|u\|_1$ is uniformly bounded on $\mathcal{C}$. (If $\tau \ge 1$ then we can also see that $\|u\|_2$ is bounded on $\mathcal{C}$ and in the following we do not need Ehrling's inequality.) Furthermore, by **Ehrling's inequality** (Theorem 1.46), for each $\varepsilon > 0$, there is an absolute constant $c(\varepsilon) > 0$ such that
$$\|u\|_2^2 \le \varepsilon\|u\|_{1,2,0}^2 + c(\varepsilon)\|u\|_1^2 \quad \text{for all} \quad u \in H_0^1(\Omega).$$
Thus
$$f(u) \ge \frac{1}{2}(1 - \varepsilon \max|k|)\|u\|_{1,2,0}^2 - \frac{1}{2}c(\varepsilon)\max|k|\|u\|_1^2.$$
If we choose ε small enough so that $1 - \varepsilon\max|k| > 0$, then $f(u) \to \infty$ as $\|u\|_{1,2,0} \to \infty$ for $u \in \mathcal{C}$. Therefore, we can apply the direct method to obtain a minimizer $u_R \in \mathcal{C}$ of f over set $\mathcal{C}$. Hence u_R is a minimizer of f with respect to the constraint $g(u) = R$. Furthermore, since $f(u) = f(|u|)$ and $g(u) = g(|u|)$ for all $u \in H_0^1(\Omega)$, $|u_R|$ will also be a minimizer of f with respect to $g(u) = R$; hence we can assume $u_R \ge 0$.

It is easily shown that f and g are both G-differentiable on $H = H_0^1(\Omega)$ and
$$\langle f'(u), v\rangle = \int_\Omega (\nabla u\cdot\nabla v - k(x)uv)\,dx, \quad \langle g'(u), v\rangle = \int_\Omega l(x)|u|^{\tau-1}uv\,dx$$
for all $u, v \in H$. Hence
$$\|f'(u) - f'(w)\|_{H^*} = \sup_{v\in H, \|v\|_H \le 1} \langle f'(u) - f'(w), v\rangle \le C\|u - w\|_H$$
and
$$\begin{aligned}\|g'(u) - g'(w)\|_{H^*} &= \sup_{v\in H, \|v\|_H\le 1}\int_\Omega l(x)(|u|^{\tau-1}u - |w|^{\tau-1}w)v\,dx \\ \le (\max l(x))\||u|^{\tau-1}u - |w|^{\tau-1}w\|_{\frac{\tau+1}{\tau}}\|v\|_{\tau+1} &\le C\||u|^{\tau-1}u - |w|^{\tau-1}w\|_{\frac{\tau+1}{\tau}}.\end{aligned}$$
Since **Nemystkii operator** $N(u) = |u|^{\tau-1}u\colon L^{\tau+1}(\Omega) \to L^{\frac{\tau+1}{\tau}}(\Omega)$ is bounded and continuous and the imbedding $H \to L^{\tau+1}(\Omega)$ is continuous, it follows that both f' and g' are continuous from H to H^* (in fact, g' is compact); hence both f and g are C^1 on $H_0^1(\Omega)$ and that $g'(u) = 0$ iff $u = 0$. Since $g(u_R) = R > 0$, clearly $u_R \ne 0$ and hence $g'(u_R) \ne 0$. Therefore, by Theorem 4.10, we conclude that there exists a real number λ_R such that $f'(u_R) = \lambda_R g'(u_R)$; hence (u_R, λ_R) is weak solution of (4.36) with $g(u_R) = R$. Moreover, if $\beta_R = f(u_R) = \min_{u\in\mathcal{C}} f(u)$, then it is easily seen that $\lambda_R = \frac{2\beta_R}{(\tau+1)R}$. □

Corollary 4.12. *For $0 < \tau < (n+2)/(n-2)$ and $\tau \ne 1$, there exists a nontrivial weak solution of*

$$\Delta u + |u|^{\tau-1}u = 0 \;\; \text{in} \;\; \Omega, \;\; u|_{\partial\Omega} = 0. \tag{4.37}$$

Proof. By Theorem 4.11, there exist $u_R \neq 0,\ \lambda_R > 0$ such that

$$\Delta u_R + \lambda_R |u_R|^{\tau-1} u_R = 0 \ \text{ in } \ \Omega, \quad u_R|_{\partial\Omega} = 0.$$

Set $u_R = kv$ ($k > 0$ is to be determined). Then $k\Delta v + \lambda_R k^\tau |v|^{\tau-1} v = 0$. Since $\tau \neq 1$, we can choose $k > 0$ to satisfy $\lambda_R k^{\tau-1} = 1$; hence, it follows that

$$\Delta v + |v|^{\tau-1} v = 0 \ \text{ in } \ \Omega, \quad v|_{\partial\Omega} = 0$$

has a nontrivial weak solution in $H_0^1(\Omega)$. □

Remarks. (i) Another method to show the existence of nontrivial weak solution to (4.37) is given later by the Mountain Pass Theorem.

(ii) The **nonexistence** of nontrivial *classical* solutions to problem (4.37) when $\tau \geq \frac{n+2}{n-2}$ will be studied later for certain domains with simple topological property.

Example 4.13. For domains like annulus, problem (4.37) always has nontrivial solutions for all $\tau > 1$. For example, let Ω be the annulus $0 < a < r < b,\ \ r = |x|$ and suppose $\tau > 1$. Prove that the BVP

$$\Delta u + |u|^{\tau-1} u = 0 \ \text{ in } \ \Omega, \quad u|_{\partial\Omega} = 0$$

has a nontrivial solution.

Proof. Minimize the functional

$$f(u) = \int_a^b (u')^2 r^{n-1} dr, \quad r = |x|$$

over all u in the set

$$C = \left\{ u \in H_0^1(a,b) \,\middle|\, \int_a^b |u|^{\tau+1} r^{n-1} dr = 1 \right\}.$$

□

4.4.2. Harmonic Maps and Liquid Crystals. We now consider the Dirichlet energy

$$I(u) = \frac{1}{2} \int_\Omega |Du|^2 \, dx$$

for $u \in H^1(\Omega; \mathbb{R}^N)$ with point-wise constraint $|u(x)| = 1$ for almost every $x \in \Omega$. Let

$$\mathcal{C} = \{ u \in \mathcal{D}_\varphi \,|\, |u(x)| = 1 \ a.e. \},$$

where $\mathcal{D}_\varphi$ is a Dirichlet class in $H^1(\Omega; \mathbb{R}^N)$. We assume $\mathcal{C}$ is non-empty. Then $\mathcal{C}$ is weakly closed in $H^1(\Omega; \mathbb{R}^N)$. We have the following result.

Theorem 4.14. *There exists $u \in \mathcal{C}$ satisfying*

$$I(u) = \min_{v \in \mathcal{C}} I(v).$$

Moreover, u satisfies

$$\int_\Omega Du \cdot Dv \, dx = \int_\Omega |Du|^2 u \cdot v \, dx \tag{4.38}$$

for each $v \in H_0^1(\Omega; \mathbb{R}^N) \cap L^\infty(\Omega; \mathbb{R}^N)$.

Remark. In this case, we see u is a *weak* solution to the **harmonic map equation**:

$$-\Delta u = |Du|^2 u \quad \text{in } \Omega.$$

The "Lagrange multiplier" in this case is the function $\lambda = |Du|^2$, instead of a constant.

Proof. The existence of minimizers follows by the direct method as above. Given any $v \in H_0^1(\Omega;\mathbb{R}^N) \cap L^\infty(\Omega;\mathbb{R}^N)$, let ε be such that $|\varepsilon|\|v\|_{L^\infty(\Omega)} \le \frac{1}{2}$. Define

$$w_\varepsilon(x) = \frac{u(x)+\varepsilon v(x)}{|u(x)+\varepsilon v(x)|}, \quad h(\varepsilon) = I(w_\varepsilon).$$

Note that $w_\varepsilon \in \mathcal{C}$ and $h(0) = I(u) = \min_{\mathcal{C}} I \le h(\varepsilon)$ for sufficiently small ε; hence, $h'(0) = 0$. Note that

$$\frac{\partial w_\varepsilon}{\partial \varepsilon} = \frac{|u+\varepsilon v|^2 v - (u\cdot v + \varepsilon|v|^2)(u+\varepsilon v)}{|u+\varepsilon v|^3}; \quad \frac{\partial w_\varepsilon}{\partial \varepsilon}|_{\varepsilon=0} = v - (u\cdot v)u.$$

Hence

$$h'(0) = \int_\Omega Du\cdot D(v-(u\cdot v)u)\,dx = \int_\Omega (Du\cdot Dv - |Du|^2 u\cdot v)\,dx,$$

where we have used $Du\cdot D((u\cdot v)u) = (u\cdot v)|Du|^2$ due to the fact that $(Du)^T u = 0$ (which follows from $|u|=1$). This yields the weak form of harmonic map equation (4.38). ☐

Liquid Crystals. Let $n = N = 3$. Then the harmonic map Dirichlet energy can be considered as a special case of the **Oseen-Frank energy** for **liquid crystals** defined earlier:

$$\begin{aligned} I(u) = \int_\Omega F(u,Du)\,dx \;&=\; \int_\Omega (\kappa_1(\operatorname{div} u)^2 + \kappa_2(u\cdot\operatorname{curl} u)^2 + \kappa_3(u\times\operatorname{curl} u)^2)\,dx \\ &+ \int_\Omega \kappa_4(\operatorname{tr}((Du)^2) - (\operatorname{div} u)^2)\,dx, \end{aligned}$$

where Ω is a bounded domain in $\mathbb{R}^3$, $u\colon \Omega\to\mathbb{R}^3$, $\operatorname{div} u = \operatorname{tr} Du$ is the divergence of u and $\operatorname{curl} u = \nabla\times u$ denotes the curl vector of u in $\mathbb{R}^3$. In the first part of the total energy $I(u)$, the κ_1-term represents the **splay energy**, κ_2-term represents the **twist energy** and κ_3-term represents the **bending energy**, corresponding to the various deformations of the **nematic director** u with $|u(x)| = 1$; the κ_4-term is a **null-Lagrangian**, depending only on the boundary data of u (see Example 5.30 in Chapter 5).

From the algebraic relation $|Du - (Du)^T|^2 = 2|\operatorname{curl} u|^2$ one easily sees that

$$|Du|^2 = |\operatorname{curl} u|^2 + \operatorname{tr}((Du)^2).$$

Furthermore, since $|u(x)| = 1$, it easily follows that

$$|\operatorname{curl} u|^2 = (u\cdot\operatorname{curl} u)^2 + |u\times\operatorname{curl} u|^2.$$

If $\kappa = \min\{\kappa_1,\kappa_2,\kappa_3\}$ then it easily follows that

$$I(u) \ge \kappa\int_\Omega |Du|^2\,dx + (\kappa_4-\kappa)\int_\Omega [\operatorname{tr}((Du)^2) - (\operatorname{div} u)^2]\,dx. \tag{4.39}$$

If $\kappa > 0$, then it can be shown that the first part of $F(u,\xi)$ is convex and quadratic in ξ, and hence the first part of the energy I is w.l.s.c. on $H^1(\Omega;\mathbb{R}^3)$ by **Tonelli's theorem** (Theorem 5.1). By (4.39), one can easily obtain that

$$\int_\Omega |Du(x)|^2\,dx \le c_0\, I(u) + C_\varphi, \quad \forall\, u\in\mathcal{D}_\varphi \subset H^1(\Omega;\mathbb{R}^3).$$

Therefore, by the direct method, we have established the following existence result for minimizers.

Theorem 4.15. *Let $\kappa_i > 0$ for $i = 1, 2, 3$. If $\mathcal{C} = \{u \in \mathcal{D}_\varphi \mid |u(x)| = 1 \text{ a.e.}\} \neq \emptyset$, then there exists at least one $u \in \mathcal{C}$ such that*

$$I(u) = \min_{v \in \mathcal{C}} I(v).$$

Remark. If all four κ_i are equal to $\frac{1}{2}$, then $I(u)$ reduces to the Dirichlet integral for **harmonic maps**: $I(u) = \frac{1}{2}\int_\Omega |Du|^2\,dx$.

4.4.3. Stokes' Problem. Let $\Omega \subset \mathbb{R}^3$ be open, bounded and simply connected. Given $f \in L^2(\Omega; \mathbb{R}^3)$, the **Stokes' problem**

$$-\Delta u = f - \nabla p, \quad \operatorname{div} u = 0 \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega,$$

can be solved by minimizing the functional

$$I(u) = \int_\Omega (\frac{1}{2}|Du|^2 - f \cdot u)\,dx$$

over the set

$$\mathcal{C} = \{u \in H_0^1(\Omega; \mathbb{R}^3) \mid \operatorname{div} u = 0 \text{ in } \Omega\}.$$

The function p in the equation is the corresponding "Lagrange multiplier", known as the pressure. Pressure does not appear in the variational problem, but arises due to the constraint $\operatorname{div} u = 0$.

4.4.4. Incompressible Elasticity. Suppose $u \colon \Omega \to \mathbb{R}^3$ represents the displacement of an elastic body occupying the domain $\Omega \subset \mathbb{R}^3$ before deformation. Assume the body is incompressible, which means

$$\det Du(x) = 1 \quad \forall\, x \in \Omega.$$

The stored energy is given by an integral functional

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

The suitable space to work in this case is $W^{1,p}(\Omega; \mathbb{R}^3)$ with $p \geq 3$. For further details, see next chapter.

4.4.5. A Nonlocal Problem in Ferromagnetism. We study a minimization problem in *ferromagnetism* which involves a constraint and a nonlocal term. Certain simplification of the Landau-Lifshitz theory leads to a model of the total micromagnetic energy $I(\mathbf{m})$ given by

$$I(\mathbf{m}) = \int_\Omega \varphi(\mathbf{m}(x))\,dx + \frac{1}{2}\int_{\mathbb{R}^n} |\nabla u(x)|^2\,dx, \tag{4.40}$$

where Ω is a bounded domain with C^1 boundary $\partial\Omega$, $\mathbf{m} \in L^\infty(\Omega; \mathbb{R}^n)$ with $|\mathbf{m}(x)| = 1$ a.e. $x \in \Omega$ (that is, $\mathbf{m} \in L^\infty(\Omega; S^{n-1})$), representing the magnetization of a ferromagnetic material occupying the domain Ω, φ is a given function representing the **anisotropy** of the material, and u is the nonlocal **stray** energy potential determined by $\mathbf{m}$ over the whole space $\mathbb{R}^n$ by the **simplified Maxwell equation**:

$$\operatorname{div}(-\nabla u + \mathbf{m}\chi_\Omega) = 0 \quad \text{in } \mathbb{R}^n, \tag{4.41}$$

where χ_Ω is the characteristic function of Ω, and $u \in H^1_{loc}(\mathbb{R}^n)$ with $\nabla u \in L^2(\mathbb{R}^n; \mathbb{R}^n)$.

The equation (4.41) is understood in the weak sense:

$$\int_{\mathbb{R}^n} \nabla u \cdot \nabla \zeta \, dx = \int_\Omega \mathbf{m} \cdot \nabla \zeta \, dx \quad \forall \, \zeta \in H_0^1(\mathbb{R}^n).$$

Note that any two weak solutions of (4.41) can only differ by a constant; therefore ∇u is uniquely determined by $\mathbf{m}$.

Let B be a fixed open ball containing $\bar{\Omega}$ in $\mathbb{R}^n$. We have the following result.

Theorem 4.16. *For each $\mathbf{m} \in L^2(\Omega; \mathbb{R}^n)$, there exists a unique weak solution $u = T\mathbf{m} \in H^1_{loc}(\mathbb{R}^n)$ to (4.41) such that $\int_B u(x)dx = 0$. Moreover $T\colon L^2(\Omega;\mathbb{R}^n) \to H^1(B)$ is linear and bounded with*

$$\|T\mathbf{m}\|_{H^1(B)} \le C \, \|\mathbf{m}\|_{L^2(\Omega;\mathbb{R}^n)}, \quad \forall \, \mathbf{m} \in L^2(\Omega; \mathbb{R}^n).$$

Proof. Let

$$X = \{u \in H^1_{loc}(\mathbb{R}^n) \mid \quad \nabla u \in L^2(\mathbb{R}^n; \mathbb{R}^n), \quad \int_B u(x) \, dx = 0\}.$$

Given $\mathbf{m} \in L^2(\Omega; \mathbb{R}^n)$, define a functional $J\colon X \to \mathbb{R}$ by

$$J(u) = \frac{1}{2} \int_{\mathbb{R}^n} |\nabla u|^2 \, dx - \int_\Omega \mathbf{m} \cdot \nabla u \, dx, \quad u \in X.$$

We solve the minimization problem: $\inf_{u \in X} J(u)$. Note that X is simply a nonempty set of functions and has no topology defined. But we can always take a minimizing sequence $u_k \in X$ such that

$$\lim_{k\to\infty} J(u_k) = \inf_{u\in X} J(u) < \infty.$$

By Cauchy's inequality with ϵ, it follows that

$$J(u) \ge \frac{1}{2} \int_{\mathbb{R}^n} |\nabla u|^2 \, dx - \epsilon \int_{\mathbb{R}^n} |\nabla u|^2 \, dx - C_\epsilon \int_\Omega |\mathbf{m}|^2 \, dx.$$

Taking $\epsilon = \frac{1}{4}$ yields

$$J(u) \ge \frac{1}{4} \int_{\mathbb{R}^n} |\nabla u|^2 \, dx - C, \quad \forall \, u \in X.$$

Therefore $\{\nabla u_k\}$ is bounded in $L^2(\mathbb{R}^n; \mathbb{R}^n)$ and hence, via a subsequence, converges weakly to some $F \in L^2(\mathbb{R}^n; \mathbb{R}^n)$; this F can be written as $F = \nabla \bar{u}$ for a unique $u \in X$. (The last fact needs a little more analysis; see Theorem 4.20 below!) We then see that $\bar{u} \in X$ is a minimizer of J over X. It also satisfies the Euler-Lagrange equation

$$\int_{\mathbb{R}^n} \nabla \bar{u} \cdot \nabla \zeta = \int_\Omega \mathbf{m} \cdot \nabla \zeta \quad \forall \zeta \in X. \tag{4.42}$$

This equation also holds for all $\zeta \in H_0^1(\mathbb{R}^n)$. From this equation, the uniqueness of minimizers and the linear dependence of $\bar{u}$ on $\mathbf{m}$ will follow. Define $\bar{u} = T\mathbf{m}$. Furthermore, using $\zeta = \bar{u} \in X$ in (4.42), we also have

$$\int_{\mathbb{R}^n} |\nabla \bar{u}|^2 \, dx = \int_\Omega \mathbf{m} \cdot \nabla \bar{u} \, dx \le \|\mathbf{m}\|_{L^2(\Omega;\mathbb{R}^n)} \|\nabla \bar{u}\|_{L^2(B;\mathbb{R}^n)}.$$

Hence

$$\|\nabla \bar{u}\|_{L^2(B;\mathbb{R}^n)} \le \|\mathbf{m}\|_{L^2(\Omega;\mathbb{R}^n)}.$$

Since $\int_B \bar{u} \, dx = 0$, by Poincaré's inequality,

$$\|\bar{u}\|_{H^1(B)} \le C \, \|\mathbf{m}\|_{L^2(\Omega;\mathbb{R}^n)}.$$

The theorem is proved. $\square$

From (4.42), we also see that

$$I(\mathbf{m}) = \int_\Omega \varphi(\mathbf{m})\,dx + \frac{1}{2}\int_\Omega \mathbf{m}\cdot\nabla\bar{u}\,dx, \tag{4.43}$$

where $\bar{u} = T\mathbf{m}$ is defined in the theorem above.

The following result is a special case of some general theorem, which can be proved easily.

Lemma 4.17. *Let $T\colon L^2(\Omega;\mathbb{R}^n) \to H^1(B)$ be defined above. Then $T\mathbf{m}_k \rightharpoonup T\mathbf{m}$ weakly in $H^1(B)$ whenever $\mathbf{m}_k \rightharpoonup \mathbf{m}$ weakly in $L^2(\Omega;\mathbb{R}^n)$.*

Theorem 4.18. *Assume $\varphi(\mathbf{m}) \geq 0$ on S^{n-1} and the minimal set $\varphi^{-1}(0)$ contains at least $\{\pm\mathbf{m}_0\}$. Then*

$$\inf_{\mathbf{m}\in L^\infty(\Omega;S^{n-1})} I(\mathbf{m}) = 0.$$

Proof. It suffices to find a sequence $\{\mathbf{m}_k\}$ in $L^\infty(\Omega;S^{n-1})$ such that

$$I(\mathbf{m}_k) = \int_\Omega \varphi(\mathbf{m}_k(x))dx + \frac{1}{2}\int_\Omega \mathbf{m}_k(x)\cdot\nabla u_k(x)\,dx \to 0$$

as $k\to\infty$, where $u_k = T\mathbf{m}_k$ is the solution in X of the simplified Maxwell equation (4.41). To this end, let $\eta \in S^{n-1}$ be such that $\eta \perp \mathbf{m}_0$ and define

$$\mathbf{m}_k(x) = \rho(kx\cdot\eta)\mathbf{m}_0 + (1-\rho(kx\cdot\eta))(-\mathbf{m}_0), \quad x\in\Omega,$$

where $\rho(t)$ is a periodic function of period 1 with

$$\rho(t) = 1 \quad (0\leq t\leq 1/2); \quad \rho(t) = 0 \quad (1/2 < t < 1).$$

It is easy to check $\mathbf{m}_k \rightharpoonup 0$ weakly in $L^2(\Omega;\mathbb{R}^n)$. Let $u_k = T\mathbf{m}_k$. Then, by Lemma 4.17 above, $u_k \rightharpoonup 0$ in $H^1(B)$. Note that, by Rellich-Kondrachov imbedding theorem (Theorem 2.33), imbeddings $H^1(B) \subset\subset L^2(B)$ and $H^1(B) \subset\subset L^2(\partial B)$ are both compact; hence $u_k \to 0$ strongly in both $L^2(\Omega)$ and $L^2(\partial\Omega)$. We now compute by the divergence theorem that

$$\int_{\Omega_j} \mathbf{m}_k\cdot\nabla u_k\,dx = \int_{\partial\Omega_j} u_k\mathbf{m}_k\cdot\nu\,dS - \int_{\Omega_j} u_k \operatorname{div}\mathbf{m}_k\,dx,$$

where ν is the unit outward normal on the boundary and the formula is valid on each piece Ω_j of $\bar{\Omega}$ where $\mathbf{m}_k$ is constant $\mathbf{m}_0$ or $-\mathbf{m}_0$; hence $\operatorname{div}\mathbf{m}_k = 0$ on each Ω_j. Moreover, $\mathbf{m}_k\cdot\nu = 0$ on $\partial\Omega_j \setminus \partial\Omega$. Hence we have

$$\left|\int_\Omega \mathbf{m}_k\cdot\nabla u_k\,dx\right| \leq \int_{\partial\Omega} |u_k(x)|\,dS \leq |\partial\Omega|^{1/2}\|u_k\|_{L^2(\partial\Omega)} \to 0 \quad \text{as } k\to\infty.$$

Finally, noting that $\varphi(\mathbf{m}_k(x)) = 0$, we arrive at

$$I(\mathbf{m}_k) = \int_\Omega \varphi(\mathbf{m}_k(x))dx + \frac{1}{2}\int_\Omega \mathbf{m}_k(x)\cdot\nabla u_k(x)\,dx \to 0$$

as $k\to\infty$. □

Theorem 4.19. *Assume $\varphi^{-1}(0) = \{\pm\mathbf{m}_0\}$. Then there exists no minimizer of $I(\mathbf{m})$ on $L^\infty(\Omega;S^{n-1})$.*

Proof. Suppose $\bar{\mathbf{m}}$ is a minimizer. Then $I(\bar{\mathbf{m}}) = \inf I = 0$ and hence $\varphi(\bar{\mathbf{m}}(x)) = 0$ and $u = T\bar{\mathbf{m}} = 0$. Therefore,

$$\bar{\mathbf{m}}(x) = \chi_E(x)\mathbf{m}_0 + (1-\chi_E(x))(-\mathbf{m}_0), \quad \operatorname{div}(\bar{\mathbf{m}}\chi_\Omega) = 0,$$

where the set $E = \{x \in \Omega \,|\, \bar{\mathbf{m}}(x) = \mathbf{m}_0\}$ is a measurable subset of Ω. Therefore

$$\int_{\mathbb{R}^n} f(x)\nabla\zeta(x)\cdot\mathbf{m}_0\,dx = 0 \quad \forall\,\zeta \in C_0^\infty(\mathbb{R}^n), \tag{4.44}$$

where $f(x) = \chi_\Omega(x)[2\chi_E(x)-1] = 2\chi_E(x) - \chi_\Omega(x)$. Let $x = x' + t\mathbf{m}_0$ where $x' \perp \mathbf{m}_0$ and write $f(x) = g(x',t)$. In (4.44), by change of variables, we have

$$\int_{\mathbb{R}^n} g(x',t)\zeta_t(x',t)\,dx'dt = 0 \quad \forall\,\zeta \in C_0^\infty(\mathbb{R}^n).$$

This implies the weak derivative $g_t(x',t) \equiv 0$ on $\mathbb{R}^n$, and hence $g(x',t) = h(x')$ is independent of t. But $h(x') = f(x'+t\mathbf{m}_0)$ vanishes for large t and hence $h \equiv 0$. So $f \equiv 0$ on $\mathbb{R}^n$. However $f(x) \in \{\pm 1\}$ for $x \in \Omega$. This is a contradiction. $\square$

4.4.6. Representation of Curl-Free Fields. Let $X = L^2(\mathbb{R}^n;\mathbb{R}^n)$ denote the Hilbert space with the inner product and norm defined by

$$(u,v) = \int_{\mathbb{R}^n}(u^1v^1 + \cdots + u^nv^n)\,dx = \int_{\mathbb{R}^n} u\cdot v\,dx; \quad \|u\| = (u,u)^{1/2}.$$

For $u \in L^2(\mathbb{R}^n;\mathbb{R}^n)$, we define $\operatorname{curl} u = (\operatorname{curl} u)_{ij}$ as distribution:

$$\langle(\operatorname{curl} u)_{ij},\varphi\rangle = -\int_{\mathbb{R}^n}(u^i\varphi_{x_j} - u^j\varphi_{x_i})\,dx, \quad \forall\ \varphi \in C_0^\infty(\mathbb{R}^n).$$

Note that, if $u \in W^{1,1}_{loc}(\mathbb{R}^n;\mathbb{R}^n)$, then $\operatorname{curl} u = Du - (Du)^T$. In the case $n=2$ or $n=3$, $\operatorname{curl} u$ can be identified as follows:

$$\operatorname{curl} u \approx \nabla^\perp\cdot u = \nabla\cdot u^\perp = \operatorname{div}(u^\perp) = u^1_{x_2} - u^2_{x_1} \quad (n=2);$$
$$\operatorname{curl} u \approx \nabla\times u = (u^3_{x_2} - u^2_{x_3}, u^1_{x_3} - u^3_{x_1}, u^2_{x_1} - u^1_{x_2}) \quad (n=3).$$

Define the subspace of curl-free fields as follows:

$$X_{\text{curl}} = \{u \in L^2(\mathbb{R}^n;\mathbb{R}^n) \,|\, \operatorname{curl} u = 0 \text{ in the sense of distribution}\}.$$

Let

$$Y = \{f \in H^1_{loc}(\mathbb{R}^n) \,|\, \|f\|_* < \infty\},$$

where $\|f\|_*$ is defined by

$$\|f\|_*^2 = \int_{\mathbb{R}^n}|\nabla f(x)|^2\,dx + \sup_{R\ge 1}\frac{1}{R^{n+2}}\int_{B_R}|f(x)|^2\,dx. \tag{4.45}$$

Clearly, $\nabla f \in X_{\text{curl}}$ for all $f \in Y$; the converse is also true.

Theorem 4.20. (Representation by Local Functions) *The space Y is a Banach space under the norm $\|f\|_*$ defined. Moreover, the gradient operator $\nabla\colon Y \to X_{\text{curl}}$ is surjective; more precisely, for any $v \in X_{\text{curl}}$, there exists a $f \in Y$ such that*

$$v = \nabla f, \quad \|f\|_* \le C_n\|v\|.$$

Proof. The proof that Y is a Banach space follows directly by the definition and will not be given here. We prove the rest of the theorem.

Given $v \in X_{\rm curl}$, let $v_\epsilon = v * \rho_\epsilon$ be the smooth approximation of v. Then $v_\epsilon \in X_{\rm curl} \cap C^\infty(\mathbb{R}^n;\mathbb{R}^n)$. Define

$$f_\epsilon(x) = \int_0^1 v_\epsilon(tx)\cdot x\,dt.$$

Then one can easily compute that

$$\begin{aligned}
\partial_{x_j} f_\epsilon(x) &= \int_0^1 v_\epsilon^j(tx)dt + \int_0^1 (\partial_{x_j} v_\epsilon)(tx)\cdot tx\,dt \\
&= \int_0^1 v_\epsilon^j(tx)dt + \int_0^1 \nabla v_\epsilon^j(tx)\cdot tx\,dt \\
&= \int_0^1 v_\epsilon^j(tx)dt + \int_0^1 \frac{d}{dt}[v_\epsilon^j(tx)]t\,dt \\
&= \int_0^1 v_\epsilon^j(tx)dt + \left[v_\epsilon^j(tx)t\Big|_0^1 - \int_0^1 v_\epsilon^j(tx)dt \right] \\
&= v_\epsilon^j(x),
\end{aligned}$$

where we have used the condition $\partial_{x_j} v_\epsilon = \nabla v_\epsilon^j$ from $\operatorname{curl} v_\epsilon = 0$. This proves $\nabla f_\epsilon(x) = v_\epsilon(x)$ for all $x \in \mathbb{R}^n$. Therefore, $v_\epsilon = \nabla f_\epsilon$ and hence, for all $x, y \in \mathbb{R}^n$,

$$f_\epsilon(x+y) - f_\epsilon(y) = \int_0^1 v_\epsilon(y+tx)\cdot x\,dt.$$

Hence

$$\begin{aligned}
|f_\epsilon(x+y) - f_\epsilon(y)|^2 &= \left| \int_0^1 v_\epsilon(y+tx)\cdot x\,dt \right|^2 \\
&\le |x|^2 \int_0^1 |v_\epsilon(y+tx)|^2\,dt.
\end{aligned}$$

Integrating this inequality over $x \in B_R(0) = B_R$, we obtain

$$\begin{aligned}
\int_{B_R(y)} |f_\epsilon(z) - f_\epsilon(y)|^2\,dz &\le R^2 \int_0^1 \left(\int_{B_R} |v_\epsilon(y+tx)|^2\,dx \right) dt \\
&= R^2 \int_0^1 \left(\int_{B_{tR}(y)} |v_\epsilon(z)|^2\,dz \right) t^{-n}\,dt \\
&= R^{n+2} \int_0^1 \left(\frac{1}{|B_{tR}(y)|} \int_{B_{tR}(y)} |v_\epsilon(z)|^2\,dz \right) dt \\
&\le R^{n+2} \mathcal{M}(|v_\epsilon|^2)(y),
\end{aligned}$$

where $\mathcal{M}(h)$ is the maximal function of h (see Stein's book). Since $|v_\epsilon|^2 \in L^1(\mathbb{R}^n)$, it follows that

$$m\{y \in \mathbb{R}^n \,|\, \mathcal{M}(|v_\epsilon|^2)(y) > \alpha\} \le \frac{5^n}{\alpha} \int_{\mathbb{R}^n} |v_\epsilon|^2\,dx \le \frac{5^n}{\alpha} \int_{\mathbb{R}^n} |v|^2\,dx.$$

Let

$$E_\epsilon = \{y \in B_1 \,|\, \mathcal{M}(|v_\epsilon|^2)(y) \le \alpha_0\},$$

where we choose

$$\alpha_0 = \frac{2 \cdot 5^n}{|B_1|} \int_{\mathbb{R}^n} |v|^2 dx.$$

Then it follows that $|E_\epsilon| \geq \frac{1}{2}|B_1|$ for all ϵ. Therefore, it is a simple exercise to show that there exists a sequence $\epsilon_k \to 0$ and a point $y_0 \in B_1$ such that $y_0 \in \cap_{k=1}^\infty E_{\epsilon_k}$; that is,

$$\mathcal{M}(|v_{\epsilon_k}|^2)(y_0) \leq \alpha_0 = \frac{2 \cdot 5^n}{|B_1|} \|v\|^2, \quad \forall\, k = 1, 2, \cdots .$$

Using this y_0 we define a new sequence

$$g_k(z) = f_{\epsilon_k}(z) - f_{\epsilon_k}(y_0), \quad z \in \mathbb{R}^n.$$

Then, for all $R \geq 1$, we have

$$\int_{B_R} |g_k(z)|^2\, dz \leq \int_{B_{2R}(y_0)} |g_k(z)|^2\, dz \leq (2R)^{n+2} \frac{2 \cdot 5^n}{|B_1|} \|v\|^2.$$

By using diagonal subsequences, there exists a subsequence g_{k_j} and a function $f \in L^2_{loc}(\mathbb{R}^n)$ such that $g_{k_j} \rightharpoonup f$ weakly as $k_j \to \infty$ on all balls $B_R(0)$, $R > 0$. This function f must satisfy $\nabla f = v \in L^2(\mathbb{R}^n; \mathbb{R}^n)$ and

$$\sup_{R \geq 1} \frac{1}{R^{n+2}} \int_{B_R} |f(x)|^2\, dx \leq C_n \int_{\mathbb{R}^n} |v(x)|^2\, dx;$$

hence $\|f\|_* \leq C\|v\|$. This completes the proof. □

Given any $u \in X = L^2(\mathbb{R}^n; \mathbb{R}^n)$, for each $R > 0$, let $B_R = B_R(0)$ and consider the following minimization problem:

$$\inf_{\varphi \in H^1_0(B_R)} \int_{B_R} |\nabla\varphi - u|^2\, dx. \tag{4.46}$$

By direct method, this problem has a unique solution, which we denote by φ_R. We also extend φ_R by zero to all $\mathbb{R}^n$. This sequence $\{\varphi_R\}$ is of course uniquely determined by $u \in X$. It also satisfies the following properties:

$$\textstyle\int_{\mathbb{R}^n} (\nabla\varphi_R - u) \cdot \nabla\zeta\, dx = 0 \quad \forall\, \zeta \in H^1_0(\Omega),\ \Omega \subseteq B_R, \tag{4.47}$$

$$\|\nabla\varphi_R\|_{L^2(\mathbb{R}^n)} \leq \|u\|. \tag{4.48}$$

Lemma 4.21. *Given $u \in X$, it follows that $\nabla\varphi_R \to v$ in X as $R \to \infty$ for some $v \in X_{\mathrm{curl}}$ uniquely determined by u. Moreover, this v satisfies*

$$\|v - u\| = \min_{v' \in X_{\mathrm{curl}}} \|v' - u\|;$$

therefore, $v = u$ if $u \in X_{\mathrm{curl}}$.

Proof. First of all, we claim $\nabla\varphi_R \rightharpoonup v$ weakly in X as $R \to \infty$. Let v', v'' be the weak limits of any two subsequences $\{\nabla\varphi_{R'}\}$ and $\{\nabla\varphi_{R''}\}$, where R', R'' are two sequences going to ∞. We would like to show $v' = v''$, which shows that $\nabla\varphi_R \rightharpoonup v$ as $R \to \infty$. Since $\nabla\varphi_R \in X_{\mathrm{curl}}$, it follows easily that $v', v'' \in X_{\mathrm{curl}}$ and, by (4.47) above,

$$\int_{\mathbb{R}^n} (v' - u) \cdot \nabla\zeta = \int_{\mathbb{R}^n} (v'' - u) \cdot \nabla\zeta = 0 \quad \forall\, \Omega \subset\subset \mathbb{R}^n,\ \forall\, \zeta \in H^1_0(\Omega). \tag{4.49}$$

This implies $\mathrm{div}(v' - u) = \mathrm{div}(v'' - u) = 0$ and hence $\mathrm{div}(v' - v'') = 0$. Since $\mathrm{curl}(v' - v'') = 0$, it follows that $\Delta(v' - v'') = 0$ in the sense of distributions on $\mathbb{R}^n$; therefore, $v' - v'' \in C^\infty(\mathbb{R}^n) \cap X$ is harmonic component-wise. By the **mean value property** of harmonic functions, it follows that $v' - v'' \equiv 0$. Hence $v' = v''$. We denote this weak limit by $v \in X_{\text{curl}}$. Note that, by (4.49), $\operatorname{div}(v-u) = 0$. If $\operatorname{curl} u = 0$ then $\operatorname{curl}(v-u) = 0$ and hence $v = u$. We now prove $\nabla\varphi_R \to v$ in X as $R \to \infty$. Taking $\zeta = \varphi_R \in H_0^1(B_R)$ in (4.49) and letting $R \to \infty$ we have

$$\int_{\mathbb{R}^n} (v-u)\cdot v\,dx = 0. \tag{4.50}$$

Using $\zeta = \varphi_R$ in (4.47), taking $R \to \infty$ and by weak limit, we have

$$\lim_{R\to\infty} \int_{\mathbb{R}^n} |\nabla\varphi_R|^2\,dx = \int_{\mathbb{R}^n} v\cdot u = \int_{\mathbb{R}^n} |v|^2\,dx.$$

This implies $\nabla\varphi_R \to v$ strongly in $L^2(\mathbb{R}^n;\mathbb{R}^n)$. We now show that

$$\|v-u\| = \min_{v' \in X_{\text{curl}}} \|v'-u\|. \tag{4.51}$$

Given any $v' \in X_{\text{curl}}$, choose the sequence $\tilde\varphi_R$ corresponding to $v'-v$. Since $\operatorname{div}(v-u) = 0$, it easily follows that

$$(v'-v, v-u) = \lim_{R\to\infty} (\nabla\tilde\varphi_R, v-u) = 0.$$

Hence $\|v'-u\|^2 = \|v-v'\|^2 + 2(v'-v, v-u) + \|v-u\|^2 \ge \|v-u\|^2$; this proves (4.51). □

Theorem 4.22. *Let $n \ge 3$ and Y_1 be the closure of $C_0^\infty(\mathbb{R}^n)$ in Y. Then,*

$$Y_1 = \{f \in L^{\frac{2n}{n-2}}(\mathbb{R}^n) \mid \nabla f \in L^2(\mathbb{R}^n;\mathbb{R}^n)\}$$

and Y_1 has the equivalent norms $\|f\|_* \approx \|\nabla f\|_{L^2(\mathbb{R}^n)}$ for all $f \in Y_1$; therefore Y_1 is the closure of $C_0^\infty(\mathbb{R}^n)$ under the norm $\|\nabla f\|_{L^2(\mathbb{R}^n)}$. Furthermore the gradient operator $\nabla\colon Y_1 \to X_{\text{curl}}$ is bijective.

Proof. Let $W = \{f \in L^{\frac{2n}{n-2}}(\mathbb{R}^n) \mid \nabla f \in L^2(\mathbb{R}^n;\mathbb{R}^n)\}$.

1. We show that $Y_1 \subseteq W$ and

$$\|f\|_* \le C\|\nabla f\|_{L^2(\mathbb{R}^n)} \quad \forall\ f \in Y_1. \tag{4.52}$$

Let $f \in Y_1$. Then there exists a sequence $f_j \in C_0^\infty(\mathbb{R}^n)$ such that $\|f_j - f\|_* \to 0$ as $j \to \infty$. Therefore $\|\nabla f_j\|_{L^2} \to \|\nabla f\|_{L^2}$. By the **Sobolev-Galiardo-Nirenberg inequality**,

$$\|f_j\|_{L^{2^*}(\mathbb{R}^n)} \le C\,\|\nabla f_j\|_{L^2(\mathbb{R}^n)} \quad \forall\ j.$$

Hence $f_j \rightharpoonup g \in L^{2^*}(\mathbb{R}^n)$. Since $f_j \to f$ in $L^2(B_R)$ for all $R > 0$. We have $f = g$. Hence $f \in W$. Furthermore, by Hölder's inequality,

$$\|f_j\|_{L^2(B_R)} \le c_n\,R^2\,\|f_j\|_{L^{2^*}(B_R)} \le C\,R^2\,\|\nabla f_j\|_{L^2(\mathbb{R}^n)}.$$

Hence, by taking limits as $j \to \infty$, it follows that

$$\sup_{R\ge 1} \frac{1}{R^{n+2}} \int_{B_R} |f|^2\,dx \le C \int_{\mathbb{R}^n} |\nabla f(x)|^2\,dx,$$

which proves (4.52).

2. We show that $W \subseteq Y_1$. Let $f \in W$. Define $f_j = f\rho_j$, where $\rho_j \in W^{1,\infty}(\mathbb{R}^n)$ defined by $\rho_j(x) = 1$ on $|x| \le j$, $\rho_j(x) = 0$ on $|x| \ge 2j$ and $\rho_j(x)$ is linear in $|x|$ for $j \le |x| \le 2j$. Then $f_j \in Y_1$. It can be easily shown that

$$\lim_{j\to\infty} \|f_j - f\|_* = 0,$$

which proves $f \in Y_1$ and hence $W \subseteq Y_1$.

3. We show that $\nabla\colon Y_1 \to X_{\mathrm{curl}}$ is surjective. Given any $u \in X_{\mathrm{curl}}$, let $\varphi_R \in H_0^1(B_R)$ be the function determined as in the minimization problem (4.46) above. Since $\operatorname{curl} u = 0$, it follows that $\nabla\varphi_R \to u$ in $X = L^2(\mathbb{R}^2;\mathbb{R}^2)$ as $R \to \infty$. Since $\varphi_R \in W$, we have $\varphi_R \in Y_1$. Inequality (4.52) implies that sequence $\{\varphi_R\}$ is a Cauchy sequence in Y_1 and hence its limit f belongs to Y_1; moreover, since $\|\varphi_R - f\|_* \approx \|\nabla\varphi_R - \nabla f\|_{L^2(\mathbb{R}^n)} \to 0$ as $R \to \infty$, we have $\nabla f = u$. This completes the proof. $\square$

4.5. Mountain Pass Theorem

4.5.1. The Palais-Smale Condition. Let $E : X \to \mathbb{R}$ be G-diff on the Banach space X. We say E satisfies the **Palais-Smale condition (PS)** if whenever $\{u_n\}$ is a sequence in X such that $E(u_n)$ is bounded and $\|E'(u_n)\| \to 0$, then $\{u_n\}$ has a convergent subsequence.

Remark. The (PS) condition is not satisfied by very smooth functions very often. For example, the function $E : \mathbb{R} \to \mathbb{R}$ with $E(u) = \cos u$ does not satisfy (PS), which can be easily seen by considering the sequence $u_n = n\pi$. Similarly, the function $E(u) = c$ does not satisfy (PS). It can be shown that if E is F-diff on the Banach space X (not necessarily reflexive) and is bounded below and satisfies (PS), then E attains its minimum value. But we study the case where functionals are neither bounded from above or below.

Lemma 4.23. (Deformation Lemma) *Let $E : X \to \mathbb{R}$ be a C^1 functional satisfying (PS). Let $c, s \in \mathbb{R}$ and define*

$$K_c = \{u \in X : \ E(u) = c, \ E'(u) = 0\}$$

$$A_s = \{u \in X : \ E(u) \le s\}.$$

Assume $K_c = \emptyset$. Then there exists an $\bar{\varepsilon} > 0$ and a continuous function $\eta : [0,1] \times X \to X$ such that for all $0 < \varepsilon \le \bar{\varepsilon}$,

(i) $\eta(0,u) = u$ *for all* $u \in X$,

(ii) $\eta(t,u) = u$ *for all* $t \in [0,1], u \notin E^{-1}([c-\varepsilon, c+\varepsilon])$,

(iii) $E(\eta(t,u)) \le E(u)$ *for all* $t \in [0,1],\ u \in X$,

(iv) $\eta(1, A_{c+\varepsilon}) \subset A_{c-\varepsilon}$.

This lemma shows that if c is not a critical level, then we can nicely deform the set $A_{c+\varepsilon}$ into $A_{c-\varepsilon}$ for some $\varepsilon > 0$. For a proof, see EVANS's book.

4.5.2. The Mountain Pass Theorem. We now prove the main theorem of this section.

Theorem 4.24. (Ambrosetti-Rabinowitz) *Let $E : X \to \mathbb{R}$ be a C^1 functional satisfying (PS) on the real Banach space X. Let $u_0, u_1 \in X,\ c_0 \in \mathbb{R}$ and $R > 0$ be such that*

(i) $\|u_1 - u_0\| > R$

(ii) $E(u_0), E(u_1) < c_0 \le E(v)$ *for all v such that $\|v - u_0\| = R$.*

Then E has a critical point u with $E(u) = c,\ c \ge c_0$; the critical value c is defined by

$$c = \inf_{p\in K} \sup_{t\in[0,1]} E(p(t)) \tag{4.53}$$

where K denotes the set of all continuous maps $p : [0,1] \to X$ with $p(0) = u_0$ and $p(1) = u_1$.

Remark. Think of the graph of E as a landscape with a low spot at u_0, surrounded by a ring of mountains. Beyond these mountains lies another low point at u_1. Note that every path p connecting u_0 to u_1 has to cross the sphere $\{v : \|v - u_0\| = R\}$ since u_1 lies outside the sphere. Moreover, on this sphere the value of E is at least c_0. Hence the maximum value of $E(p(t))$ for any such path p is at least c_0. Hence $c \geq c_0$. An important aspect of the MPT is that the critical point u at level c is distinct from u_0 and u_1. Hence, if u_0 already satisfies $E'(u_0) = 0$ by some other method, then u will give a second solution of $E'(u) = 0$.

Proof. Let c be as defined in (4.53). If it is not a critical value, then $K_c = \emptyset$. Let η and $\bar{\varepsilon}$ be as in the deformation lemma. Now by condition (ii) of the theorem, we can choose ε small enough so that $0 < \varepsilon < \bar{\varepsilon}$ and $E(u_0), E(u_1) \notin [c-\varepsilon, c+\varepsilon]$ (since $c \geq c_0$). Let $p \in K$ and define the path $\xi : [0,1] \to X$ by

$$\xi(t) = \eta(1, p(t)).$$

Since $p(0) = u_0$ and $p(1) = u_1$, it follows, by the choice of ε, that

$$\xi(0) = \eta(1, u_0) = u_0,\ \xi(1) = \eta(1, u_1) = u_1$$

using condition (ii) of the lemma. Thus $\xi \in K$. Now, we can choose $p \in K$ such that

$$\max_{t\in[0,1]} E(p(t)) < c + \varepsilon.$$

Since $p(t) \in A_{c+\varepsilon}$, by (iv) of the lemma, $\xi(t) \in A_{c-\varepsilon}$. Thus

$$\max_{t\in[0,1]} E(\xi(t)) \leq c - \varepsilon$$

which contradicts the definition of c. Hence $K_c \neq \emptyset$. $\square$

4.5.3. Saddle Point Solutions. In order to illustrate these ideas we apply the MPT to the problem

$$\Delta u + f(x,u) = 0 \ \text{ in } \ \Omega, \quad u|_{\partial\Omega} = 0 \tag{4.54}$$

where $f \in \text{Car}$ satisfies the growth condition

$$|f(x,z)| \leq c(1+|z|^\tau),\ |f_z(x,z)| \leq c(1+|z|^{\tau-1}) \tag{4.55}$$

with $1 < \tau < (n+2)/(n-2)$. Moreover, we assume $f(x,0) = 0$ and

$$0 \leq F(x,z) \leq \gamma f(x,z) z \tag{4.56}$$

for some constant $\gamma < 1/2$, where $F(x,z) = \int_0^z f(x,s)ds$. Finally we assume

$$a|z|^{\tau+1} \leq |F(x,z)| \leq A|z|^{\tau+1} \tag{4.57}$$

for constants $0 < a \leq A$.

Remarks. (i) A typical example is given by $f(x,z) = l(x)|z|^{\tau-1}z$, where $l \in C(\bar{\Omega})$ and $l(x) \geq \alpha > 0$. In this case, $F(x,z) = \frac{1}{\tau+1} l(x)|z|^{\tau+1}$ satisfies (4.56) and (4.57) with $\gamma = \frac{1}{\tau+1} < \frac{1}{2}$.

(ii) Note that assumption (4.57) implies that

$$F_z(x,0) = f(x,0) = 0.$$

Hence $u \equiv 0$ is a trivial solution to (4.54).

Associated with (4.54) is the functional I defined by

$$I(u) = \frac{1}{2}\int_\Omega |\nabla u|^2\,dx - J(u),$$

where
$$J(u)=\int_\Omega F(x,u)dx.$$
The following lemma simplifies some of the technical details that follow.

Lemma 4.25. *Let I and J be defined on $H=H_0^1(\Omega)$ as above. Then both I and J are of C^1 on H and*

(a) *$J'\colon H\to H^*$ is compact.*

(b) *If $\{u_n\}$ is a bounded sequence in H such that $I'(u_n)\to 0$ as $n\to\infty$, then $\{u_n\}$ has a convergent subsequence.*

Proof. The norm $\int_\Omega |\nabla u|^2dx$ is obviously of C^∞ on H. J is a Nemytskii operator with F being C^1 in u, and hence is of C^1 on H with $J'\colon H\to H^*$ given by
$$\langle J'(u),v\rangle=\int_\Omega f(x,u(x))\,v(x)\,dx,\quad \forall\ u,\,v\in H.$$
We have the estimate
$$\begin{aligned}\|J'(u)-J'(w)\|_{H^*} &= \sup_{v\in H,\,\|v\|\le 1}|\langle J'(u)-J'(w),v\rangle|\\ &\le \|f(\cdot,u)-f(\cdot,w)\|_{(\tau+1)/\tau}\|v\|_{\tau+1}\\ &\le C\,\|f(\cdot,u)-f(\cdot,w)\|_{(\tau+1)/\tau},\quad \forall\ u,\,w\in H,\end{aligned}$$
where we have used the Sobolev inequality $\|v\|_{\tau+1}\le C\|v\|_H\le C$ since $\tau+1<2^*=\frac{2n}{n-2}$.

Let $\{u_n\}$ be a bounded sequence in H. Then by choosing a subsequence if necessary, $u_n\rightharpoonup u$ in H and $u_n\to u$ in $L^{\tau+1}(\Omega)$. By the continuity of Nemyskii operators, $f(\cdot,u_n)\to f(\cdot,u)$ in $L^{(\tau+1)/\tau}$. So, by the estimate above, $J'(u_n)$ converges to $J'(u)$ in H^* and hence $J'\colon H\to H^*$ is compact.

Let $A:H\to H^*$ denote the **duality map** defined by
$$\langle Au,v\rangle=\int_\Omega \nabla u\cdot\nabla v\,dx\quad\text{for all}\quad u,v\in H.$$
Then $\langle I'(u),v\rangle=\langle Au,v\rangle-\langle J'(u),v\rangle$ for all $u,v\in H$. Note that, by the Riesz representation or Lax-Milgram theorem, A^{-1} exists and is bounded on H^*; hence, it follows that
$$A^{-1}I'(u)=u-A^{-1}J'(u).$$
Let $\{u_n\}$ be a bounded sequence in H with $I'(u_n)\to 0$ as $n\to\infty$. Since A^{-1} is linear and continuous and J' is compact, by passing to a subsequence if necessary, we have $u_n=A^{-1}I'(u_n)+A^{-1}J'(u_n)$ converges. □

Remark. We only need the assumption (4.55). Also, from the proof we see that the compactness of the embedding $H=H_0^1(\Omega)\subset L^{\tau+1}(\Omega)$ is necessary.

Theorem 4.26. *The boundary value problem (4.54) has at least one nontrivial solution $u\in H_0^1(\Omega)$.*

Proof. Consider the C^1 functional $I(u)$ above, which we rewrite as
$$I(u)=\frac{1}{2}\|u\|_{1,2,0}^2-J(u),\quad\text{where } \|u\|_{1,2,0}^2=\textstyle\int_\Omega|\nabla u|^2\,dx,$$
for $u\in H_0^1(\Omega)$. In order to apply MPT, we first verify the Palais-Smale condition. Let $\{u_n\}\subset H_0^1(\Omega)$ be a sequence such that $|I(u_n)|\le c$ for all n and $I'(u_n)\to 0$. If we can show

that $\{\|u_n\|_{1,2,0}\}$ is bounded, then according to Lemma 4.25, I satisfies (PS). For the given sequence

$$I(u_n) = \frac{1}{2}\|u_n\|^2_{1,2,0} - \int_\Omega F(x, u_n)dx$$

and

$$\langle I'(u_n), u_n\rangle = \|u_n\|^2_{1,2,0} - \int_\Omega u_n f(x, u_n)dx$$

and so, by (4.56),

$$\begin{aligned}\frac{1}{2}\|u_n\|^2_{1,2,0} &\le I(u_n) + \gamma \int_\Omega u_n f(x, u_n)dx \\ &\le c + \gamma(\|u_n\|^2_{1,2,0} + \|I'(u_n)\|\ \|u_n\|_{1,2,0}).\end{aligned}$$

Since $\gamma < 1/2$ and $\|I'(u_n)\| \to 0$, this inequality implies that $\|u_n\|^2_{1,2,0}$ is bounded. Hence (PS) condition is proved. To verify the conditions (i), (ii) in the Mountain Pass Theorem, we choose $u_0 = 0$. Clearly, $I(u_0) = 0$. Now we show that $I|_{\partial B(0,R)} \ge c_0$ for some $R, c_0 > 0$. In view of the embedding $H^1_0(\Omega) \subset L^{\tau+1}(\Omega)$, we have, by (4.57),

$$J(u) \le A \int_\Omega |u|^{\tau+1}dx \le c\|u\|^{\tau+1}_{1,2,0}.$$

Thus for all u satisfying $\|u\|_{1,2,0} = R$

$$I(u) \ge \frac{R^2}{2} - cR^{\tau+1} \ge c_0 > 0$$

provided we take R sufficiently small. (Here we need $\tau > 1$, which is not needed before.) Next, let $v = tu$, where $u \ne 0$ is a fixed function in $H^1_0(\Omega)$, and $t > 0$ is to be selected. Then, by (4.57),

$$\begin{aligned} I(v) &= \frac{t^2}{2}\|u\|^2_{1,2,0} - J(tu) \\ &\le \frac{t^2}{2}\|u\|^2_{1,2,0} - at^{\tau+1}\int_\Omega |u|^{\tau+1}dx \\ &< 0 \end{aligned}$$

if $t > 0$ is sufficiently large. (Again $\tau > 1$ is needed.) Moreover, $v \in H^1_0(\Omega)\backslash \bar{B}(0, R)$. Therefore, by MPT, there exists a critical point $u \ne 0$ of I which is a weak solution of (4.54). □

4.6. Nonexistence and Radial Symmetry

In this section, we always assume Ω is a bounded domain in $\mathbb{R}^n$ and $\partial\Omega$ is C^1. We will show that for some of such domains the boundary value problem (4.36) (with $k = 0, l = 1$) does not have nontrivial classical solutions if $\tau > (n + 2)/(n - 2)$.

4.6.1. Pohozaev's Identity. We shall need the following fundamental identity:

Theorem 4.27. (Pohozaev) *Let $u \in C^2(\bar{\Omega})$, $u|_{\partial\Omega} = 0$. Then*

$$\int_{\partial\Omega} u^2_\nu (x\cdot\nu)dS = \int_\Omega [2(\nabla u\cdot x) + (n-2)u]\Delta u dx \tag{4.58}$$

where $\nu = (\nu_1, \cdots, \nu_n)$ denotes the outward unit normal to $\partial\Omega$ and $u_\nu = \frac{\partial u}{\partial \nu} = \nabla u\cdot\nu = \gamma_1(u)$ denotes the normal derivative of u on $\partial\Omega$.

Proof. By an easy calculation we have the following identity

$$2\Delta u(\nabla u \cdot x) = \operatorname{div}[2(\nabla u \cdot x)\nabla u - |\nabla u|^2 x] + (n-2)|\nabla u|^2.$$

An application of the divergence theorem gives

$$\int_{\partial\Omega} [2(\nabla u \cdot x)u_\nu - |\nabla u|^2 x \cdot \nu] dS = \int_\Omega [2\Delta u(\nabla u \cdot x) - (n-2)|\nabla u|^2] dx.$$

Since $u = 0$ on $\partial\Omega$, $\int_\Omega u\Delta u dx + \int_\Omega |\nabla u|^2 dx = 0$ and $u_{x_i} = \nu_i u_\nu$, i.e., $\nabla u = u_\nu \nu$, which combined with the above integral identity yields (4.58). □

Corollary 4.28. *Let $u \in C^2(\bar\Omega)$ be a solution of the boundary value problem*

$$\Delta u + f(u) = 0 \ \text{ in } \ \Omega, \quad u|_{\partial\Omega} = 0.$$

Let $F(u) = \int_0^u f(t)dt$. Then u satisfies the equation

$$\frac{1}{2}\int_{\partial\Omega} u_\nu^2 (x \cdot \nu) dS = n\int_\Omega F(u)dx + \frac{2-n}{2}\int_\Omega uf(u)dx. \tag{4.59}$$

Proof. From (4.58) we have

$$\frac{1}{2}\int_{\partial\Omega} u_\nu^2 (x \cdot \nu) dS = \int_\Omega \Delta u(\nabla u \cdot x)dx + \frac{2-n}{2}\int_\Omega uf(u)dx$$

so it suffices to show that

$$\int_\Omega \Delta u(\nabla u \cdot x)dx = n\int_\Omega F(u)dx.$$

In this direction, we first note that $\nabla(F(u)) = f(u)\nabla u$; hence

$$\Delta u(\nabla u \cdot x) = -f(u)(\nabla u \cdot x) = -\nabla(F(u)) \cdot x.$$

Since $\nabla \cdot (xF(u)) = nF(u) + x \cdot \nabla(F(u))$, we have

$$\int_\Omega \Delta u(\nabla u \cdot x)dx = \int_\Omega nF(u)dx - \int_\Omega \nabla \cdot (xF(u))dx.$$

However, the second integral on the right is zero, as may be seen by applying the divergence theorem and noting that $F(u) = 0$ on $\partial\Omega$ since $u = 0$ on $\partial\Omega$. □

Example 4.29. (H. Weinberger) Let Ω be a bounded domain in $\mathbb{R}^n$ with smooth boundary $\partial\Omega$. Assume $u \in C^3(\bar\Omega)$ is a solution to $\Delta u = -1$ in Ω and $u = 0$ on $\partial\Omega$. If the normal derivative u_ν is a constant c on $\partial\Omega$, then Ω must be a ball of radius nc.

Proof. Use (4.59) with $f(u) = 1$ and $F(u) = u$ and we have

$$c^2\int_{\partial\Omega}(x \cdot \nu)dS = (n+2)\int_\Omega u\, dx.$$

But $\int_{\partial\Omega}(x \cdot \nu)dS = \int_{\partial\Omega} \nabla(\frac{|x|^2}{2}) \cdot \nu dS = \int_\Omega \Delta(\frac{|x|^2}{2})\, dx = n|\Omega|$; so we obtain

$$(n+2)\int_\Omega u\, dx = nc^2|\Omega|. \tag{4.60}$$

On the other hand, by Cauchy-Schwarz inequality, we have

$$1 = -\Delta u \le \sum_{i=1}^n |u_{x_i x_i}| \le \sqrt{n}(\sum_{i=1}^n u_{x_i x_i}^2)^{1/2} \le \sqrt{n}(\sum_{i,j=1}^n u_{x_i x_j}^2)^{1/2}. \tag{4.61}$$

Consider function $v = |\nabla u|^2 + \frac{2}{n}u$. From $\Delta u = -1$, it follows that, by (4.61),

$$\Delta v = 2\sum_{i,j=1}^{n} u_{x_i x_j}^2 - \frac{2}{n} \geq 0 \quad \text{in } \Omega.$$

From $u = 0$ on $\partial\Omega$, we have $|\nabla u|^2 = u_\nu^2 = c^2$ on $\partial\Omega$; hence $v = c^2$ on $\partial\Omega$. By strong maximum principle, we conclude that either $v < c^2$ in Ω or $v \equiv c^2$ in Ω. Suppose $v < c^2$ in Ω. Then

$$c^2|\Omega| > \int_\Omega v\,dx = \int_\Omega |\nabla u|^2\,dx + \frac{2}{n}\int_\Omega u\,dx = -\int_\Omega u\Delta u\,dx + \frac{2}{n}\int_\Omega u\,dx = \frac{n+2}{n}\int_\Omega u\,dx,$$

which contradicts (4.60). Hence $v \equiv c^2$ in Ω. This implies $\Delta v = 2\sum_{i,j=1}^n u_{x_i x_j}^2 - \frac{2}{n} = 0$ in Ω; hence $n\sum_{i,j=1}^n u_{x_i x_j}^2 = 1$. Therefore, inequalities in (4.61) are all equalities. This implies $u_{x_i x_j} = -\frac{1}{n}\delta_{ij}$ in Ω. So $u(x) = -\frac{1}{2n}|x|^2 + p\cdot x + k$ in Ω, where $p \in \mathbb{R}^n$ and $k \in \mathbb{R}$ are constants; thus u can be written as

$$u(x) = \frac{1}{2n}(A - |x - x_0|^2)$$

for some $x_0 \in \mathbb{R}^n$ and $A \in \mathbb{R}$. Since the set $\{u = 0\}$ is nonempty, we must have $A > 0$. Again, by strong maximum principle, $u > 0$ in Ω; hence $\Omega = \{x \in \mathbb{R}^n : |x - x_0| < \sqrt{A}\}$. From $u_\nu = c$ on $\partial\Omega$, we have the radius of the ball Ω is $\sqrt{A} = nc$. □

4.6.2. Star-shaped Domains. Let Ω be an open set containing 0. We say Ω is **star-shaped** (with respect to 0) if, for each $x \in \bar\Omega$, the line segment $\{\lambda x : 0 \leq \lambda \leq 1\}$ lies in $\bar\Omega$.

Remark. Clearly if Ω is convex and $0 \in \Omega$, then Ω is star-shaped. An annulus is not star-shaped since $x \cdot \nu < 0$ on the boundary of the inner circle.

Lemma 4.30. *Assume $\partial\Omega$ is C^1 and Ω is star-shaped with respect to $0 \in \Omega$. Then $x\cdot\nu(x) \geq 0$ for all $x \in \partial\Omega$, where $\nu(x)$ is the outward unit normal at $x \in \partial\Omega$.*

Proof. Given $x_0 \in \partial\Omega$, since $\partial\Omega$ is C^1, there exists a ball $B = B_\varepsilon(x_0)$ and a C^1 function ϕ on B such that

$$\Omega \cap B = \{x \in B \,|\, \phi(x) < 0\}, \quad \partial\Omega \cap B = \{x \in B \,|\, \phi(x) = 0\}.$$

Note that the outward unit normal at x_0 is now given by $\nu(x_0) = \frac{\nabla\phi(x_0)}{|\nabla\phi(x_0)|}$. Let $\delta > 0$ be sufficiently small so that $\lambda x_0 \in B$ for all $\lambda \in [1-\delta, 1]$; hence, from Ω being star-shaped, $\lambda x_0 \in \bar\Omega \cap B$ for all $\lambda \in [1-\delta, 1]$. Consider $h(\lambda) = \phi(\lambda x_0)$ defined on $\lambda \in [1-\delta, 1]$. Then h has the maximum 0 at right-end point 1 and hence

$$h'(1^-) = \nabla\phi(x_0)\cdot x_0 \geq 0,$$

which proves $x_0 \cdot \nu(x_0) \geq 0$ for all $x_0 \in \partial\Omega$. □

4.6.3. Nonexistence of Classical Solutions.

Theorem 4.31. *Let Ω be star-shaped with respect to $x = 0$. Then the problem*

$$\Delta u + |u|^{\tau-1}u = 0 \ \text{ in } \Omega, \ \ u|_{\partial\Omega} = 0 \qquad (\tau > \frac{n+2}{n-2}) \tag{4.62}$$

has no nontrivial $C^2(\bar\Omega)$ solution. Furthermore, the problem

$$\Delta u + u^\tau = 0 \ \text{ in } \Omega, \ \ u|_{\partial\Omega} = 0 \qquad (\tau \geq \frac{n+2}{n-2}) \tag{4.63}$$

has no positive $C^2(\bar{\Omega})$ solution.

Proof. Suppose u is a nontrivial $C^2(\bar{\Omega})$ solution of (4.62). Applying formula (4.59) with $f(u) = |u|^{\tau-1}u$ (and thus $F(u) = \frac{1}{\tau+1}|u|^{\tau+1}$), and using Lemma 4.30, we obtain

$$
\begin{aligned}
&(4.64) \qquad \frac{1}{2}\int_{\partial\Omega} u_\nu^2(x\cdot\nu)dS &&= n\int_\Omega \frac{|u|^{\tau+1}}{\tau+1}dx + \frac{2-n}{2}\int_\Omega |u|^{\tau+1}dx \\
&(4.65) &&= \left(\frac{n}{\tau+1} - \frac{n-2}{2}\right)\int_\Omega |u|^{\tau+1}dx \geq 0,
\end{aligned}
$$

which yields $\frac{n}{\tau+1} \geq \frac{n-2}{2}$ and hence a contradiction: $\tau \leq \frac{n+2}{n-2}$.

If u is a positive $C^2(\bar{\Omega})$ solution of (4.63) with $\tau = \frac{n+2}{n-2}$. Since $u = 0$ on $\partial\Omega$, by a sharp maximum principle (see **Serrin's Maximum Principle**, Lemma 4.32 below), we have $u_\nu \leq -\sigma$ on $\partial\Omega$ for a positive number σ. By (4.65), we obtain

$$\sigma^2\int_{\partial\Omega} x\cdot\nu\, dS \leq \int_{\partial\Omega} u_\nu^2(x\cdot\nu)\, dS = 0,$$

from which we have a desired contradiction:

$$0 = \int_{\partial\Omega} x\cdot\nu\, dS = \int_\Omega \Delta(\frac{|x|^2}{2})\, dx = n|\Omega|.$$

This completes the proof. □

4.6.4. Radial Symmetry of Solutions in a Ball. Let $\Omega = B$ be the open unit ball in $\mathbb{R}^n$. We shall study the nonlinear problem

$$(4.66) \qquad \Delta u + f(u) = 0 \text{ in } \Omega, \quad u|_{\partial\Omega} = 0.$$

We are interested in positive solutions: $u > 0$ in Ω. Assume that $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz, but is otherwise arbitrary. Our intention is to prove that u is necessarily radial, i.e., $u(x)$ depends only on $r = |x|$. This is the famous theorem of Gidas, Ni & Nirenberg and is a quite remarkable conclusion, as we are making essentially no assumptions on the nonlinearity.

The technique of proof is based on an extension of the maximum principle and the **method of moving planes**.

Lemma 4.32. (Serrin's Maximum Principle) *Let Ω be a bounded domain in $\mathbb{R}^n$. Assume $u \in C^2(\bar{\Omega})$ satisfies*

$$\Delta u + a(x)u \leq 0, \;\; u \geq 0 \;(\not\equiv 0) \text{ in } \Omega$$

where $a(x)$ is bounded. If $x_0 \in \partial\Omega$, $u(x_0) = 0$, and Ω satisfies the interior ball condition at x_0, then the normal derivative $u_n(x_0) < 0$. Moreover, $u > 0$ in Ω.

Proof. Set $a = a^+ + a^-$, where $a^+ = \max(a, 0)$, $a^- = \min(a, 0)$. Thus

$$\Delta u + a^-(x)u \leq -a^+u \leq 0 \text{ in } \Omega$$

and the conclusions follows from the strong maximum principle and the boundary point principle. □

4.6.5. Moving Plane Method. Let Ω_+ denote the open upper half ball $\Omega \cap \{x_n > 0\}$, and Ω_- the open lower half ball $\Omega \cap \{x_n < 0\}$.

Lemma 4.33. *Let $u \in C^2(\bar\Omega)$ be a positive solution of* (4.66). *Then*

$$u_{x_n} < 0 \text{ in } \Omega \text{ near } \partial\Omega_+.$$

Proof. Fix any point $x_0 \in \partial\Omega_+$ and let $\nu = (\nu_1, ..., \nu_n)$ denote the outer unit normal to $\partial\Omega_+$ at x_0. Note that $\nu_n > 0$. We claim that $u_{x_n} < 0$ in Ω near x_0.

We shall give the proof under the assumption $f(0) \geq 0$. Then

$$\begin{aligned} 0 &= -\Delta u - f(u) + f(0) - f(0) \\ &\leq -\Delta u - \int_0^1 \frac{\partial}{\partial s} f(su(x))ds \\ &\leq -\Delta u + cu \end{aligned}$$

for $c(x) = -\int_0^1 f'(su(x))ds$. According to Lemma 4.32, $u_\nu(x_0) = \nabla u(x_0) \cdot \nu < 0$. Since ∇u is parallel to ν on $\partial\Omega$ and $\nu_n > 0$, we conclude that $u_{x_n}(x_0) < 0$, and thus $u_{x_n} < 0$ in Ω near x_0. □

Theorem 4.34. *Let $u \in C^2(\bar B)$ be a positive solution of* (4.66). *Then u is radial, i.e.,*

$$u(x) = v(r) \quad (r = |x|)$$

for some strictly decreasing function $v : [0, 1] \to [0, \infty)$.

Proof. We apply the method of moving planes, following the important work of Gidas, Ni & Nirenberg.

For $0 \leq \lambda \leq 1$, set $P_\lambda \equiv \{x \in \mathbb{R}^n : x_n = \lambda\}$. For $x = (x_1, \cdots, x_n) \in \bar B$, let

$$x_\lambda \equiv (x_1, \ldots, x_{n-1}, 2\lambda - x_n)$$

denote the reflection of x with respect to P_λ. Let $E_\lambda \equiv \{x \in \Omega : \lambda < x_n < 1\}$. For each $0 \leq \lambda < 1$, consider the statement

$$u(x) < u(x_\lambda) \quad \text{for all} \quad x \in E_\lambda. \tag{4.67}$$

According to Lemma 4.33, we see that this statement is valid for each $\lambda < 1$, λ sufficiently close to 1. Set

$$\lambda_0 \equiv \inf\{\mu \in [0, 1) \colon (4.67) \text{ holds for each } \lambda \in [\mu, 1)\}.$$

We claim that $\lambda_0 = 0$. Suppose to the contrary that $\lambda_0 > 0$. Set

$$w(x) = u(x_{\lambda_0}) - u(x) \quad (x \in E_{\lambda_0}).$$

Then

$$-\Delta w = f(u(x_{\lambda_0})) - f(u(x)) = -cw \text{ in } E_{\lambda_0}$$

for

$$c(x) = -\int_0^1 f'(su(x_{\lambda_0}) + (1 - s)u(x))dx.$$

As $w \geq 0$ in E_{λ_0}, we deduce from Lemma 4.32 (applied to E_{λ_0}) that

$$w > 0 \text{ in } E_{\lambda_0}, \quad w_{x_n} > 0 \text{ on } P_{\lambda_0} \cap \Omega.$$

Thus

$$u(x) < u(x_{\lambda_0}) \text{ in } E_{\lambda_0} \tag{4.68}$$

and

$$u_{x_n} < 0 \text{ on } P_{\lambda_0} \cap \Omega. \tag{4.69}$$

Using (4.69) and Lemma 4.33, we conclude

$$u(x_n) < 0 \text{ on } P_{\lambda_0 - \varepsilon} \cap \Omega \quad \text{for all} \quad 0 \leq \varepsilon \leq \varepsilon_0 \tag{4.70}$$

if ε_0 is sufficiently small. Then (4.68) and the continuity of u imply

$$u(x) < u(x_{\lambda_0 - \varepsilon}) \text{ in } E_{\lambda_0 - \varepsilon} \quad \text{for all} \quad 0 \leq \varepsilon \leq \varepsilon_0 \tag{4.71}$$

if ε_0 is small enough. Assertion (4.71) contradicts our choice of λ_0.

Since $\lambda_0 = 0$, we see that

$$u(x_1, \ldots, x_{n-1}, -x_n) \geq u(x_1, \ldots, x_n) \quad \text{for all} \quad x \in \Omega_+.$$

A similar argument in Ω_- proves

$$u(x_1, \ldots, x_{n-1}, -x_n) \leq u(x_1, \ldots, x_n) \quad \text{for all} \quad x \in \Omega_+.$$

Thus u is symmetric with respect to the plane P_0 and $u_{x_n} = 0$ on P_0.

This argument applies as well after any rotation of coordinate axes, and so the theorem follows. $\square$

Chapter 5

Weak Lower Semicontinuity on Sobolev Spaces

As discussed before, in many variational problems, weak lower semicontinuity is essential for using direct method to establish the existence of minimizers. In this chapter, we study the conditions for weak lower semicontinuity of a multiple integral functional $I(u)$ on the Sobolev space $W^{1,p}(\Omega;\mathbb{R}^N)$. Assume

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

Recall that I is called **weakly lower semicontinuous** on $W^{1,p}(\Omega;\mathbb{R}^N)$ if

$$I(\bar u) \le \liminf_{\nu\to\infty} I(u_\nu) \quad \text{whenever } u_\nu \rightharpoonup \bar u \text{ weakly on } W^{1,p}(\Omega;\mathbb{R}^N). \tag{5.1}$$

5.1. The Convex Case

5.1.1. Tonelli's Theorem. We first prove a semicontinuity result of Tonelli.

Theorem 5.1. *Let $F(x,s,\xi)\ge 0$ be* **smooth** *and* **convex** *in ξ. Assume F, F_ξ are both* continuous *in (x,s,ξ). Then the functional $I(u)$ defined above is* sequentially weakly (weakly star if $p=\infty$) lower semicontinuous *on $W^{1,p}(\Omega;\mathbb{R}^N)$ for all $1\le p\le\infty$.*

Proof. We need only to prove $I(u)$ is *w.l.s.c.* on $W^{1,1}(\Omega;\mathbb{R}^N)$. To this end, assume $\{u_\nu\}$ is a sequence weakly convergent to u in $W^{1,1}(\Omega;\mathbb{R}^N)$. We need to show

$$I(u) \le \liminf_{\nu\to\infty} I(u_\nu).$$

By the Sobolev embedding theorem it follows that (via a subsequence) $u_\nu \to u$ in $L^1(\Omega;\mathbb{R}^N)$. We can also assume $u_\nu(x)\to u(x)$ for almost every $x\in\Omega$. Now, for any given $\delta>0$ we choose a compact set $K\subset\Omega$ such that

(i) $u_\nu\to u$ uniformly on K and $|\Omega\setminus K|<\delta$ (by Egorov*'s theorem*);

(ii) u, Du are continuous on K (by Lusin*'s theorem*).

Since $F(x,s,\xi)$ is *smooth* and *convex* in ξ, it follows that

$$F(x,s,\eta) \geq F(x,s,\xi) + F_{\xi_i^k}(x,s,\xi)\,(\eta_i^k - \xi_i^k) \quad \forall \xi,\ \eta \in \mathbb{M}^{N\times n}.$$

Therefore, since $F \geq 0$,

$$\begin{aligned}
I(u_\nu) &\geq \int_K F(x,u_\nu,Du_\nu)\,dx \\
&\geq \int_K \Big[F(x,u_\nu,Du) + F_{\xi_i^k}(x,u_\nu,Du)\,(D_iu_\nu^k - D_iu^k)\Big] \\
&= \int_K F(x,u_\nu,Du) + \int_K F_{\xi_i^k}(x,u,Du)\,(D_iu_\nu^k - D_iu^k) \\
&+ \int_K [F_{\xi_i^k}(x,u_\nu,Du) - F_{\xi_i^k}(x,u,Du)]\,(D_iu_\nu^k - D_iu^k).
\end{aligned}$$

Since $F(x,s,\xi)$ and $F_\xi(x,s,\xi)$ are both uniformly continuous on bounded sets and $u_\nu(x) \to u(x)$ uniformly on K we have

$$\lim_{\nu\to\infty} \int_K F(x,u_\nu,Du)\,dx = \int_K F(x,u,Du)\,dx,$$

$$\lim_{\nu\to\infty} \|F_{\xi_i^k}(x,u_\nu,Du) - F_{\xi_i^k}(x,u,Du)\|_{L^\infty(K)} = 0.$$

Now since $F_{\xi_i^k}(x,u,Du)$ is bounded on K and $D_iu_\nu^k \rightharpoonup D_iu^k$ weakly in $L^1(\Omega)$ as $\nu \to \infty$, we thus have

$$\lim_{\nu\to\infty} \int_K F_{\xi_i^k}(x,u,Du)\,(D_iu_\nu^k - D_iu^k)\,dx = 0.$$

From these estimates, noting that, for any two sequences $\{a_\nu\}, \{b_\nu\}$,

$$\liminf_{\nu\to\infty}(a_\nu + b_\nu) \geq \liminf_{\nu\to\infty} a_\nu + \liminf_{\nu\to\infty} b_\nu, \tag{5.2}$$

we have

$$\liminf_{\nu\to\infty} I(u_\nu) \geq \int_K F(x,u,Du)\,dx.$$

If $F(x,u,Du) \in L^1(\Omega)$, i.e., $I(u) < \infty$, then for any given $\epsilon > 0$, we use *Lebesgue's absolute continuity theorem* to determine $\delta > 0$ so that

$$\int_E F(x,u,Du)dx \geq \int_\Omega F(x,u,Du)dx - \epsilon, \quad \forall E \subset \Omega,\ |\Omega \setminus E| < \delta.$$

On the other hand, if $I(u) = \infty$ then for any given large number $M > 0$ we choose $\delta > 0$ so that

$$\int_E F(x,u,Du)\,dx > M, \quad \forall E \subset \Omega,\ |\Omega \setminus E| < \delta.$$

In any of these two cases, using this $\delta > 0$ with $E = K$, where K is determined according to (i) and (ii) above, we conclude, by setting either $\epsilon \to 0$ or $M \to \infty$, that

$$\liminf_{\nu\to\infty} I(u_\nu) \geq I(u).$$

The theorem is proved. □

5.1.2. Existence in the Convex Case. Using the theorem, we obtain the following existence result for convex functionals.

Theorem 5.2. *In addition to the hypotheses of the previous theorem, assume there exists* $1 < p < \infty$ *such that*

$$F(x, s, \xi) \geq c\,|\xi|^p - C(x),$$

where $c > 0$, $C \in L^1(\Omega)$ *are given. If for some* $\varphi \in W^{1,p}(\Omega; \mathbb{R}^N)$, $I(\varphi) < \infty$, *then minimization problem* $\inf_{u \in \mathcal{D}_\varphi} I(u)$ *has a minimizer in the Dirichlet class* $\mathcal{D}_\varphi$.

Proof. This follows from the abstract existence Theorem 4.7 in the previous chapter. □

Remark. Both theorems in this section hold for more general functions $F(x, s, \xi)$. For example, we can replace the continuity condition by the *Carathéodory* condition.

5.2. Morrey's Quasiconvexity

In this section, we will derive a condition which, under the mild general assumption, will be the "right" (*necessary and sufficient*) condition for the weak lower semicontinuity for integral functionals on Sobolev spaces. This will be Morrey's **quasiconvexity** condition; see Morrey [**15, 16**]. Please be aware that there is at least one other *quasiconvexity* in the analysis that has a totally different meaning.

5.2.1. Lipschitz Convergence. Note that $W^{1,\infty}(\Omega; \mathbb{R}^N)$ can be identified with the space of all *Lipschitz maps* from Ω to $\mathbb{R}^N$. A sequence $\{u_\nu\}$ converges to u in the weak star topology of $W^{1,\infty}(\Omega; \mathbb{R}^N)$ if and only if $\{u_\nu\}$ converges to u in the sense of **Lipschitz convergence**; that is,

1) $u_\nu \to u$ uniformly in $C(\bar{\Omega}; \mathbb{R}^N)$;
2) the Lipschitz norms of u_ν and u are bounded.

5.2.2. Quasiconvexity as Necessary Condition. The following result, mainly due to Morrey [**15**], gives the necessary condition for the lower semicontinuity under the *Lipschitz convergence* of the multiple integral

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

Theorem 5.3. (Morrey) *Assume* $F(x, s, \xi)$ *is continuous on* $\bar{\Omega} \times \mathbb{R}^N \times \mathbb{M}^{N \times n}$. *Assume the functional* $I(u)$ *is* s.l.s.c. *with respect to Lipschitz convergence on* $W^{1,\infty}(\Omega; \mathbb{R}^N)$. *Then the following condition holds for all* $x_0 \in \Omega$, $s_0 \in \mathbb{R}^N$, $\xi_0 \in \mathbb{M}^{N \times n}$ *and all* $\phi \in W_0^{1,\infty}(\Omega; \mathbb{R}^N)$:

$$F(x_0, s_0, \xi_0) \leq \fint_\Omega F(x_0, s_0, \xi_0 + D\phi(x))\, dx. \tag{5.3}$$

Proof. Let $\phi \in W_0^{1,\infty}(\Omega; \mathbb{R}^N)$ be given. Let Q be a fixed open cube containing $\bar{\Omega}$ with center $\bar{x}$ and side-length $2L$. We extend ϕ by zero onto Q; then $\phi \in W_0^{1,\infty}(Q; \mathbb{R}^N)$.

Let $x_0 \in \Omega$, $s_0 \in \mathbb{R}^N$ and $\xi_0 \in \mathbb{M}^{N \times n}$ be given. Let $\tilde{u}(x) = s_0 + \xi_0 \cdot (x - x_0)$. Assume $Q' \subset\subset \Omega$ is an arbitrarily given cube containing x_0 with side-length $2l$. For any positive integer ν we divide each side of Q' into 2^ν intervals of equal length, each being equal to

$2^{-\nu+1}l$. This divides Q' into $2^{n\nu}$ small cubes $\{Q_j^\nu\}$ with $j=1,2,...,2^{n\nu}$. Denote the center of each cube Q_j^ν by $\bar{x}_j^\nu$ and define a function $u_\nu\colon \Omega\to\mathbb{R}^N$ as follows:

$$u_\nu(x)=\begin{cases} \tilde{u}(x) & \text{if } x\in\Omega\setminus\cup_{j=1}^{2^{n\nu}}Q_j^\nu;\\ \tilde{u}(x)+\frac{2^{-\nu}l}{L}\,\phi\big(\bar{x}+\frac{2^\nu L}{l}\,(x-\bar{x}_j^\nu)\big) & \text{if } x\in Q_j^\nu,\ 1\le j\le 2^{n\nu}.\end{cases}$$

We easily see that $u_\nu\to\tilde{u}$ uniformly on Ω. Moreover,

$$Du_\nu(x)=\begin{cases} \xi_0 & \text{if } x\in\Omega\setminus\cup_{j=1}^{2^{n\nu}}Q_j^\nu;\\ \xi_0+D\phi\big(\bar{x}+\frac{2^\nu L}{l}\,(x-\bar{x}_j^\nu)\big) & \text{if } x\in Q_j^\nu,\ 1\le j\le 2^{n\nu}.\end{cases}$$

Therefore $\{Du_\nu\}$ is uniformly bounded and, by definition, $\{u_\nu\}$ converges to $\tilde{u}$ as $\nu\to\infty$ in the sense of Lipschitz convergence. Note that

$$I(\tilde{u})=\int_\Omega F(x,\tilde{u}(x),\xi_0)\,dx$$

and that

$$\begin{aligned} I(u_\nu)&=\int_\Omega F(x,u_\nu(x),Du_\nu)\,dx\\ &=\int_{\Omega\setminus Q'}F(x,\tilde{u},\xi_0)\,dx+\int_{Q'}F(x,u_\nu,Du_\nu)\,dx.\end{aligned}$$

Therefore, by the lower semicontinuity of I, we have

$$\int_{Q'}F(x,\tilde{u},\xi_0)\,dx\le\liminf_{\nu\to\infty}\int_{Q'}F(x,u_\nu,Du_\nu)\,dx. \tag{5.4}$$

From the uniform continuity of $F(x,s,\xi)$ on bounded sets and the fact that $u_\nu\to\tilde{u}$ uniformly on Ω we have

$$\liminf_{\nu\to\infty}\int_{Q'}F(x,u_\nu,Du_\nu)\,dx=\liminf_{\nu\to\infty}\int_{Q'}F(x,\tilde{u},Du_\nu)\,dx. \tag{5.5}$$

We now compute

$$\begin{aligned}\int_{Q'}F(x,\tilde{u},Du_\nu)\,dx&=\sum_{j=1}^{2^{n\nu}}\int_{Q_j^\nu}F\Big(x,\tilde{u},\xi_0+D\phi\big(\bar{x}+\frac{2^\nu L}{l}\,(x-\bar{x}_j^\nu)\big)\Big)\,dx\\ &=\sum_{j=1}^{2^{n\nu}}\int_{Q_j^\nu}F\Big(\bar{x}_j^\nu,\tilde{u}(\bar{x}_j^\nu),\xi_0+D\phi\big(\bar{x}+\frac{2^\nu L}{l}\,(x-\bar{x}_j^\nu)\big)\Big)\,dx+o(1)\\ &=\sum_{j=1}^{2^{n\nu}}\Big(\frac{l}{2^\nu L}\Big)^n\int_Q F\big(\bar{x}_j^\nu,\tilde{u}(\bar{x}_j^\nu),\xi_0+D\phi(y)\big)\,dy+o(1)\end{aligned}$$

$$=\sum_{j=1}^{2^{n\nu}}\tilde{F}(\bar{x}_j^\nu)\,|Q_j^\nu|+o(1), \tag{5.6}$$

where $o(1)\to 0$ as $\nu\to\infty$, and

$$\tilde{F}(x)=\fint_Q F(x,\tilde{u}(x),\xi_0+D\phi(y))\,dy.$$

This function is continuous on Q' and the sum in (5.6) is simply the *Riemann sum* of the integral of $\tilde{F}$ over Q'. Therefore, we arrive at

$$\lim_{\nu\to\infty}\int_{Q'}F(x,\tilde{u},Du_\nu)\,dx=\int_{Q'}\tilde{F}(x)\,dx,$$

which by (5.5) implies
$$\int_{Q'} F(x, \tilde{u}(x), \xi_0)\, dx \le \int_{Q'} \tilde{F}(x)\, dx.$$
This inequality holds for any cube $Q' \subset\subset \Omega$ containing x_0; therefore,
$$F(x_0, \tilde{u}(x_0), \xi_0) \le \tilde{F}(x_0).$$
This is nothing but
$$F(x_0, s_0, \xi_0) \le \fint_Q F(x_0, s_0, \xi_0 + D\phi(y))\, dy.$$
From this (5.3) follows since $\phi = 0$ on $Q \setminus \Omega$; hence the proof of Theorem 5.3 is complete. □

Motivated by this theorem, we have the following definition of **quasiconvex functions** in the sense of Morrey.

Definition. A function $F\colon \mathbb{M}^{N\times n} \to \bar{\mathbb{R}}$ is called **quasiconvex** (in the sense of Morrey) if

$$F(\xi) \le \fint_\Omega F(\xi + D\phi(x))\, dx \tag{5.7}$$

holds for all $\phi \in W_0^{1,\infty}(\Omega; \mathbb{R}^N)$.

The following result, due to N. Meyers, will be useful to relax the zero boundary condition on ϕ.

Theorem 5.4. *Let $F\colon \mathbb{M}^{N\times n} \to \mathbb{R}$ be continuous and quasiconvex. For every bounded set $Q \subset \mathbb{R}^n$ and every sequence $\{z_\nu\}$ in $W^{1,\infty}(Q; \mathbb{R}^N)$ converging to zero in the sense of Lipschitz convergence, we have*
$$F(\xi) \le \liminf_{\nu\to\infty} \fint_Q F(\xi + Dz_\nu(x))\, dx$$
for every $\xi \in \mathbb{M}^{N\times n}$.

Proof. Let $Q_k = \{x \in Q \mid \operatorname{dist}(x, \partial Q) > 1/k\}$. Then $Q_k \subset\subset Q$ and $|Q \setminus Q_k| \to 0$ as $k \to \infty$. Choose a cut-off function $\zeta_k \in C_0^\infty(Q)$ such that
$$0 \le \zeta_k \le 1, \quad \zeta_k\big|_{Q_k} = 1, \quad M_k = \|D\zeta_k\|_{L^\infty} < \infty.$$
Since $z_\nu \to 0$ uniformly on Q we can choose a subsequence $\{\nu_k\}$ such that
$$\|z_{\nu_k}\|_{L^\infty} \le (M_k + 1)^{-1} \quad \forall\, k = 1, 2, \ldots$$
and we may also assume
$$\lim_{k\to\infty} \int_Q F(\xi + Dz_{\nu_k}(x))\, dx = \liminf_{\nu\to\infty} \int_Q F(\xi + Dz_\nu(x))\, dx.$$
Define $\phi_k = \zeta_k\, z_{\nu_k}$. Then $\phi_k \in W_0^{1,\infty}(Q; \mathbb{R}^N)$ and we can use them as test functions in the definition of quasiconvexity to obtain
$$\begin{aligned} |Q|\, F(\xi) &\le \int_Q F(\xi + D\phi_k(x))\, dx \\ &= \int_{Q_k} F(\xi + Dz_{\nu_k}) + \int_{Q\setminus Q_k} F(\xi + \zeta_k\, Dz_{\nu_k} + z_{\nu_k} \otimes D\zeta_k) \\ &= \int_Q F(\xi + Dz_{\nu_k}(x))\, dx + \epsilon_k, \end{aligned}$$

where

$$\epsilon_k = \int_{Q\setminus Q_k} [F(\xi + \zeta_k\, Dz_{\nu_k} + z_{\nu_k} \otimes D\zeta_k u) - F(\xi + Dz_{\nu_k}(x))]\, dx.$$

Since $F(\xi)$ is bounded on bounded sets and $|Q \setminus Q_k| \to 0$ as $k \to \infty$, we easily have $\epsilon_k \to 0$ as $k \to \infty$. Therefore,

$$|Q|\, F(\xi) \leq \liminf_{\nu\to\infty} \int_Q F(\xi + Dz_\nu(x))\, dx.$$

This completes the proof. □

5.2.3. Quasiconvexity as Sufficient Condition. We now prove the sufficiency of quasiconvexity for the lower semicontinuity of the functional

$$I(u) = \int_\Omega F(x, u(x), Du(x))\, dx$$

under the Lipschitz convergence on Ω.

Theorem 5.5. (Morrey) *Assume $F(x, s, \xi)$ is continuous on $\bar{\Omega} \times \mathbb{R}^N \times \mathbb{M}^{N\times n}$ and is* quasiconvex *in ξ. Then the functional I defined above is s.l.s.c. with respect to Lipschitz convergence on Ω.*

Proof. Let $\{z_k\}$ be any sequence converging to 0 in the sense of Lipschitz convergence on Ω, and let $u \in W^{1,\infty}(\Omega; \mathbb{R}^N)$ be any given function. We need to show

$$\int_\Omega F(x, u, Du) \leq \liminf_{k\to\infty} \int_\Omega F(x, u + z_k, Du + Dz_k). \tag{5.8}$$

For any given $\epsilon > 0$, since both function $F(x, u, Du)$ and sequence $\{F(x, u+z_k, Du+Dz_k)\}$ are bounded, we choose finitely many disjoint cubes Q_j contained in Ω such that

$$I(u) \leq \int_{\cup Q_j} F(x, u, Du)\, dx + \epsilon$$

and

$$I(u + z_k) \geq \int_{\cup Q_j} F(x, u + z_k, Du + Dz_k)\, dx - \epsilon,$$

for all $k = 1, 2, \cdots$. In what follows, we prove for each cube $Q = Q_j$

$$I_Q(u) \equiv \int_Q F(x, u, Du)\, dx \leq \liminf_{k\to\infty} I_Q(u + z_k).$$

This, by (5.2), will certainly imply the conclusion of the theorem. To this end, for each positive integer ν, we divide Q into small cubes $\{Q_j^\nu\}$ with center $\bar{x}_j^\nu$ as in the proof of Theorem 5.3:

$$Q = \bigcup_{j=1}^{2^{n\nu}} Q_j^\nu \cup E, \quad |E| = 0.$$

Define

$$(u)_j^\nu = \fint_{Q_j^\nu} u(x)\, dx, \quad (Du)_j^\nu = \fint_{Q_j^\nu} Du(x)\, dx,$$

and

$$U^\nu(x) = \sum_{j=1}^{2^{n\nu}} (u)_j^\nu \cdot \chi_{Q_j^\nu}(x), \quad M^\nu = \sum_{j=1}^{2^{n\nu}} (Du)_j^\nu \cdot \chi_{Q_j^\nu}(x).$$

Note that

$$\|U^\nu\|_{L^\infty} + \|M^\nu\|_{L^\infty} \le \|u\|_{W^{1,\infty}}$$

and that the sequences $\{U^\nu\}$ and $\{M^\nu\}$ converge almost everywhere to u and Du on Q as $\nu \to \infty$, respectively. We now estimate $I_Q(u + z_k)$.

$$I_Q(u+z_k) = \int_Q F(x, u+z_k, Du+Dz_k) = a_k + b_k^\nu + c_k^\nu + d^\nu + I_Q(u),$$

where

$$\begin{aligned}
a_k &= \int_Q [F(x, u+z_k, Du+Dz_k) - F(x, u, Du+Dz_k)]\,dx,\\
b_k^\nu &= \sum_{j=1}^{2^{n\nu}} \int_{Q_j^\nu} [F(x, u, Du+Dz_k) - F(\bar{x}_j^\nu, (u)_j^\nu, (Du)_j^\nu + Dz_k)]\,dx,\\
c_k^\nu &= \sum_{j=1}^{2^{n\nu}} \int_{Q_j^\nu} [F(\bar{x}_j^\nu, (u)_j^\nu, (Du)_j^\nu + Dz_k) - F(\bar{x}_j^\nu, (u)_j^\nu, (Du)_j^\nu)]\,dx,\\
d^\nu &= \sum_{j=1}^{2^{n\nu}} \int_{Q_j^\nu} [F(\bar{x}_j^\nu, (u)_j^\nu, (Du)_j^\nu) - F(x, u, Du)]\,dx.
\end{aligned}$$

By the uniform continuity of $F(x, s, \xi)$ on bounded sets and the pointwise convergence of $\{U^\nu\}$ and $\{M^\nu\}$ we have

$$\lim_{k\to\infty} a_k = 0, \quad \lim_{\nu\to\infty} d^\nu = 0$$

and $\lim_{\nu\to\infty} b_k^\nu = 0$ uniformly with respect to k. We apply Theorem 5.4 to each Q_j^ν to obtain, by (5.2),

$$\liminf_{k\to\infty} c_k^\nu \ge 0$$

for all $\nu = 1, 2, \cdots$. Therefore, again by (5.2),

$$\liminf_{k\to\infty} I_Q(u+z_k) \ge I_Q(u),$$

as desired. The proof is complete. □

5.2.4. Weak Lower Semicontinuity on Sobolev Spaces. Quasiconvexity is also the "right" condition for weak lower semicontinuity of integral functionals on $W^{1,p}(\Omega; \mathbb{R}^N)$. A most general theorem in this direction is the following theorem due to Acerbi & Fusco [**1**]; see the reference for proof.

Theorem 5.6. (Acerbi & Fusco) *Let $F(x, s, \xi)$ be a* Carathéodory *function. Assume for some $1 \le p < \infty$*

$$0 \le F(x, s, \xi) \le a(x) + C\,(|s|^p + |\xi|^p),$$

where $C > 0$ is a constant and $a(x) \ge 0$ is a locally integrable *function in Ω. Then functional $I(u) = \int_\Omega F(x, u, Du)\,dx$ is w.s.l.s.c. on $W^{1,p}(\Omega; \mathbb{R}^N)$* **if and only if** *$F(x, s, \xi)$ is quasiconvex in ξ.*

5.2.5. Existence in the General Case.

Theorem 5.7. (Existence of minimizers) *Let $F(x,s,\xi)$ be* Carathéodory *and* quasiconvex *in ξ and satisfy*

$$\max\{0,\, c\,|\xi|^p - C(x)\} \le F(x,s,\xi) \le a(x) + C\,(|s|^p + |\xi|^p)$$

for some $1 < p < \infty$, where $c > 0$ is a constant and $C(x)$, $a(x)$ are given integrable functions in Ω. Then, for any $\varphi \in W^{1,p}(\Omega;\mathbb{R}^N)$, the minimization problem

$$\min_{u\in\mathcal{D}_\varphi} \int_\Omega F(x,u(x),Du(x))\,dx$$

has a minimizer in the Dirichlet class $\mathcal{D}_\varphi$.

5.3. Properties of Quasiconvex Functions

5.3.1. Domain Independence. We prove that quasiconvexity is independent of the domain Ω.

Theorem 5.8. *Let $F\colon \mathbb{M}^{N\times n} \to \bar{\mathbb{R}}$ be such that (5.7) holds for all $\phi \in W_0^{1,\infty}(\Omega;\mathbb{R}^N)$. Then for any bounded open set $G \subset \mathbb{R}^n$ with $|\partial G| = 0$ one has*

$$F(\xi) \le \fint_G F(\xi + D\psi(y))\,dy, \quad \forall \xi \in \mathbb{M}^{N\times n} \tag{5.9}$$

holds for all $\psi \in W_0^{1,\infty}(G;\mathbb{R}^N)$.

Proof. Let $G \subset \mathbb{R}^n$ be any bounded open set with $|\partial G| = 0$, and $\psi \in W_0^{1,\infty}(G;\mathbb{R}^N)$. Assume $\bar{y} \in G$. For any $x \in \Omega$ and $\epsilon > 0$ let

$$G(x,\epsilon) = \{z \in \mathbb{R}^n \mid z = x + \epsilon\,(y - \bar{y}) \text{ for some } y \in G\}.$$

Then there exists an $\epsilon_x > 0$ such that $x \in \overline{G(x,\epsilon)} \subset \Omega$ for all $x \in \Omega$ and $0 < \epsilon < \epsilon_x$. This means the family

$$\{\overline{G(x,\epsilon)} \mid x \in \Omega,\ 0 < \epsilon < \epsilon_x\}$$

covers Ω in the sense of **Vitali covering**. Therefore, there exists a countable *disjoint* subfamily $\{G(x_j,\epsilon_j)\}$ and a set E of measure zero such that

$$\Omega = \bigcup_{j=1}^\infty \overline{G(x_j,\epsilon_j)} \cup E. \tag{5.10}$$

We now define a function $\phi\colon \Omega \to \mathbb{R}^N$ as follows.

$$\phi(x) = \begin{cases} 0 & \text{if } x \in \bigcup_{j=1}^\infty \partial[G(x_j,\epsilon_j)] \cup E, \\ \epsilon_j\,\psi\big(\bar{y} + \frac{x-x_j}{\epsilon_j}\big) & \text{if } x \in G(x_j,\epsilon_j) \text{ for some } j. \end{cases}$$

One can verify that $\phi \in W_0^{1,\infty}(\Omega;\mathbb{R}^N)$ and

$$D\phi(x) = D\psi\big(\bar{y} + \frac{x-x_j}{\epsilon_j}\big) \quad \forall x \in G(x_j,\epsilon_j).$$

Therefore, from (5.7), it follows that

$$\begin{aligned} F(\xi)\,|\Omega| &\le \int_\Omega F(\xi + D\phi(x))\,dx \\ &= \sum_{j=1}^\infty \int_{G(x_j,\epsilon_j)} F\left(\xi + D\psi\big(\bar{y} + \frac{x-x_j}{\epsilon_j}\big)\right)\,dx \end{aligned}$$

$$= \sum_{j=1}^{\infty} \epsilon_j^n \int_G F(\xi + D\psi(y))\, dy = \frac{|\Omega|}{|G|} \int_G F(\xi + D\psi(y))\, dy,$$

where the last equality follows since, by (5.10), $\sum_{j=1}^{\infty} \epsilon_j^n = |\Omega|/|G|$. We have thus proved (5.9). $\square$

In the following, let Σ be the unit cube in $\mathbb{R}^n$; that is

$$\Sigma = \{x \in \mathbb{R}^n \mid 0 < x_\alpha < 1,\ \alpha = 1, 2, \cdots, n\}.$$

Note that $|\Sigma| = 1$.

Let $f\colon \mathbb{R}^n \to \mathbb{R}^N$ be any given function. We say f is **Σ-periodic** if $f(\cdots, x_\alpha, \cdots)$ is 1-periodic in x_α for all $\alpha = 1, 2, \cdots, n$. Quasiconvexity can be also characterized by the following condition.

Theorem 5.9. *Let $F\colon \mathbb{M}^{N\times n} \to \mathbb{R}$ be continuous. Then F is quasiconvex if and only if*

$$F(\xi) \le \int_\Sigma F(\xi + D\phi(x))\, dx \quad \forall \xi \in \mathbb{M}^{N\times n} \tag{5.11}$$

for all Σ-periodic Lipschitz functions $\phi \in W^{1,\infty}(\mathbb{R}^n; \mathbb{R}^N)$.

Proof. Since any function $\psi \in W_0^{1,\infty}(\Sigma; \mathbb{R}^N)$ can be extended as a Σ-periodic function on $\mathbb{R}^n$, we easily see that (5.11) implies (5.9) for $G = \Sigma$ thus the quasiconvexity of F. We have only to prove (5.11) holds if F is quasiconvex. Let $\phi \in W^{1,\infty}(\mathbb{R}^n; \mathbb{R}^N)$ be a Σ-periodic function. Define

$$\phi_j(x) = \frac{1}{j}\, \phi(jx)$$

for all $j = 1, 2, \cdots$. It is easily seen that $\phi_j \to 0$ in the sense of *Lipschitz convergence* on $W^{1,\infty}(\Sigma; \mathbb{R}^N)$. Therefore the theorem of Meyers, Theorem 5.4, and the quasiconvexity of F implies

$$F(\xi) \le \liminf_{j\to\infty} \int_\Sigma F(\xi + D\phi_j(x))\, dx.$$

Note that

$$\begin{aligned} \int_\Sigma F(\xi + D\phi_j(x))\, dx &= \int_\Sigma F(\xi + D\phi(jx))\, dx \\ &= j^{-n} \int_{j\Sigma} F(\xi + D\phi(y))\, dy \end{aligned}$$

and that, besides a set of measure zero,

$$j\Sigma = \bigcup_{\nu=1}^{j^n} (\bar{x}_\nu + \Sigma),$$

where $\bar{x}_\nu$ are the left-lower corner points of the subcubes obtained by dividing the sides of $j\Sigma$ into j-equal subintervals. Since $D\phi(x)$ is Σ-periodic, we thus have

$$\int_{j\Sigma} F(\xi + D\phi(y))dy = \sum_{\nu=1}^{j^n} \int_{\bar{x}_\nu+\Sigma} F(\xi + D\phi(y))dy = j^n \int_\Sigma F(\xi + D\phi(x))dx,$$

and therefore

$$F(\xi) \le \int_\Sigma F(\xi + D\phi(x))\, dx$$

as needed; the proof is complete. $\square$

5.3.2. Convexity vs Quasiconvexity.

Lemma 5.10. (Jensen's inequality) *Let (E,μ) be a measure space with total mass $\mu(E)=1$ and let $h\colon E\to\mathbb{R}^L$ be an integrable function on E. If $G\colon\mathbb{R}^L\to\mathbb{R}$ is a convex function, then*
$$G\left(\int_E h(x)\,d\mu\right)\le\int_E G(h(x))\,d\mu.$$

Proof. Let $F=\int_E h(x)\,d\mu$. Since $G\colon\mathbb{R}^L\to\mathbb{R}$ is convex, there exists $l_F\in\mathbb{R}^L$ (note that $l_F=DG(F)$ for almost every F) such that
$$G(A)\ge G(F)+l_F\cdot(A-F)\quad\forall A\in\mathbb{R}^L.$$
Then $G(h(x))\ge G(F)+l_F\cdot(h(x)-F)$ for all $x\in E$, and integrating over $x\in E$ yields
$$\int_E G(h(x))\,d\mu\ge G(F)+l_F\cdot\int_E(h(x)-F)\,d\mu=G(F),$$
which proves Jensen's inequality. □

Theorem 5.11. *If $F\colon\mathbb{M}^{N\times n}\to\mathbb{R}$ is convex, then F is quasiconvex.*

Proof. This follows easily from Jensen's inequality and the divergence theorem. □

5.3.3. Quasiconvexity vs Rank-one Convexity. Recall that $F\colon\mathbb{M}^{N\times n}\to\bar{\mathbb{R}}$ is **rank-one convex** if for any $\xi\in\mathbb{M}^{N\times n}$, $q\in\mathbb{R}^N$, $p\in\mathbb{R}^n$ the function $f(t)=F(\xi+tq\otimes p)$ is a convex function of $t\in\mathbb{R}$.

We have the following result.

Theorem 5.12. *Every finite-valued quasiconvex function is rank-one convex.*

Proof. Let F be quasiconvex. We need to show that for any $\xi\in\mathbb{M}^{N\times n}$, $q\in\mathbb{R}^N$, $p\in\mathbb{R}^n$ the function $f(t)=F(\xi+tq\otimes p)$ is a convex function of $t\in\mathbb{R}$; that is, for all $0<\theta<1$ and t, $s\in\mathbb{R}$,
$$f(\theta t+(1-\theta)s)\le\theta f(t)+(1-\theta)f(s). \tag{5.12}$$
which is equivalent to
$$F(\tilde\xi)\le\theta F(\tilde\xi+aq\otimes p)+(1-\theta)F(\tilde\xi+bq\otimes p), \tag{5.13}$$
where
$$\tilde\xi=\xi+[\theta t+(1-\theta)s]q\otimes p,\quad a=(1-\theta)(t-s),\quad b=\theta(s-t).$$
Assume $t>s$; so $a>0$ and $b<0$. Let $\zeta(\tau)$ be the periodic Lipschitz function of peroid 1 on $\mathbb{R}$ satisfying $\zeta(\tau)=a\tau$ for $0\le\tau\le\theta$ and $\zeta(\tau)=b(\tau-1)$ for $\theta\le\tau\le1$. Let G be a cube of *unit volume* which is bounded between two planes $\{x\cdot p=0\}$ and $\{x\cdot p=1\}$. For $x\in\mathbb{R}^n$, define
$$|x|_\infty=\max\{|x_i|\,|\,1\le i\le n\},\quad\delta(x)=\operatorname{dist}_\infty(x,\partial G)=\inf\{|x-y|_\infty\,|\,y\in\partial G\}.$$
Then $\delta\in W_0^{1,\infty}(G)$ and $\nabla\delta(x)\in\{\pm e_i\,|\,i=1,2,\cdots,n\}$. Define functions by
$$u_k(x)=k^{-1}\zeta(kx\cdot p)\,q,\quad\phi_k(x)=\min\{k^{-1}\zeta(kx\cdot p),\,\delta(x)\}\,q,\quad k=1,2,\cdots.$$
Note that $Du_k(x)\in\{aq\otimes p,\,bq\otimes p\}$, $\phi_k\in W_0^{1,\infty}(G;\mathbb{R}^N)$ and
$$\lim_{k\to\infty}|\{x\in G\,|\,Du_k(x)\ne D\phi_k(x)\}|=0.$$

Hence, by (5.7) and noting that $Du_k(x)$ and $D\phi_k(x)$ take only on finitely many values,

$$\begin{aligned} F(\tilde{\xi}) &\leq \lim_{k\to\infty} \int_G F(\tilde{\xi} + D\phi_k(x))\,dx \\ &= \lim_{k\to\infty} \int_G F(\tilde{\xi} + Du_k(x))\,dx \\ &= \theta F(\tilde{\xi} + aq\otimes p) + (1-\theta)F(\tilde{\xi} + bq\otimes p). \end{aligned}$$

This completes the proof. □

Remark. Another proof in the case of continuous F is as follows. Let F be quasiconvex. If F is of class C^2, then

$$f(t) = \int_\Omega F(\xi + tD\phi(x))\,dx$$

takes its minimum at $t = 0$. Therefore $f''(0) \geq 0$; that is,

$$\int_\Omega F_{\xi_i^k \xi_j^l}(\xi)\, D_i\phi^k(x) D_j\phi^l(x)\,dx \geq 0$$

for all $\phi \in C_0^\infty(\Omega; \mathbb{R}^N)$. This implies the *weak* Legendre-Hadamard condition:

$$F_{\xi_i^k \xi_j^l}(\xi)\, q^k q^l\, p_i p_j \geq 0 \quad \forall\, q \in \mathbb{R}^N,\ p \in \mathbb{R}^n,$$

which is equivalent to that F is *rank-one convex.* For a continuous F, let $F^\epsilon = F * \rho_\epsilon$ be the regularization of F. Then F^ϵ is of class C^∞ and can be shown to be quasiconvex, and hence F^ϵ satisfies (5.13) for all $\epsilon > 0$; letting $\epsilon \to 0$ yields that F satisfies (5.13) and thus is rank-one convex.

The following result shows that Theorem 5.12 does not hold for extended valued functions.

Example 5.13. Let $n \geq 2$ and let $A,\, B \in \mathbb{M}^{N\times n}$ be such that $\operatorname{rank}(A - B) = 1$. Define

$$F(\xi) = \begin{cases} 0 & \xi \in \{A,\, B\}, \\ \infty & \xi \notin \{A,\, B\}. \end{cases}$$

Then F is quasiconvex convex, but not rank-one convex.

Proof. The rank-one convexity of F would imply

$$0 \leq F(\lambda A + (1-\lambda)B) \leq \lambda F(A) + (1-\lambda)F(B) = 0$$

for all $0 < \lambda < 1$. Hence F is not rank-one convex. To see F is quasiconvex, given $\xi \in \mathbb{M}^{N\times n}$ and $\phi \in W_0^{1,\infty}(\Sigma; \mathbb{R}^N)$, we need to show

$$F(\xi) \leq \int_\Sigma F(\xi + D\phi(x))\,dx. \tag{5.14}$$

Since the integral on the right-hand side takes only two values of $\{0, \infty\}$, we only need to prove the inequality when

$$\int_\Sigma F(\xi + D\phi(x))\,dx = 0.$$

In this case, we must have $\xi + D\phi(x) \in \{A,\, B\}$ for almost every $x \in \Sigma$. We claim, in this case, one must have $\xi \in \{A,\, B\}$ and hence (5.14) is valid. To prove this, let $\xi + D\phi(x) = \chi_E(x)A + (1 - \chi_E(x))B$, where $E = \{x \in \Sigma \,|\, \xi + D\phi(x) = A\}$. Integrating this over Σ yields $\xi = \lambda A + (1-\lambda)B$, where $\lambda = |E| \in [0,\, 1]$. Hence

$$D\phi(x) \in \{A - \xi,\, B - \xi\} = \{(1-\lambda)(A - B),\, \lambda(A - B)\} \quad a.e. \quad x \in \Sigma.$$

Let $A - B = q \otimes p$ with $p \in S^{n-1}$. Let $e \in S^{n-1}$ be a vector perpendicular to p. Then $D\phi(x)\, e = 0$ for almost every $x \in \Sigma$. This implies $\phi(x + te)$ is independent of t as long as $x + te \in \bar{\Sigma}$; hence, choosing t such that $x + te \in \partial\Sigma$, we have $\phi(x) = 0$ for all $x \in \Sigma$. Hence, $\xi \in \{A, B\}$, as claimed. □

Theorem 5.14. *Every finite-valued quasiconvex function is locally Lipschitz continuous.*

Proof. Let $F\colon \mathbb{M}^{N\times n} \to \mathbb{R}$ be quasiconvex. Then $f(t) = F(\xi + tq \otimes p)$ is a finite convex function of $t \in \mathbb{R}$; hence f is locally Lipschitz continuous. From this the theorem follows. □

Lemma 5.15. *If F is a polynomial of degree two (quadratic polynomial) then F is quasiconvex if and only if F is rank-one convex.*

Proof. Assume F is a rank-one convex quadratic polynomial. We show F is quasiconvex. Since subtraction of an affine function from a function does not change the quasiconvexity or rank-one convexity, we thus assume F is a homogeneous quadratic polynomial given by

$$F(\xi) = A_{ij}^{kl} \xi_i^k \xi_j^l \quad \text{(summation notation is used here and below)}$$

with A_{ij}^{kl} are constants. Note that the rank-one convexity is equivalent to the weak **Legendre-Hadamard** condition:

$$\sum_{i,j=1}^{n} \sum_{k,l=1}^{N} A_{ij}^{kl} q^k q^l \, p_i p_j \geq 0.$$

Using this condition and the Fourier transform as before (see proof of Lemma 3.12), we can show that

$$\int_{\mathbb{R}^n} A_{ij}^{kl} D_i\phi^k(x) D_j\phi^l(x)\, dx \geq 0$$

for all $\phi \in C_0^\infty(\mathbb{R}^n; \mathbb{R}^N)$. Note that

$$F(\xi + \eta) = F(\xi) + A_{ij}^{kl}\xi_i^k \eta_j^l + A_{ij}^{kl}\xi_j^l \eta_i^k + A_{ij}^{kl}\eta_i^k \eta_j^l.$$

Hence

$$\int_\Sigma F(\xi + D\phi(x))\, dx = F(\xi) + \int_{\mathbb{R}^n} A_{ij}^{kl} D_i\phi^k(x) D_j\phi^l(x)\, dx \geq F(\xi),$$

holds for all $\phi \in C_0^\infty(\Sigma; \mathbb{R}^N)$. This proves the quasiconvexity of F. □

Using the definition of **null-Lagrangians** given later, we have the following.

Lemma 5.16. *A rank-one convex third degree polynomial must be a null-Lagrangian and thus quasiconvex.*

Proof. Let F be a rank-one convex *third degree* polynomial. Then the polynomial $f(t) = F(\xi + tq \otimes p)$ is convex and of degree ≤ 3 in t, and hence the degree of $f(t)$ cannot be 3. Note that the coefficient of t^2 term in f is half of

$$F_{\xi_i^k \xi_j^l}(\xi)\, p_i p_j \, q^k q^l \geq 0, \tag{5.15}$$

which holds for all ξ, p, q. Since $F_{\xi_i^k \xi_j^l}(\xi)$ is linear in ξ, condition (5.15) implies

$$F_{\xi_i^k \xi_j^l}(\xi)\, p_i p_j \, q^k q^l \equiv 0.$$

Therefore $f(t) = F(\xi + tq \otimes p)$ is affine in t and hence F is *rank-one affine.* Consequently, the result follows from the fact that a rank-one affine function must be a null-Lagrangian. □

5.3.4. Šverák's Example. The following example of V. Šverák [**22**] settles a long-standing open problem raised by C. B. Morrey [**15**].

Theorem 5.17. **(Šverák)** *If $n \geq 2$, $N \geq 3$ then there exists a rank-one convex function $F\colon \mathbb{M}^{N\times n} \to \mathbb{R}$ which is not quasiconvex.*

Proof. We only prove the theorem for $n = 2$, $N = 3$. Consider the Σ-periodic function $u\colon \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$u(x) = \frac{1}{2\pi}(\sin 2\pi x_1,\, \sin 2\pi x_2,\, \sin 2\pi(x_1 + x_2)).$$

Then

$$Du(x) = \begin{pmatrix} \cos 2\pi x_1 & 0 \\ 0 & \cos 2\pi x_2 \\ \cos 2\pi(x_1 + x_2) & \cos 2\pi(x_1 + x_2) \end{pmatrix} \in L,$$

where L is the 3-dimensional linear subspace of $\mathbb{M} = \mathbb{M}^{3\times 2}$ defined by

$$L = \left\{ [r,s,t] \equiv \begin{pmatrix} r & 0 \\ 0 & s \\ t & t \end{pmatrix} \;\Big|\; r,\, s,\, t \in \mathbb{R} \right\}.$$

Note that a matrix $\xi = [r,s,t] \in L$ is of rank ≤ 1 if and only if at most one of $\{r,s,t\}$ is nonzero. Define $g\colon L \to \mathbb{R}$ by $g([r,s,t]) = -rst$. Using formula $2\cos\alpha \cdot \cos\beta = \cos(\alpha+\beta) + \cos(\alpha-\beta)$, we easily obtain

$$g(Du(x)) = -\frac{1}{4} - \frac{1}{4}\big(\sin 4\pi(x_1 + x_2) + \cos 4\pi x_1 + \cos 4\pi x_2\big)$$

and by a direct computation it follows that

$$\int_\Sigma g(Du(x))\, dx = \int_0^1 \int_0^1 g(Du(x))\, dx_1 dx_2 = -\frac{1}{4}. \tag{5.16}$$

We now extend g to the whole $\mathbb{M}$. Let $P\colon \mathbb{M} \to L$ be the orthogonal projection onto L and, for $k > 0$, $\epsilon > 0$, consider *fourth degree* polynomials:

$$F_{\epsilon,k}(\xi) = g(P\xi) + \epsilon\,(|\xi|^2 + |\xi|^4) + k\,|\xi - P\xi|^2. \tag{5.17}$$

Lemma 5.18. *For each $\epsilon > 0$ there exists a $k = k_\epsilon > 0$ such that F_{ϵ,k_ϵ} is rank-one convex.*

Proof. We use contradiction method. Suppose there exists an $\epsilon_0 > 0$ such that $F_{\epsilon_0,k}$ is not rank-one for all $k > 0$. Hence there exist sequences $\xi_k \in \mathbb{M}$, $p^k \in \mathbb{R}^2$, $q_k \in \mathbb{R}^3$ with $|p^k| = |q_k| = 1$ such that

$$\frac{\partial^2 F_{\epsilon_0,k}(\xi_k)}{\partial \xi_i^\alpha \partial \xi_j^\beta}\, q_k^\alpha q_k^\beta\, p_i^k p_j^k \equiv D^2 F_{\epsilon_0,k}(\xi_k)\,[q_k \otimes p^k,\, q_k \otimes p^k] \leq 0, \quad \forall\ k = 1, 2, \cdots. \tag{5.18}$$

Given $\xi, \eta \in \mathbb{M}$, let $f(t) = F_{\epsilon,k}(\xi + t\eta)$; then computing $f''(0)$ yields

$$\begin{aligned} f''(0) &= D^2 F_{\epsilon,k}(\xi)\,[\eta,\eta] \\ &= D^2 g(P\xi)\,[P\eta, P\eta] + 2\epsilon\,|\eta|^2 + \epsilon\,(4|\xi|^2|\eta|^2 + 8(\xi\cdot\eta)^2) + 2k\,|\eta - P\eta|^2. \end{aligned}$$

The term $D^2 g(P\xi)$ is linear in ξ; the third term is quadratic and positive definite in ξ if $\eta \neq 0$ (this is the reason the $|\xi|^4$-term is needed for $F_{\epsilon,k}$). Using this formula with $\xi = \xi_k$ and $\eta = q_k \otimes p^k$ and from (5.18), we deduce $\{\xi_k\}$ is bounded as $k \to \infty$. Assume, via subsequence,

$$\xi_k \to \bar\xi,\ q_k \to \bar q,\ p^k \to \bar p \quad \text{as } k \to \infty.$$

Note that $D^2F_{\epsilon,j}(\xi)\,[\eta,\,\eta] \le D^2F_{\epsilon,k}(\xi)\,[\eta,\,\eta]$ for all $k \ge j$. Hence we deduce

$$D^2g(P\bar{\xi})\,[P(\bar{q}\otimes\bar{p}),\,P(\bar{q}\otimes\bar{p})] + 2\epsilon_0 + 2j\,|P(\bar{q}\otimes\bar{p}) - \bar{q}\otimes\bar{p}\,| \le 0 \tag{5.19}$$

for all $j = 1, 2, \cdots$. Hence $P(\bar{q}\otimes\bar{p}) = \bar{q}\otimes\bar{p}$; that is, $\bar{q}\otimes\bar{p} \in L$. This implies $\bar{q}\otimes\bar{p} = [a, b, c]$, where at most one of a, b, c is nonzero. Therefore, function

$$t \mapsto g(P(\bar{\xi} + t\bar{q}\otimes\bar{p})) = g(P\bar{\xi} + t\bar{q}\otimes\bar{p})$$

is affine in t, and hence the first term in (5.19) vanishes. This yields the desired contradiction $\epsilon_0 \le 0$. The lemma is proved. □

We now complete the proof of Šverák's theorem. Let u be the periodic function above. We choose $\epsilon > 0$ small enough such that

$$\epsilon \int_\Sigma (|Du(x)|^2 + |Du(x)|^4)\,dx < \frac{1}{4}.$$

Let $F_\epsilon(\xi) = F_{\epsilon,k_\epsilon}(\xi)$ be a rank-one function determined by the previous lemma. Since $Du(x) \in L$, it follows by (5.16) that

$$\int_\Sigma F_\epsilon(Du(x))\,dx = \int_\Sigma g(Du) + \epsilon \int_\Sigma (|Du|^2 + |Du|^4) < 0 = F_\epsilon(0).$$

This shows that F_ϵ is not quasiconvex by Theorem 5.9 above. The theorem is now proved. □

5.4. Polyconvex Functions and Null-Lagrangians

Unlike the convexity and rank-one convexity, quasiconvexity is a global property since the inequality (5.7) is required to hold for all test functions. It is thus generally impossible to verify whether a given function $F(\xi)$ is quasiconvex.

We have already seen that every convex function is quasiconvex. However, there is a class of functions which are quasiconvex but not necessarily convex. This class, mainly due to C. B. Morrey, has been called the **polyconvex** functions in J. M. Ball [**4**]. In order to introduce the polyconvex functions, we need some notation.

5.4.1. Determinant and Adjugate Matrix. Let ξ be a $n\times n$ square matrix. We denote by $\det\xi$ and $\operatorname{adj}\xi$ the **determinant** and **adjugate matrix** of ξ, respectively, which satisfy the following relation:

$$\xi\,(\operatorname{adj}\xi) = (\operatorname{adj}\xi)\,\xi = (\det\xi)\,I,$$

where I is the $n\times n$ identity matrix. From this relation, we have

$$\frac{\partial \det\xi}{\partial\xi} = (\operatorname{adj}\xi)^T, \tag{5.20}$$

where η^T is the transpose matrix of η.

Lemma 5.19. *For all $u \in C^2(\bar{\Omega};\mathbb{R}^n)$ it follows that $\operatorname{div}(\operatorname{adj} Du(x))^T = 0$; that is,*

$$\sum_{i=1}^n D_i[(\operatorname{adj} Du(x))^i_k] = 0 \quad (k = 1, 2, \cdots, n).$$

Proof. Note that by the identity above, we have $(\det \xi)I = \xi^T(\operatorname{adj}\xi)^T$. For $\xi = Du(x)$ this implies

$$(\det Du(x))\,\delta_{ij} = \sum_{k=1}^{n} D_i(u^k(x))\,(\operatorname{adj} Du(x))_k^j, \quad i,\, j = 1, 2, \cdots, n.$$

Differentiating this identity with respect to x_j and summing over $j = 1, 2, \cdots, n$, we have

$$\sum_{j,k,m=1}^{n} \delta_{ij}(\operatorname{adj} Du)_k^m D_j(D_m u^k) = \sum_{k,j=1}^{n} (D_j D_i u^k)\,(\operatorname{adj} Du)_k^j + (D_i u^k)\, D_j[(\operatorname{adj} Du)_k^j]$$

for $i = 1, 2, \cdots, n$. This identity simplifies to read

$$\sum_{k=1}^{n}(D_i u^k)\left(\sum_{j=1}^{n} D_j[(\operatorname{adj} Du)_k^j]\right) = 0 \quad (i = 1, 2, \cdots, n).$$

In short, this can be written as

$$Du(x)^T[\operatorname{div}(\operatorname{adj} Du(x))^T] = 0, \quad x \in \Omega. \tag{5.21}$$

Now, if $\det Du(x_0) \neq 0$ then by (5.21), $\operatorname{div}(\operatorname{adj} Du(x_0))^T = 0$. If instead $\det Du(x_0) = 0$, we choose a sequence $\epsilon_\nu \to 0$ such that $\det(Du(x_0) + \epsilon_\nu I) \neq 0$ for all ν. Use (5.21) with $\tilde{u} = u + \epsilon_\nu x$ we have $\det D\tilde{u}(x_0) \neq 0$ and hence

$$0 = \operatorname{div}(\operatorname{adj} D\tilde{u}(x_0))^T = \operatorname{div}(\operatorname{adj}(Du(x_0) + \epsilon_\nu I))^T = 0$$

for all $\epsilon_\nu \to 0$. Hence $\operatorname{div}(\operatorname{adj}(Du(x_0)))^T = 0$. $\square$

5.4.2. Subdeterminants. Let $\sigma = \min\{n, N\}$. Given an integer $k \in [1, \sigma]$, for any two ordered k-tuples of integers

$$1 \le i_1 < i_2 < \cdots < i_k \le N, \quad 1 \le \alpha_1 < \alpha_2 < \cdots < \alpha_k \le n,$$

let $J^{i_1 i_2 \cdots i_k}_{\alpha_1 \alpha_2 \cdots \alpha_k}(\xi)$ be the *determinant* of the $k \times k$ matrix whose (q, p) position element is $\xi^{i_q}_{\alpha_p}$ for all $1 \le p,\, q \le k$. Note that, by the usual notation,

$$J^{i_1 i_2 \cdots i_k}_{\alpha_1 \alpha_2 \cdots \alpha_k}(Du(x)) = \frac{\partial(u^{i_1}, u^{i_2}, \cdots, u^{i_k})}{\partial(x_{\alpha_1}, x_{\alpha_2}, \cdots, x_{\alpha_k})} = \det\left(\frac{\partial u^{i_q}}{\partial x_{\alpha_p}}\right).$$

Let $\mathcal{J}(\xi)$ be a fixed arrangement of $J^{i_1 i_2 \cdots i_k}_{\alpha_1 \alpha_2 \cdots \alpha_k}(\xi)$ for all $k \in [1, \sigma]$ and k-tuples $\{i_q\}$, $\{\alpha_p\}$. In this way, $\mathcal{J}$ defines a function from $\mathbb{M}^{N\times n}$ to $\mathbb{R}^L$, where

$$L = L(n, N) = \sum_{k=1}^{\sigma}\begin{pmatrix} N \\ k \end{pmatrix}\begin{pmatrix} n \\ k \end{pmatrix}.$$

Theorem 5.20. *Let $\mathcal{J}(\xi)$ be defined as above, and let Σ be the unit cube in $\subset \mathbb{R}^n$. Then it follows that*

$$\int_\Sigma \mathcal{J}\big(\xi + D\phi(x)\big)\, dx = \mathcal{J}(\xi)$$

for all $\phi \in C_0^\infty(\Sigma; \mathbb{R}^N)$ and $\xi \in \mathbb{M}^{N\times n}$.

Proof. Since each $\mathcal{J}(\xi)$ is given by a $k \times k$-determinant, without loss of generality, we only prove this identity for $J_k(\xi) = J^{12\cdots k}_{12\cdots k}(\xi)$, where $1 \le k \le \sigma = \min\{n,\, N\}$. For simplicity, set $u(x) = \xi x + \phi(x)$. Let

$$x' = (x_1, \cdots, x_k), \quad x'' = (x_{k+1}, \cdots, x_n) \quad \text{if } k + 1 \le n.$$

Let Σ', Σ'' be the unit cubes in x', x'' variables, respectively. Fix $x'' \in \Sigma''$, for $t \geq 0$, consider maps $V_t, U_t \colon \Sigma' \to \mathbb{R}^k$ such that

$$V_t^i(x') = tx_i + (\xi x)^i, \quad U_t^i(x') = tx_i + u^i(x', x'').$$

We can choose $t > 0$ sufficiently large so that V_t, U_t are both *diffeomorphisms* on Σ', and therefore

$$\int_{\Sigma'} \det(DU_t(x'))\, dx' = \int_{U_t(\Sigma')} dy' = \int_{V_t(\Sigma')} dy' = \int_{\Sigma'} \det(DV_t(x'))\, dx'.$$

Since both sides are polynomials of t of degree k, it follows that this equality holds for all t. When $t = 0$ this implies

$$\int_{\Sigma'} J_k(\xi + D\phi(x', x''))\, dx' = \int_{\Sigma'} J_k(\xi)\, dx'. \tag{5.22}$$

Integrating (5.22) over $x'' \in \Sigma''$ we deduce

$$\int_{\Sigma} J_k(\xi + D\phi(x))\, dx = J_k(\xi), \tag{5.23}$$

completing the proof. □

5.4.3. Polyconvex Functions. A (finite-valued) function $F \colon \mathbb{M}^{N\times n} \to \mathbb{R}$ is called a **polyconvex** function if there exists a *convex* function $G \colon \mathbb{R}^L \to \mathbb{R}$ such that $F(\xi) = G(\mathcal{J}(\xi))$ for all $\xi \in \mathbb{M}^{N\times n}$; that is, $F = G \circ \mathcal{J}$ on $\mathbb{M}^{N\times n}$.

Remark. For a polyconvex function we may have different convex functions in its representation. For example, let $n = N = 2$ and $F(\xi) = |\xi|^2 - \det \xi$. In this case, let $\mathcal{J}(\xi) = (\xi, \det \xi) \in \mathbb{R}^5$. Then we have

$$F(\xi) = G_1(\mathcal{J}(\xi)), \quad F(\xi) = G_2(\mathcal{J}(\xi)),$$

where

$$G_1(\xi, t) = |\xi|^2 - t, \quad G_2(\xi, t) = (\xi_1^1 - \xi_2^2)^2 + (\xi_1^2 + \xi_2^1)^2 + t$$

are both convex functions of (ξ, t).

Theorem 5.21. *A polyconvex function is quasiconvex.*

Proof. Let $F \colon \mathbb{M}^{N\times n} \to \mathbb{R}$ be a polyconvex function. Then there exists a convex function $G \colon \mathbb{R}^L \to \mathbb{R}$ such that $F(\xi) = G(\mathcal{J}(\xi))$ for all ξ. Given $\xi \in \mathbb{M}^{N\times n}$ and $\phi \in C_0^\infty(\Sigma; \mathbb{R}^N)$, let $h(x) = \mathcal{J}(\xi + D\phi(x))$. Then Jensen's inequality implies

$$G\left(\int_{\Sigma} h(x)\, dx\right) \leq \int_{\Sigma} G(h(x))\, dx.$$

By the theorem above, the left-hand side is $G(\mathcal{J}(\xi)) = F(\xi)$ and therefore

$$F(\xi) \leq \int_{\Sigma} G(h(x))\, dx = \int_{\Sigma} G(\mathcal{J}(\xi + D\phi(x)))\, dx = \int_{\Sigma} F(\xi + D\phi(x))\, dx,$$

proving that F is quasiconvex. □

5.4.4. Null-Lagrangians. A smooth function $F\colon \Omega\times\mathbb{R}^N\times\mathbb{M}^{N\times n}\to\mathbb{R}$ is called a **null-Lagrangian** if the system of Euler-Lagrange equations (4.4) for the energy

$$I(u)=\int_\Omega F(x,u,Du)\,dx$$

is satisfied by all smooth functions $u:\Omega\to\mathbb{R}^N$; that is, for all $u\in C^2(\bar\Omega;\mathbb{R}^N)$, the equation

$$\sum_{i=1}^n \frac{\partial}{\partial x_i} F_{\xi_i^k}(x,u(x),Du(x)) = F_{s^k}(x,u(x),Du(x)),\quad \forall\ x\in\bar\Omega, \tag{5.24}$$

holds for all $k=1,2,\cdots,N$.

Example 5.22. Let $f\colon\mathbb{R}^n\to\mathbb{R}$ be C^1, and let $F(s,\xi)=f(s)\det\xi$ for $s\in\mathbb{R}^n$, $\xi\in\mathbb{M}^{n\times n}$. Then, by (5.20),

$$F_{s^k}(s,\xi)=f_{s^k}(s)\ \det\xi,\quad F_{\xi_i^k}(s,\xi)=f(s)\,(\operatorname{adj}\xi)^i_k.$$

Hence for any $u\in C^2(\bar\Omega;\mathbb{R}^n)$, by Lemma 5.19 above, we have

$$\begin{aligned}\sum_{i=1}^n \frac{\partial}{\partial x_i}F_{\xi_i^k}(u(x),Du(x)) &= \sum_{i=1}^n\frac{\partial}{\partial x_i}[f(u(x))\,(\operatorname{adj}Du(x))^i_k]\\ &= \sum_{i,j=1}^n f_{s^j}(u(x))\,u^j_{x_i}\,(\operatorname{adj}Du(x))^i_k\\ &= \sum_{j=1}^n f_{s^j}(u(x))\,(\det Du(x))\delta_{jk}\\ &= f_{s^k}(u(x))\,(\det Du(x))\end{aligned}$$

for all $k=1,2,\cdots,n$. Hence $F(s,\xi)=f(s)\det\xi$ is a null-Lagrangian.

We now prove the following boundary dependence property for null-Lagrangians.

Theorem 5.23. *Let $I(u)=\int_\Omega F(x,u,Du)\,dx$ and F be a null-Lagrangian. Assume $u,\,v\in C^2(\bar\Omega;\mathbb{R}^N)$ satisfy $u(x)=v(x)$ for all $x\in\partial\Omega$. Then $I(u)=I(v)$. Moreover, if F satisfies $|F(x,s,\xi)|\le c(|s|^p+|\xi|^p)+C(x)$ with $c>0$ and $C\in L^1(\Omega)$, then $I(u)=I(v)$ for all $u,\,v\in W^{1,p}(\Omega;\mathbb{R}^N)$ with $u-v\in W^{1,p}_0(\Omega;\mathbb{R}^N)$.*

Proof. Define $h(t)=I(U(t,\cdot))$ for $t\in[0,1]$, where $U(t,x)=tu(x)+(1-t)v(x)$. Then

$$h'(t)=\int_\Omega\Big(F_{\xi_i^k}(x,U(t,x),DU(t,x))D_i(u^k-v^k)+F_{s^k}(x,U(t,x),DU(t,x))(u^k-v^k)\Big)\,dx.$$

Since $u-v$, $U(t,\cdot)\in C^2(\bar\Omega;\mathbb{R}^N)$ and $u-v=0$ on $\partial\Omega$, using the divergence theorem and (5.24) with $u=U(t,\cdot)$ we have $h'(t)=0$ for all $t\in(0,1)$ and hence $h(0)=h(1)$ and the first part of the result follows. The second part follows by the continuity of I on $W^{1,p}(\Omega;\mathbb{R}^N)$ under the given assumption. □

We now study the null-Lagrangians depending only on variable in $\mathbb{M}^{N\times n}$.

Theorem 5.24. *Let $F\colon\mathbb{M}^{N\times n}\to\mathbb{R}$ be continuous and let*

$$I_\Omega(u)=\int_\Omega F(Du(x))\,dx,$$

where Ω is any smooth bounded domain in $\mathbb{R}^n$. Then the following conditions are equivalent:

(1) *F is of C^2 and is a* null-Lagrangian;

(2) *$F(\xi)=\fint_\Sigma F(\xi+D\varphi(x))\,dx$ for all $\xi\in\mathbb{M}^{N\times n}$ and $\varphi\in C_0^\infty(\Sigma;\mathbb{R}^N)$;*

(3) *$I_\Omega(u)=I_\Omega(u+\phi)$ for all $u\in C^1(\bar\Omega;\mathbb{R}^N)$ and $\phi\in C_0^\infty(\Omega;\mathbb{R}^N)$;*

(4) *I_Ω is* continuous *with respect to the Lipschitz convergence on $W^{1,\infty}(\Omega;\mathbb{R}^N)$;*

(5) *there exists a linear function $\mathcal{L}\colon\mathbb{R}^L\to\mathbb{R}$ such that $F(\xi)=\mathcal{L}(\mathcal{J}(\xi))$ for all $\xi\in\mathbb{M}^{N\times n}$.*

Proof. We will not show the equivalence of (5) to other conditions, but it is important to know that null-Lagrangians can only be the linear combination of subdeterminants. We prove other equivalent conditions.

It is easy to see that (3) implies (2). Note that (2) is equivalent to that both F and $-F$ are quasiconvex; the latter is equivalent to that both I_Ω and $-I_\Omega$ are lower semicontinuous with respect to the Lipschitz convergence on $W^{1,\infty}(\Omega;\mathbb{R}^N)$. Therefore, (4) is equivalent to (2). It remains to show (2) implies (1) implies (3). Let us first show that (1) implies (3). To this end, given $u\in C^2(\bar\Sigma;\mathbb{R}^N)$ and $\phi\in C_0^\infty(\Sigma;\mathbb{R}^N)$, let

$$f(t)=\int_\Sigma F(Du(x)+tD\phi(x))\,dx.$$

Then, by (1),

$$f'(t)=\int_\Sigma F_{\xi^i_\alpha}(Du(x)+tD\phi(x))\,D_\alpha\phi^i(x)\,dx=0.$$

Hence f is a constant function; hence $f(0)=f(1)$, which proves (3). The proof of that (2) implies (1) will follow from several lemmas proved below. □

Lemma 5.25. *If F is of C^2, then (2) implies (1).*

Proof. Note that (2) implies

$$\int_\Omega F(\xi+D\phi(x))\,dx=F(\xi)\,|\Omega|\quad\forall\ \phi\in C_0^\infty(\Omega;\mathbb{R}^N).$$

Since F is of C^2, this implies that, for any $\phi,\,\psi\in C_0^\infty(\Omega;\mathbb{R}^N)$, the function

$$f(t)=\int_\Omega F(t\,D\phi(x)+D\psi(x))\,dx$$

is constant and of C^2. Therefore, $f'(0)=0$, which gives

$$\int_\Omega F_{\xi^k_i}(D\psi(x))\,D_i\phi^k(x)\,dx=0.$$

Now given $u\in C^2(\bar\Omega;\mathbb{R}^N)$, we can select a sequence $\{\psi_\nu\}$ in $C_0^\infty(\Omega;\mathbb{R}^N)$ such that $\psi_\nu\to u$ in $C^2(\operatorname{supp}\phi;\mathbb{R}^N)$. Using the identity above with $\psi=\psi_\nu$ and letting $\nu\to\infty$ yield

$$\int_\Omega F_{\xi^k_i}(Du(x))\,D_i\phi^k(x)\,dx=0,$$

which, by the divergence theorem, shows the Euler-Lagrange equation for I is satisfied by $u\in C^2(\bar\Omega;\mathbb{R}^N)$, and hence F is a null-lagrangian. □

We say F is **rank-one affine** if $F(\xi+tq\otimes p)$ is affine in t for all $\xi\in\mathbb{M}^{N\times n}$, $q\in\mathbb{R}^N$, $p\in\mathbb{R}^n$.

Lemma 5.26. *If F satisfies* (2) *then F is rank-one affine.*

Proof. Since (2) is equivalent to that F and $-F$ are rank-one convex, which is equivalent to that F is rank-one affine. $\square$

We need some notation. Let $\mu^i_\alpha = e^i \otimes e_\alpha$, where $\{e^i\}$ and $\{e_\alpha\}$ are the standard bases of $\mathbb{R}^N$ and $\mathbb{R}^n$, respectively. For each $1 \le k \le \sigma = \min\{n, N\}$ and $1 \le i_1, \cdots, i_k \le N$, $1 \le \alpha_1, \cdots, \alpha_k \le n$, we define, inductively,

$$F^{i_1}_{\alpha_1}(\xi) = F(\xi + \mu^{i_1}_{\alpha_1}) - F(\xi),$$

$$F^{i_1\cdots i_k}_{\alpha_1\cdots\alpha_k}(\xi) = F^{i_1\cdots i_{k-1}}_{\alpha_1\cdots\alpha_{k-1}}(\xi + \mu^{i_k}_{\alpha_k}) - F^{i_1\cdots i_{k-1}}_{\alpha_1\cdots\alpha_{k-1}}(\xi).$$

Note that if F is a polynomial it follows

$$F^{i_1\cdots i_k}_{\alpha_1\cdots\alpha_k}(\xi) = \partial^k F(\xi)/\partial\xi^{i_1}_{\alpha_1}\cdots\partial\xi^{i_k}_{\alpha_k}.$$

Indeed, we have the same permutation invariance property.

Lemma 5.27. *Let $\{1', \cdots, k'\}$ be any permutation of $\{1, \cdots, k\}$. Then*

$$F^{i_1\cdots i_k}_{\alpha_1\cdots\alpha_k}(\xi) = F^{i_{1'}\cdots i_{k'}}_{\alpha_{1'}\cdots\alpha_{k'}}(\xi).$$

Proof. We use induction on k. Assume this is true for all $k \le s-1$. Let $\{1', 2', \cdots, s'\}$ be a permutation of $\{1, 2, \cdots, s\}$. We need to show

$$F^{i_1\cdots i_s}_{\alpha_1\cdots\alpha_s}(\xi) = F^{i_{1'}\cdots i_{s'}}_{\alpha_{1'}\cdots\alpha_{s'}}(\xi). \tag{5.25}$$

By definition

$$F^{i_1\cdots i_s}_{\alpha_1\cdots\alpha_s}(\xi) = F^{i_1\cdots i_{s-1}}_{\alpha_1\cdots\alpha_{s-1}}(\xi + \mu^{i_s}_{\alpha_s}) - F^{i_1\cdots i_{s-1}}_{\alpha_1\cdots\alpha_{s-1}}(\xi).$$

By induction assumption, (5.25) holds if $s' = s$. We thus assume $s' < s$. In this case, by induction assumption,

$$F^{i_1\cdots i_s}_{\alpha_1\cdots\alpha_s}(\xi) = F^{i_1\cdots\hat{i}_{s'}\cdots i_{s'}}_{\alpha_1\cdots\hat{\alpha}_{s'}\cdots\alpha_{s'}}(\xi + \mu^{i_s}_{\alpha_s}) - F^{i_1\cdots\hat{i}_{s'}\cdots i_{s'}}_{\alpha_1\cdots\hat{\alpha}_{s'}\cdots\alpha_{s'}}(\xi)$$

(where the $\hat{m}$ means omitting m)

$$= F^{i_1\cdots\hat{i}_{s'}\cdots}_{\alpha_1\cdots\hat{\alpha}_{s'}\cdots}(\xi + \mu^{i_s}_{\alpha_s} + \mu^{i_{s'}}_{\alpha_{s'}}) - F^{i_1\cdots\hat{i}_{s'}\cdots}_{\alpha_1\cdots\hat{\alpha}_{s'}\cdots}(\xi + \mu^{i_s}_{\alpha_s})$$

$$-F^{i_1\cdots\hat{i}_{s'}\cdots}_{\alpha_1\cdots\hat{\alpha}_{s'}\cdots}(\xi + \mu^{i_{s'}}_{\alpha_{s'}}) + F^{i_1\cdots\hat{i}_{s'}\cdots}_{\alpha_1\cdots\hat{\alpha}_{s'}\cdots}(\xi)$$

$$= F^{i_1\cdots\hat{i}_{s'}\cdots i_s}_{\alpha_1\cdots\hat{\alpha}_{s'}\cdots\alpha_s}(\xi + \mu^{i_{s'}}_{\alpha_{s'}}) - F^{i_1\cdots\hat{i}_{s'}\cdots i_s}_{\alpha_1\cdots\hat{\alpha}_{s'}\cdots\alpha_s}(\xi),$$

which, by induction assumption, equals

$$F^{i_{1'}\cdots i_{(s-1)'}}_{\alpha_{1'}\cdots\alpha_{(s-1)'}}(\xi + \mu^{i_{s'}}_{\alpha_{s'}}) - F^{i_{1'}\cdots i_{(s-1)'}}_{\alpha_{1'}\cdots\alpha_{(s-1)'}}(\xi) = F^{i_{1'}\cdots i_{s'}}_{\alpha_{1'}\cdots\alpha_{s'}}(\xi).$$

This proves the induction procedure and hence the lemma. $\square$

Lemma 5.28. *If F is rank-one affine, then all $F^{i_1\cdots i_k}_{\alpha_1\cdots\alpha_k}$ are also rank-one affine. Moreover,*

$$F^{i_1\cdots i_k}_{\alpha_1\cdots\alpha_k}(\xi + t\,\mu^j_\beta) = F^{i_1\cdots i_k}_{\alpha_1\cdots\alpha_k}(\xi) + t\,F^{i_1\cdots i_k j}_{\alpha_1\cdots\alpha_k\beta}(\xi). \tag{5.26}$$

Proof. Use induction again on k. $\square$

Lemma 5.29. *If F satisfies (2) then $F(\xi)$ is a polynomial in ξ.*

Proof. If F satisfies (2) then F is rank-one affine. Write

$$\xi = \sum_{i=1}^{N} \sum_{\alpha=1}^{n} \xi_\alpha^i \, \mu_\alpha^i .$$

Then a successive use of the previous lemma shows that $F(\xi)$ is a polynomial of degree at most nN in ξ with coefficients determined by $F_{\alpha_1\cdots\alpha_k}^{i_1\cdots i_k}(0)$. $\square$

EXAMPLE 5.30. The function $F(\xi) = (\operatorname{tr}\xi)^2 - \operatorname{tr}\xi^2$ is a null-Lagrangian on $\mathbb{M}^{n\times n}$. To see this, note that

$$F(\xi) = (\sum_{i=1}^{n} \xi_i^i)(\sum_{k=1}^{n} \xi_k^k) - \sum_{i,k=1}^{n} \xi_i^k \xi_k^i = \sum_{i,k=1}^{n} P_{ik}(\xi),$$

where $P_{ik}(\xi) = \xi_i^i \xi_k^k - \xi_k^i \xi_i^k$. For each pair (i,k), $i \neq k$, $P_{ik}(\xi)$ is a 2×2 subdeterminant of ξ. Hence $F(\xi)$ is a null-Lagrangian.

5.4.5. Compensated Compactness of Null-Lagrangians. We prove a compensated compactness property of the null-Lagrangians; for general results on compensated compactness, see [**7, 24**].

Theorem 5.31. *Let $J_k(Du)$ be any $k\times k$ subdeterminant. Let $\{u_j\}$ be any sequence weakly convergent to $\bar u$ in $W^{1,k}(\Omega;\mathbb{R}^N)$ as $j\to\infty$. Then $J_k(Du_j)\to J_k(D\bar u)$ in the sense of distribution in Ω; that is,*

$$\lim_{j\to\infty} \int_\Omega J_k(Du_j(x))\,\phi(x)\,dx = \int_\Omega J_k(D\bar u(x))\,\phi(x)\,dx$$

for all $\phi \in C_0^\infty(\Omega)$.

Proof. We prove this by induction on k. Obviously, the theorem is true when $k=1$. Assume it holds for J_s with $s \le k-1$. We need to show it also holds for $s=k$. Without loss of generality, we may assume

$$J_k(Du(x)) = \frac{\partial(u^1, u^2, \cdots, u^k)}{\partial(x_1, x_2, \cdots, x_k)}.$$

For any smooth function u, we observe that $J_k(Du)$ is actually a divergence:

$$J_k(Du(x)) = \sum_{\nu=1}^{k} \frac{\partial}{\partial x_\nu}\left(u^1 \frac{(-1)^{\nu+1}\partial(u^2,\cdots,u^k)}{\partial(x_1,\cdots,\hat{x_\nu},\cdots,x_k)}\right), \tag{5.27}$$

where $\hat{x_\nu}$, again, means deleting x_ν. Let

$$J_{k-1}^{(\nu)}(Du(x)) = \frac{(-1)^{\nu+1}\partial(u^2,\cdots,u^k)}{\partial(x_1,\cdots,\hat{x_\nu},\cdots,x_k)}.$$

Then (5.27) implies

$$\int_\Omega J_k(Du(x))\,\phi(x)\,dx = \sum_{\nu=1}^{k} \int_\Omega u^1(x)\,J_{k-1}^{(\nu)}(Du(x))\,D_\nu\phi(x)\,dx. \tag{5.28}$$

By density argument, this identity still holds if $u \in W^{1,k}(\Omega;\mathbb{R}^N)$. Suppose $u_j \rightharpoonup \bar u$ in $W^{1,k}(\Omega;\mathbb{R}^N)$. By the Sobolev embedding theorem, $u_j \to \bar u$ in $L^k(\Omega;\mathbb{R}^N)$. Moreover, by the induction assumption, for each ν, $J_{k-1}^{(\nu)}(Du_j) \to J_{k-1}^{(\nu)}(D\bar u)$ in the sense of distribution. Note that since sequence $\{J_{k-1}^{(\nu)}(Du_j)\}$ is also bounded in $L^{\frac{k}{k-1}}(\Omega)$ it also weakly converges in

$L^{\frac{k}{k-1}}(\Omega)$; hence, the weak limit must be equal to the distribution limit $J^{(\nu)}_{k-1}(D\bar u)$. Therefore, we have proved that $u^1_j D_\nu\phi \to \bar u^1 D_\nu \phi$ strongly in $L^k(\Omega)$ and $J^{(\nu)}_{k-1}(Du_j) \rightharpoonup J^{(\nu)}_{k-1}(D\bar u)$ weakly in $L^{\frac{k}{k-1}}(\Omega)$; hence

$$\lim_{j\to\infty}\int_\Omega u^1_j(x) D_\nu\phi(x) J^{(\nu)}_{k-1}(Du_j(x))\,dx = \int_\Omega \bar u^1(x) D_\nu\phi(x) J^{(\nu)}_{k-1}(D\bar u(x))\,dx.$$

From this and (5.28) we conclude

$$\lim_{j\to\infty}\int_\Omega J_k(Du_j(x))\,\phi(x)\,dx = \sum_{\nu=1}^n \int_\Omega \bar u^1(x) D_\nu\phi(x) J^{(\nu)}_{k-1}(D\bar u(x))\,dx = \int_\Omega J_k(D\bar u(x))\,\phi(x)\,dx,$$

as desired. The proof is complete. □

Corollary 5.32. *Let $J_k(Du)$ be any $k\times k$ subdeterminant and $p>k$ be a number. Let $\{u_j\}$ be any sequence weakly convergent to $\bar u$ in $W^{1,p}(\Omega;\mathbb{R}^N)$ as $j\to\infty$. Then $J_k(Du_j)\rightharpoonup J_k(D\bar u)$ weakly in $L^{\frac{p}{k}}(\Omega)$.*

Proof. Note that $f_j = J_k(Du_j)$ is bounded in $L^{\frac{p}{k}}(\Omega)$ and hence, via a subsequence, $f_j \rightharpoonup \bar f$ for some $\bar f$ weakly in $L^{\frac{p}{k}}(\Omega)$, which also implies $f_j \to \bar f$ in distribution. Hence $\bar f \equiv J_k(Du)$ and the whole sequence $f_j \rightharpoonup J_k(Du)$ in $L^{\frac{p}{k}}(\Omega)$. □

Example 5.33. In general, the weak convergence $u_j \rightharpoonup \bar u$ in $W^{1,k}(\Omega;\mathbb{R}^N)$ does not imply the weak convergence $J_k(Du_j)\rightharpoonup J_k(D\bar u)$ in $L^1(\Omega)$. For example, let $\Omega = B$ be the unit open ball in $\mathbb{R}^n$. Consider, for $j=1,2,\cdots$, the *radial* mappings

$$u_j(x) = \frac{U_j(|x|)}{|x|}x,\quad U_j(r) = \begin{cases} jr & \text{if } 0\le r\le 1/j,\\ 2-jr & \text{if } 1/j\le r\le 2/j,\\ 0 & \text{if } 2/j\le r<1.\end{cases}$$

Computation shows that $u_j \rightharpoonup 0$ in $W^{1,n}(B;\mathbb{R}^n)$ as $j\to\infty$. But

$$\det Du_j(x) = (U_j(r)/r)^{n-1} U_j'(r)$$

for a.e. $x\in B$, where $r=|x|$, and hence

$$\int_{|x|<2/j} |\det Du_j(x)|\,dx = C$$

is a constant independent of j. This shows that the sequence $\{\det Du_j\}$ is not equi-integrable in B, and therefore it does not converges weakly in $L^1(B)$. Therefore, although $u_j \rightharpoonup \bar u$ in $W^{1,k}(\Omega;\mathbb{R}^N)$, $J_k(Du_j)\to J_k(D\bar u)$ in the sense of distribution, and $\{J_k(Du_j)\}$ is bounded in $L^1(\Omega)$, but it is **not true** that $J_k(Du_j)\rightharpoonup J_k(D\bar u)$ weakly in $L^1(\Omega)$.

However, if we assume $\det Du_j(x)\ge 0$ a.e. in Ω and $u_j\to\bar u$ weakly in $W^{1,n}(\Omega;\mathbb{R}^n)$ then $\det Du_j \rightharpoonup \det D\bar u$ weakly in $L^1_{loc}(\Omega)$; this follows from a well-known theorem of S. Müller that $\det Du \ln(1+\det Du)\in L^1_{loc}(\Omega)$ if $u\in W^{1,n}(\Omega;\mathbb{R}^n)$ satisfies $\det Du(x)\ge 0$ a.e. $x\in\Omega$. We will not prove this result here, but prove the following similar result which is useful below.

Theorem 5.34. *Let $u_j\in\mathcal{D}_\varphi$, the Dirichlet class of $\varphi\in W^{1,n}(\Omega;\mathbb{R}^n)$. Assume $u_j\rightharpoonup\bar u$ in $W^{1,n}(\Omega;\mathbb{R}^n)$ and $\det Du_j(x)\ge 0$ a.e. in Ω. Then $\det Du_j\rightharpoonup\det D\bar u$ weakly in $L^1(\Omega)$.*

Proof. Note that, since $\det Du_j$ is bounded in $L^1(\Omega)$, the weak convergence is equivalent to

$$\lim_{j\to\infty}\int_E \det Du_j(x)\,dx = \int_E \det D\bar{u}(x)\,dx \tag{5.29}$$

for all measurable sets $E \subseteq \Omega$. By Theorem 5.23, we have

$$\int_\Omega \det Du_j(x)\,dx = \int_\Omega \det D\bar{u}(x)\,dx = \int_\Omega \det D\varphi(x)\,dx. \tag{5.30}$$

Hence (5.29) holds when $E = \Omega$. Let $E \subset \Omega$. Let

$$F(x,\xi) = \chi_E(x)\,\max\{0, \det\xi\},$$

where χ_E is the **characteristic function** of E. Then $F(x,\xi)$ is polyconvex and hence quasiconvex in ξ. Note that

$$0 \le F(x,\xi) \le C\,|\xi|^n.$$

Hence by Theorem 5.6 the functional

$$I(u) = \int_\Omega F(x,Du)dx$$

is *w.l.s.c.* on $W^{1,n}(\Omega;\mathbb{R}^n)$. Using the fact

$$F(x,Du_j(x)) = \chi_E(x)\det Du_j(x), \quad F(x,D\bar{u}(x)) \ge \chi_E(x)\det D\bar{u}(x) \quad \forall\ a.e.\ x\in\Omega,$$

this w.l.s.c. implies

$$\int_E \det D\bar{u}(x)\,dx \le \liminf_{j\to\infty}\int_E \det Du_j(x)\,dx. \tag{5.31}$$

This is also valid for $E^c = \Omega\setminus E$; that is,

$$\int_{E^c} \det D\bar{u}(x)\,dx \le \liminf_{j\to\infty}\int_{E^c} \det Du_j(x)\,dx. \tag{5.32}$$

Note that

$$\int_{E^c}\det Du_j\,dx = \int_\Omega \det Du_j\,dx - \int_E \det Du_j\,dx.$$

Hence (5.30) and (5.32)imply

$$\int_E \det D\bar{u}\,dx \ge \limsup_{j\to\infty}\int_E \det Du_j\,dx,$$

which, combined with (5.31), implies (5.29). $\square$

5.5. Existence in Nonlinear Elasticity

5.5.1. Hyperelastic Materials. In nonlinear elasticity, a **hyperelastic material** refers to a material that possesses a **stored energy density function** $F(x,u,Du)$ to characterize all the continuum mechanical properties under the deformation u of the material. In particular, the **total stored energy** is defined by

$$I(u) = \int_\Omega F(x,u(x),Du(x))\,dx,$$

where $\Omega\subset\mathbb{R}^3$ is a bounded smooth domain occupied by the hyperelastic material in the *reference configuration*, and $u\colon\Omega\to\mathbb{R}^3$ is the deformation ($u(x)$ representing displacement of material point $x\in\Omega$ in the deformed configuration).

The stored energy density function $F(x,s,\xi)$ should satisfy several assumptions due to continuum mechanics principles and material properties; some of these assumptions exclude the possibility of F being convex in ξ and almost all practical problems require that $F(x,s,\xi)$ satisfy the following:

$$F(x,s,\xi)=\infty \quad \text{if } \det\xi\le 0, \qquad \lim_{\det\xi\to 0^+} F(x,s,\xi)=\infty. \tag{5.33}$$

The first is due to the **orientation-preserving** assumption and the second is due to the fact that infinite energy is needed to press a solid material into zero-volume.

Therefore, for nonlinear elasticity problems, we often need to consider **extended-valued** energy functionals; theorems proved above for finite-valued functionals may not be applicable to these problems.

5.5.2. Polyconvex Energy Density Functions. We assume the energy density function $F(x,s,\xi)$ satisfies the following assumptions:

(1) $F(x,s,\xi)=+\infty$ if and only if $\det\xi\le 0$, and $\lim_{\det\xi\to 0^+} F(x,s,\xi)=\infty$ uniformly on (x,s);

(2) $F(x,s,\xi)$ is continuous in $(x,s)\in\bar\Omega\times\mathbb{R}^3$ and *polyconvex* on $\xi\in\mathbb{M}^{3\times 3}$, $\det\xi>0$ in the sense that there exists a continuous function $W(x,s,J)$ on $\bar\Omega\times\mathbb{R}^3\times\mathbb{R}^{19}_+$, convex in $J\in\mathbb{R}^{19}_+=\mathbb{M}^{3\times3}\times\mathbb{M}^{3\times3}\times\mathbb{R}^+$, such that
$$F(x,s,\xi)=W(x,s,\xi,\operatorname{adj}\xi,\det\xi)$$
for all $x\in\bar\Omega$, $s\in\mathbb{R}^3$, $\xi\in\mathbb{M}^{3\times3}$ with $\det\xi>0$; furthermore, $W(x,s,J)$ is smooth in $J\in\mathbb{R}^{19}_+$ and $W_J(x,s,J)$ is continuous on $\bar\Omega\times\mathbb{R}^3\times\mathbb{R}^{19}_+$;

(3) $F(x,s,\xi)\ge c_0|\xi|^3$ for all $\xi\in\mathbb{M}^{3\times3}$, where $c_0>0$ is a constant.

5.5.3. Existence Theorem. Let $I(u)$ be the energy functional defined by
$$I(u)=\int_\Omega F(x,u(x),Du(x))\,dx,$$
where F satisfies the previous assumptions. Let $\varphi\in W^{1,3}(\Omega;\mathbb{R}^3)$ be a given function such that $I(\varphi)<\infty$. Then we have the following existence result.

Theorem 5.35. (Existence in Nonlinear Elasticity) *There exists at least one minimizer $\bar u\in\mathcal{D}_\varphi$ with $\det D\bar u(x)>0$ for almost every $x\in\Omega$ such that*
$$I(\bar u)=\min_{u\in\mathcal{D}_\varphi} I(u).$$

Proof. Let $\{u_\nu\}$ be a minimizing sequence in $\mathcal{D}_\varphi$. Since $\lim I(u_\nu)=\inf_{\mathcal{D}_\varphi} I\le I(\varphi)<\infty$, from Assumption (1), it follows that
$$\det Du_\nu(x)>0 \quad a.e.\ \Omega.$$
Now Assumption (3) implies that $\{|Du_\nu|\}$ is bounded in $L^3(\Omega)$ and hence, having the fixed Dirichlet boundary condition, $\{u_\nu\}$ is bounded in $W^{1,3}(\Omega;\mathbb{R}^3)$. Therefore, via a subsequence, we assume $u_\nu\rightharpoonup\bar u$ in $W^{1,3}(\Omega;\mathbb{R}^3)$ as $\nu\to\infty$. Then, by compact imbedding, Corollary 5.32 and Theorem 5.34,

$$u_\nu\to\bar u \ \text{ in } L^3(\Omega), \quad \operatorname{adj} Du_\nu\rightharpoonup\operatorname{adj} D\bar u \ \text{ in } L^{3/2}(\Omega), \quad \det Du_\nu\rightharpoonup\det D\bar u \ \text{ in } L^1(\Omega). \tag{5.34}$$

It is easy to see that $\det D\bar u(x)\ge 0$ for almost every $x\in\Omega$. We show $\det D\bar u(x)>0$ a.e. in Ω. We follow an idea of Pedregal. Let
$$h(\tau)=\inf\{F(x,s,\xi)\,|\,(x,s,\xi)\in\bar\Omega\times\mathbb{R}^3\times\mathbb{M}^{3\times3},\ \det\xi=\tau\}.$$

Then, by Assumption (1), $\lim_{\tau \to 0^+} h(\tau) = \infty$. Note that

$$\sup_\nu \int_\Omega h(\det Du_\nu(x))\, dx \le \sup_\nu \int_\Omega F(x, u_\nu(x), Du_\nu(x))\, dx < \infty.$$

Let $E = \{x \in \Omega \mid \det D\bar{u}(x) = 0\}$. We show $|E| = 0$. Suppose $|E| > 0$. By weak convergence,

$$\lim_{\nu\to\infty} \int_E \det Du_\nu(x) dx = \int_E \det D\bar{u}(x) dx = 0.$$

So $\det Du_\nu \to 0$ in $L^1(E)$. Hence, $\det Du_{\nu'}(x) \to 0$ a.e. $x \in E$ along a subsequence $\nu' \to \infty$. Therefore, $h(\det Du_{\nu'}(x)) \to \infty$ a.e. $x \in E$. However, by Fatou's lemma,

$$\int_E \liminf_{\nu'\to\infty} h(\det Du_{\nu'}(x))\, dx \le \liminf_{\nu'\to\infty} \int_E h(\det Du_{\nu'}(x)) dx < \infty.$$

This implies $\liminf_{\nu'\to\infty} h(\det Du_{\nu'}(x)) < \infty$ a.e. $x \in E$, a contradiction. Hence $|E| = 0$; so, $\det D\bar{u}(x) > 0$ a.e. in Ω.

By Assumption (2), $F(x, s, \xi) = W(x, s, J(\xi))$ for all x, s, ξ with $\det \xi > 0$, where $J(\xi) = (\xi, \operatorname{adj} \xi, \det \xi)$. Hence, at every point x where $\det Du_\nu(x) > 0$ and $\det D\bar{u}(x) > 0$, it follows that

$$\begin{aligned} & F(x, u_\nu, Du_\nu) - F(x, \bar{u}, D\bar{u}) \\ = \;& [W(x, u_\nu, J(D\bar{u})) - W(x, \bar{u}, J(D\bar{u}))] + [W(x, u_\nu, J(Du_\nu)) - W(x, u_\nu, J(D\bar{u}))] \\ \ge \;& [W(x, u_\nu, J(D\bar{u})) - W(x, \bar{u}, J(D\bar{u}))] + W_J(x, u_\nu, J(D\bar{u})) \cdot (J(Du_\nu) - J(D\bar{u})). \end{aligned}$$

We can then follow the similar steps as in the proof of the **Tonelli theorem** (Theorem 5.1) to conclude that

$$I(\bar{u}) \le \liminf_{\nu\to\infty} I(u_\nu).$$

For example, by (5.34), (via a subsequence) we can also assume $u_\nu(x) \to \bar{u}(x)$ for almost every $x \in \Omega$. Now, for any given $\delta > 0$, we choose a compact set $K \subset \Omega \setminus E$, where $E = \{x \in \Omega \mid \det D\bar{u}(x) = 0\}$, such that

(i) $u_\nu \to \bar{u}$ uniformly on K and $|\Omega \setminus K| < \delta$ (by **Egorov's theorem**);

(ii) $\bar{u}$, $D\bar{u}$ are continuous on K (by **Lusin's theorem**);

(iii) $\det D\bar{u}(x) \ge \delta_0 > 0$ on K for some constant $\delta_0 > 0$ (as $\det D\bar{u} > 0$ on K).

Since $F \ge 0$, it follows that

$$\begin{aligned} I(u_\nu) \;\ge\; & \int_K F(x, u_\nu, Du_\nu)\, dx \ge \int_K F(x, \bar{u}, D\bar{u}) dx \\ + \; & \int_K [W(x, u_\nu, J(D\bar{u})) - W(x, \bar{u}, J(D\bar{u}))] dx \\ + \; & \int_K W_J(x, u_\nu, J(D\bar{u})) \cdot (J(Du_\nu) - J(D\bar{u})) dx. \end{aligned}$$

Since $W(x, s, J)$ and $W_J(x, s, J)$ are both uniformly continuous on compact subsets of $\Omega \times \mathbb{R}^3 \times \mathbb{R}^{19}_+$ and $u_\nu(x) \to \bar{u}(x)$ uniformly on K, it follows that $W(x, u_\nu, J(D\bar{u})) \to W(x, \bar{u}, J(D\bar{u}))$ and $W_J(x, u_\nu, J(D\bar{u})) \to W_J(x, \bar{u}, J(D\bar{u}))$ both uniformly on K. Therefore, by (5.34), we conclude that

$$\infty > \liminf_{\nu\to\infty} I(u_\nu) \ge \int_K F(x, \bar{u}, D\bar{u})\, dx.$$

This, holding for a sequence of compact subsets K of Ω approaching Ω, yields that $I(\bar{u}) = \int_\Omega F(x, \bar{u}, D\bar{u})dx < \infty$. Hence, for each $\epsilon > 0$, **Lebesgue's absolute continuity theorem** determines a number $\delta > 0$ such that

$$\int_S F(x, \bar{u}, D\bar{u})dx \geq \int_\Omega F(x, \bar{u}, D\bar{u})dx - \epsilon, \quad \forall\, S \subset \Omega,\ |\Omega \setminus S| < \delta.$$

Using this $\delta > 0$ with $S = K$, where K is determined according to (i, (ii) and (iii) above, we conclude, by setting $\epsilon \to 0$, that

$$\liminf_{\nu\to\infty} I(u_\nu) \geq I(\bar{u}).$$

Hence $\bar{u}$ is a minimizer. This completes the proof. □

Remark. For general existence results with Assumption (3) weakened to $F(x, s, \xi) \geq c_0(|\xi|^p + |\operatorname{adj}\xi|^q)$ for certain $p < 3$ and $q > 1$, see Müller, Tang & Yan [**18**].

5.6. Relaxation Principle and Existence for Nonconvex Problems

There are many application problems that do not satisfy the convexity conditions. In such cases, minimizers may not be found as we did before using the direct method, but the variational methods may still help to study the problem.

5.6.1. Non-quasiconvex Problems. In general, we consider the multiple integral functional

$$I(u) = \int_\Omega F(x, u(x), Du(x))\, dx,$$

where Ω is a bounded domain in $\mathbb{R}^n$, $u\colon \Omega \to \mathbb{R}^N$, and $F(x, s, \xi)$ is a given function *not quasiconvex* in ξ. We give two examples of one-dimensional problem.

Example 5.36. Consider the classical example of Bolza: to minimize the functional

$$I(u) = \int_0^1 [((u')^2 - 1)^2 + u^2]dx$$

among all Lipschitz functions u satisfying boundary conditions $u(0) = u(1) = 0$. Let

$$u_k(x) = \frac{1}{2k} - \left| x - \frac{[kx]}{k} - \frac{1}{2k} \right|,$$

where $[a]$ stands for the integer part of $a \in \mathbb{R}$. Then $u_k \to 0$ in the Lipschitz convergence, and $I(u_k) \to 0$ as $k \to \infty$. Hence the infimum of I over all such u is 0. Note that $I(0) = 1$. Hence I is *not* lower semicontinuous in the Lipschitz convergence. Moreover, I does not have minimizers since $I(u) = 0$ would imply $|u'| = 1$ a.e and hence $u \equiv 0$.

Example 5.37. Let $f(\xi) = (\xi^2 - 1)^2$ for $\xi \in \mathbb{R}$ and consider the functional

$$I(u) = \int_0^1 f(u'(x))\, dx = \int_0^1 ((u')^2 - 1)^2 dx$$

over all Lipschitz functions u satisfying $u(0) = 0$, $u(1) = a$. Then this nonconvex problem has minimizers for all values of a; moreover, the minimizer is unique if $|a| \geq 1$ and there exist infinitely many minimizers if $|a| < 1$.

Proof. Let first of all $|a| < 1$. Then it can be easily seen that there are infinitely many Lipschitz functions u such that $u'(x) \in \{-1, 1\}$ with $u(0) = 0$ and $u(1) = a$; any of such functions will be a minimizer of I with minimum value zero. Now let $|a| \geq 1$. Note that the strict inequality

$$f(\xi) - [f(a) + f'(a)(\xi - a)] > 0$$

holds for all $\xi \neq a$ if $|a| > 1$, or holds for all $\xi \notin \{-1, 1\}$ if $|a| = 1$. Therefore, for all Lipschitz functions u with $u(0) = 0, u(1) = a$ it follows that

$$I(u) = \int_0^1 f(u'(x))\,dx \geq f(a) + \int_0^1 f'(a)(u'(x) - a)\,dx = f(a) = I(\bar{u}),$$

where $\bar{u}(x) = ax$; hence the minimum value of I is $f(a)$ and, certainly, $\bar{u}(x) = ax$ is a minimizer. We show that $\bar{u}$ is the unique minimizer. Suppose v is any minimizer of I over all Lipschitz functions u with $u(0) = 0, u(1) = a$; that is,

$$I(v) = \int_0^1 f(v'(x))\,dx = f(a)$$

with $v(0) = 0$ and $v(1) = a$. If $a = 1$ then since $f(1) = 0$ it must follow that $v'(x) \in \{-1, 1\}$ a.e. $x \in (0, 1)$ and $\int_0^1 (1 - v'(x))\,dx = 0$ and hence $v'(x) = 1$ on $(0, 1)$. This implies $v(x) = x$. Similarly if $a = -1$ then $v(x) = -x$. Now assume $|a| > 1$. Since $f(\xi) - f(a) - f'(a)(\xi - a) > 0$ for all $\xi \neq a$, from

$$\int_0^1 [f(v'(x)) - f(a) - f'(a)(v'(x) - a)]\,dx = 0,$$

it follows that $v'(x) = a$ for a.e. $x \in (0, 1)$; henc $v(x) = ax$. Therefore, in all cases, $v = \bar{u}$. This completes the proof. $\square$

From this example, we see that the function

$$g(\xi) = \inf_{w \in W_0^{1,\infty}(0,1)} \int_0^1 f(\xi + w'(x))\,dx = f^c(\xi) = \begin{cases} f(\xi) & \text{if } |\xi| \geq 1 \\ 0 & \text{if } |\xi| < 1 \end{cases}$$

is the **convexification** of function f; that is, the largest convex function less than or equal to f.

5.6.2. Quasiconvexification. Given a function $F \colon \mathbb{M}^{N\times n} \to \mathbb{R}$, we define the largest quasiconvex function less than or equal to F to be the **quasiconvexication** or **quasiconvex envelope** of F. We denote the quasiconvexification of F by F^{qc}.

Theorem 5.38. *Let $F \geq 0$ be continuous. Then*

$$F^{qc}(\xi) = \inf_{\phi \in W_0^{1,\infty}(\Sigma;\mathbb{R}^N)} \int_\Sigma F(\xi + D\phi(x))\,dx, \quad \xi \in \mathbb{M}^{N\times n}. \tag{5.35}$$

Proof. For any quasiconvex function $G \leq F$,

$$G(\xi) \leq \int_\Sigma G(\xi + D\phi(x))\,dx \leq \int_\Sigma F(\xi + D\phi(x))\,dx.$$

Hence $G(\xi) \leq F^{qc}(\xi)$ by the definition of F^{qc}. It remains to prove that F^{qc} is itself quasiconvex. We first observe that

$$F^{qc}(\xi) = \inf_{\phi \in W_0^{1,\infty}(\Omega;\mathbb{R}^N)} ⨍_\Omega F(\xi + D\phi(x))\,dx \tag{5.36}$$

for any open set $\Omega \subset \mathbb{R}^n$ with $|\partial\Omega| = 0$. This can be proved by using the Vitali covering argument as before. We now claim that, for all piecewise affine Lipschitz continuous functions $\phi \in W_0^{1,\infty}(\Sigma; \mathbb{R}^N)$, it follows that

$$F^{qc}(\xi) \le \fint_{\Sigma} F^{qc}(\xi + D\phi(x))\, dx. \tag{5.37}$$

To see this, let $\Sigma = \cup_i \Omega_i \cup E$ be such that $|E| = 0$ and each Ω_i is an open set with $|\partial\Omega_i| = 0$ and such that $\phi \in W_0^{1,\infty}(\Sigma; \mathbb{R}^N)$ takes constant gradients $D\phi = M_i$ on each Ω_i. Let $\epsilon > 0$ be given. By the above remark on the definition of F^{qc}, there exists $\psi_i \in W_0^{1,\infty}(\Omega_i; \mathbb{R}^N)$ such that

$$F^{qc}(\xi + M_i) \ge \fint_{\Omega_i} F(\xi + M_i + D\psi_i(x))\, dx - \epsilon.$$

With each ψ_i being extended by zero to Σ, we set $\psi = \phi + \sum_i \psi_i$. Then $\psi \in W_0^{1,\infty}(\Sigma; \mathbb{R}^N)$ and we have

$$\begin{aligned}
\int_{\Sigma} F^{qc}(\xi + D\phi(x))\, dx &= \sum_i |\Omega_i| F^{qc}(\xi + M_i) \\
&\ge \sum_i \left[\int_{\Omega_i} F(\xi + M_i + D\psi_i(x))\, dx - \epsilon |\Omega_i| \right] \\
&= \int_{\Sigma} F(\xi + D\psi(x))\, dx - \epsilon \ge F^{qc}(\xi) - \epsilon,
\end{aligned}$$

and hence the ineqquality (5.37) follows. From this we conclude that function F^{qc} is rank-one convex (using "sawtooth" like piecewise affine functions as before) and is thus locally Lipschitz continuous. To show (5.37) holds for general test functions $\phi \in W_0^{1,\infty}(\Omega; \mathbb{R}^N)$, we use approximation of ϕ by piecewise affine functions. An approximation theorem (see, e.g., [**9**, Corollary 10.13]) asserts that there exists a sequence of piecewise affine functions $\phi_k \in W_0^{1,\infty}(\Omega; \mathbb{R}^N)$ such that

$$\|D\phi_k\|_{L^\infty} \le C, \quad \|D\phi_k - D\phi\|_{L^p(\Omega)} \to 0,$$

where $1 < p < \infty$. Using such an approximation and the continuity of F^{qc}, one can show that (5.37) holds for $\phi \in W_0^{1,\infty}(\Sigma; \mathbb{R}^N)$. Hence F^{qc} is quasiconvex. This completes the proof. □

5.6.3. Weak Lower Semicontinuous Envelope. Suppose $F(x, s, \xi)$ is a Caratheodory function and satisfies

$$0 \le F(x, s, \xi) \le C\,(|\xi|^p + |s|^p + 1) \quad \forall x \in \Omega,\ s \in \mathbf{R}^N,\ \xi \in \mathbb{M}^{N\times n}$$

for some $1 \le p < \infty$. We are interested in the *largest w.l.s.c.* functional $\tilde{I}(u)$ on $W^{1,p}(\Omega; \mathbb{R}^N)$ which is less than or equal to the functional

$$I(u) = \int_{\Omega} F(x, u(x), Du(x))\, dx.$$

This functional $\tilde{I}(u)$ is called the **weak lower semicontinuous envelope** or **relaxation** of I in the weak topology of $W^{1,p}(\Omega; \mathbb{R}^N)$. It turns out that under some mild conditions $\tilde{I}(u)$ is given by the integral functional of the quasiconvexification of F with respect to ξ.

Theorem 5.39. *Let $F^{qc}(x,s,\cdot)$ be the quasiconvexification of $F(x,s,\cdot)$ for given (x,s). In addition, assume F^{qc} is also a Caratheodory function. Then the envelope $\tilde{I}(u)$ of I in the weak topology of $W^{1,p}(\Omega;\mathbb{R}^N)$ is given by*

$$\tilde{I}(u) = \int_\Omega F^{qc}(x,u(x),Du(x))\,dx.$$

Proof. Let $\hat{I}(u)$ be the integral of $F^{qc}(x,u(x),Du(x))$ over Ω. Since F^{qc} is quasiconvex and satisfies the same growth condition as F, the functional $\hat{I}(u)$ is thus (sequential) w.l.s.c. on $W^{1,p}(\Omega;\mathbb{R}^N)$ by the theorem of Acerbi & Fusco. Therefore, $\hat{I} \le \tilde{I}$. To prove the other direction, we first assume there exists a Caratheodory function $g(x,s,\xi)$ such that

$$\tilde{I}(u) = \int_\Omega g(x,u(x),Du(x))\,dx \quad \forall\ u \in W^{1,p}(\Omega;\mathbb{R}^N). \tag{5.38}$$

Then $g(x,s,\cdot)$ must be quasiconvex and $g \le F$, and thus $g \le F^{qc}$; this proves $\tilde{I} \le \hat{I}$. However, the proof of integral representation (5.38) is beyond the scope of this lecture and is omitted; see e.g., Acerbi & Fusco [**1**], Buttazzo [**6**], or Dacorogna [**8**]. □

5.6.4. Relaxation Principle. We have the following theorem; the proof of the theorem is also omitted; see the references above.

Theorem 5.40. (Relaxation Principle) *Assume $F(x,s,\xi)$ and $F^{qc}(x,s,\xi)$ are both Caratheodory and that*

$$0 \le F(x,s,\xi) \le C\left(|\xi|^p + |s|^p + 1\right)$$

holds for some constants $C>0$, $p>1$. Then

$$\inf_{u\in\mathcal{D}_\varphi} \int_\Omega F(x,u,Du)\,dx = \inf_{u\in\mathcal{D}_\varphi} \int_\Omega F^{qc}(x,u,Du)\,dx$$

for any $\varphi \in W^{1,p}(\Omega;\mathbb{R}^N)$, where $\mathcal{D}_\varphi$ is the Dirichlet class of φ.

Remarks. (a) The passage from F to F^{qc} is called **relaxation**. The relaxation principle replaces a nonconvex problem that may not have any solution by a quasiconvex (relaxed) problem that will have solutions if in addition we assume the coercivity condition $F(x,s,\xi) \ge c|\xi|^p$ for some constant $c>0$. But, there are costs for this: we lose some useful information about the minimizing sequences; e.g., the finer and finer patterns of minimizing sequences.

(b) In phase transition problems for certain materials, the finer and finer patterns of minimizing sequences account for the *microstructures*, while the minimizers of relaxed problem may only capture the *macroscopic* (or effective) properties of the material. In certain studies, microstructures are usually characterized by **Young measures**; for more information, we refer to Müller [**17**] and the references therein.

(c) By the Relaxation Principle, if u_0 is a minimizer of $\int_\Omega F(x,u,Du)\,dx$ then it must satisfy the **differential inclusion**:

$$Du_0(x) \in K(x,u_0(x)) \quad a.e.\ x\in\Omega, \tag{5.39}$$

with set-valued function $K(x,s) = \{\xi \mid F(x,s,\xi) = F^{qc}(x,s,\xi)\}$. There are whole lots of researches on differential inclusions like (5.39), which could be a totally new topics course.

5.6.5. Existence of Minimizers for a Nonconvex Problem. Consider

$$I(u) = \int_\Omega f(\nabla u(x))\,dx$$

where Ω is a bounded domain in $\mathbb{R}^n$, $u\colon \Omega \to \mathbb{R}$ is a scalar function, and $f\colon \mathbb{R}^n \to \mathbb{R}$ is a given continuous function. We define the **subgradient** of f at a point ξ by

$$\partial f(\xi) = \{l \in \mathbb{R}^n \mid f(\eta) \geq f(\xi) + l \cdot (\eta - \xi), \quad \forall\, \eta \in \mathbb{R}^n\}.$$

Each $l \in \partial f(\xi)$ is called a **subdifferential** of f at ξ. Note that $\partial f(\xi)$ may be an empty set for some ξ.

For a set of points $\xi_1, \xi_2, \cdots, \xi_q$ in $\mathbb{R}^n$, we denote by $\operatorname{co}\{\xi_1, \cdots, \xi_q\}$ the **convex hull** defined by

$$\operatorname{conv}\{\xi_1, \cdots, \xi_q\} = \left\{ \sum_{i=1}^q \lambda_i \xi_i \,\middle|\, \forall\, \lambda_i \geq 0, \sum_{i=1}^q \lambda_i = 1 \right\}.$$

The following theorem indicates that even without convexity assumption the minimization may still have solutions. However, the existence can not follow from the direct method, but has to rely on different methods.

Theorem 5.41. *Given $\xi \in \mathbb{R}^n$, the minimization problem*

$$\inf_{u \in W^{1,1}(\Omega),\, u|_{\partial\Omega} = \xi x} \int_\Omega f(\nabla u)dx$$

has a minimizer if and only if either $\partial f(\xi) \neq \emptyset$, or there exist $\xi_1, \xi_2, \cdots, \xi_q$ such that

$$\xi \in \operatorname{int} \operatorname{conv}\{\xi_1, \cdots, \xi_q\} \quad \textit{and} \quad \cap_{i=1}^q \partial f(\xi_i) \neq \emptyset. \tag{5.40}$$

Proof. This form of the theorem is due to Sychev [**23**] and we follow his proof too.

1. **"Sufficiency."** First, if $\partial f(\xi) \neq \emptyset$, then one easily shows that $u_\xi \equiv \xi x$ is a minimizer. Indeed, let $\eta \in \partial f(\xi)$; then

$$f(\nabla u(x)) \geq f(\xi) + \eta \cdot (\nabla u(x)) - \xi) \quad \forall\, u \in W^{1,1}(\Omega).$$

Hence $I(u) \geq I(u_\xi)$ for all such u with $u|_{\partial\Omega} = u_\xi$.

Now suppose there exist $\xi_1, \xi_2, \cdots, \xi_q$ such that

$$\xi \in \operatorname{int} \operatorname{conv}\{\xi_1, \cdots, \xi_q\} \quad \text{and} \quad \cap_{i=1}^q \partial f(\xi_i) \neq \emptyset.$$

We assume that $\xi_1, \cdots, \xi_r$ are the **extreme points** of $\operatorname{conv}\{\xi_1, \cdots, \xi_q\}$; hence $\operatorname{conv}\{\xi_1, \cdots, \xi_r\} = \operatorname{conv}\{\xi_1, \cdots, \xi_q\}$. Since $\xi \in \operatorname{int} \operatorname{conv}\{\xi_1, \cdots, \xi_r\}$, we can show that the function

$$w(x) = \max_{1 \leq i \leq r} \{(\xi_i - \xi) \cdot x - 1\}$$

is Lipschitz, $\nabla w(x) \in \{\xi_i - \xi \colon i = 1, \cdots, r\}$ a.e., and $w|_{\partial P} = 0$, where

$$P = \{x \in \mathbb{R}^n \colon w(x) \leq 0\}$$

is a compact polyhedral set with nonempty interior containing 0. (Due to the special shape of its graph, function $-w(x)$ is usually called a **pyramid function** on P.) By Vitali covering argument, we can write $\Omega = \cup_{i=1}^\infty (y_i + s_i P) \cup N$, where $|N| = 0$, $y_i \in \Omega$, $s_i > 0$ and $\{y_i + s_i P\}$ are disjoint subsets of Ω. We now define

$$u_0(x) = \begin{cases} \xi x + s_i w(\frac{x - y_i}{s_i}) & \text{for } x \in y_i + s_i P,\ i = 1, 2, \cdots, \\ \xi x & \text{elsewhere on } \Omega. \end{cases}$$

Then $u_0 \in W^{1,\infty}(\Omega)$ and $u_0|_{\partial\Omega} = \xi x$. We show that u_0 is minimizer. Indeed, let $\eta \in \cap_{i=1}^r \partial f(\xi_i)$. Then, $f(\xi_j) \geq L_i(\xi_j)$ for all $i, j = 1, 2, \cdots, r$, where $L_i(v) = f(\xi_i) + \eta \cdot (v - \xi_i)$; this implies that all L_i are equal and $f(v) \geq L_i(v)$ for all $v \in \mathbb{R}^n$. Therefore, for all $u \in W^{1,1}(\Omega)$ with $u|_{\partial\Omega} = \xi x$, using $\nabla u_0(x) \in \{\xi_1, \xi_2, \cdots, \xi_r\}$, it follows that

$$\begin{aligned} I(u) - I(u_0) &= \int_\Omega (f(\nabla u) - f(\nabla u_0) - \eta \cdot (\nabla u - \nabla u_0))\, dx \\ &= \int_\Omega (f(\nabla u) - L_1(\nabla u))\, dx \geq 0. \end{aligned}$$

2. **"Necessity."** Let u_0 be a minimizer. Let f^{**} be the **convexification** of f. Then a similar argument of the relaxation principle shows that $I(u_0) \leq f^{**}(\xi)|\Omega|$. Let $\eta \in \partial f^{**}(\xi)$. Then

$$I(u_0) - f^{**}(\xi)|\Omega| = \int_\Omega (f(\nabla u_0(x)) - f^{**}(\xi) - \eta \cdot (\nabla u_0(x) - \xi))\, dx \geq 0.$$

Hence $I(u_0) = f^{**}(\xi)|\Omega|$. This implies $\nabla u_0(x) \in P_\eta$ for a.e. $x \in \Omega$, where

$$P_\eta = \{v \in \mathbb{R}^n \colon f(v) - f^{**}(\xi) - \eta \cdot (v - \xi) = 0\}.$$

If $f^{**}(\xi) = f(\xi)$, then $\partial f(\xi) \neq \emptyset$. We now assume $f^{**}(\xi) < f(\xi)$. We claim $\xi \in \operatorname{int} \operatorname{conv} P_\eta$. Suppose this is not the case. Then, by the **Hahn-Banach theorem** in convex analysis, there exists $a \in \mathbb{R}^n$ such that $\xi \cdot a \geq v \cdot a$ for all $v \in \operatorname{conv} P_\eta$. Hence $\xi \cdot a \geq \nabla u_0(x) \cdot a$ for a.e. $x \in \Omega$. Since $\int_\Omega \xi \cdot a\, dx = \int_\Omega \nabla u_0(x) \cdot a\, dx$, it follows that $\xi \cdot a = \nabla u_0(x) \cdot a$ for a.e. $x \in \Omega$; hence the directional derivative

$$\frac{\partial (u_0 - \xi x)}{\partial a} = (\nabla u_0 - \xi) \cdot a = 0 \quad a.e.\ \Omega.$$

So, for a.e. $x \in \Omega$, the function $h(t) = u_0(x + ta) - \xi \cdot (x + ta)$ must be constant on each open interval of t where the function is defined; the endpoints t_0 of any such interval must satisfy $x + t_0 a \in \partial\Omega$. Since $u_0(x) = \xi x$ on $x \in \partial\Omega$, this implies $u_0(x) \equiv \xi x$; hence $\xi \in P_\eta$ and $f^{**}(\xi) = f(\xi)$, a contradiction. So $\xi \in \operatorname{int} \operatorname{conv} P_\eta$ and thus there exist $\xi_1, \xi_2, \cdots, \xi_q \in P_\eta$ such that $\xi \in \operatorname{int} \operatorname{conv}\{\xi_1, \xi_2, \cdots, \xi_q\}$. Since $\xi_i \in P_\eta$, we have $f(\xi_i) = f^{**}(\xi) + \eta \cdot (\xi_i - \xi)$; so, for all $v \in \mathbb{R}^n$, as $\eta \in \partial f^{**}(\xi)$, it follows that

$$\begin{gathered} f(v) \geq f^{**}(v) \geq f^{**}(\xi) + \eta \cdot (v - \xi) \\ = f^{**}(\xi) + \eta \cdot (\xi_i - \xi) + \eta \cdot (v - \xi_i) = f(\xi_i) + \eta \cdot (v - \xi_i), \end{gathered}$$

which shows that $\eta \in \partial f(\xi_i)$ for all $i = 1, 2, \cdots, q$. So $\cap_{i=1}^q \partial f(\xi_i) \neq \emptyset$. This completes the proof. □

Remark. Condition (5.40) is essentially equivalent to the condition that the face of epi f^{**} whose relative interior contains $(\xi, f^{**}(\xi))$ has dimension n, which is the original assumption of Cellina. Standard notation and theorems in convex analysis can be found in Rockafellar's book [**20**].

Bibliography

[1] E. Acerbi and N. Fusco, *Semicontinuity problems in the calculus of variations,* Arch. Rational Mech. Anal., **86** (1984), 125–145.

[2] R. A. Adams and J. Fournier, “Sobolev Spaces,” Second Edition. Academic Press, New York, 2003.

[3] A. Bruckner, J. Bruckner and B. Thomson, “Real Analysis,” Prentice-Hall, 1997.

[4] J. M. Ball, *Convexity conditions and existence theorems in nonlinear elasticity*, Arch. Rational Mech. Anal., **63** (1977), 337–403.

[5] V. Barbu and Th. Precupanu, “Convexity and Optimization in Banach Spaces,” Reidel, Boston, 1986.

[6] G. Buttazzo, “Semicontinuity, Relaxation and Integral Representation in the Calculus of Variations,” Pitman Res. Notes Math. Ser. **207**, Harlow, 1989.

[7] R. Coifman, P.-L. Lions, Y. Meyer and S. Semmes, *Compensated compactness and Hardy spaces,* J. Math. Pures Appl., **72(9)** (1993), 247–286.

[8] B. Dacorogna, “Direct Methods in the Calculus of Variations,” second edition, Springer, New York, 2008.

[9] B. Dacorogna and P. Marcellini, “Implicit Partial Differential Equations,” Birkhäuser, Boston, 1999.

[10] D. Dunninger, Lecture Notes on Applied Analysis, (unpublished), Michigan State University

[11] L. C. Evans, “Partial Differential Equations,” AMS-GSM, Vol. **19**, 1998.

[12] M. Giaquinta, “Multiple Integrals in the Calculus of Variations and Nonlinear Elliptic Systems,” Princeton University Press, 1983.

[13] M. Giaquinta, “Introduction to Regularity Theory for Nonlinear Elliptic Systems,” Birkhäuser, 1993.

[14] D. Gilbarg and N. S. Trudinger, “Elliptic Partial Differential Equations of Second Order,” 2nd ed., Springer-Verlag, Berlin, 1984.

[15] C. B. Morrey, Jr., *Quasiconvexity and the lower semicontinuity of multiple integrals,* Pacific J. Math., **2** (1952), 25–53.

[16] C. B. Morrey, Jr., “Multiple Integrals in the Calculus of Variations,” Springer, New York, 1966.

[17] S. Müller, *Variational models for microstructure and phase transitions,* in “Calculus of Variations and Geometric Evolution Problems,” (Cetraro, 1996), pp. 85–210, Lecture Notes in Math. **1713**, Springer, 1999.

[18] S. Müller, Q. Tang and B. Yan, *On a new class of elastic deformations not allowing for cavitation,* Ann. Inst. Henri Poincaré, Analyse non linéaire, **11**(2) (1994), 217–243.

[19] P. Pedregal, “Parametrized Measure and Variational Principles.” Birkhauser, Basel, 1997.

[20] R. T. Rockafellar, “Convex Analysis.” Princeton University Press, Princeton, 1972.

[21] E. Stein, “Harmonic Analysis.” Princeton University Press, Princeton, 1993.

[22] V. Šverák, *Rank-one convexity does not imply quasiconvexity,* Proc. Roy. Soc. Edinburgh Sect. A, **120** (1992), 185–189.

[23] M. Sychev, *Characterization of homogeneous scalar variational problems solvable for all boundary data,* Proceedings of the Royal Society of Edinburgh, **130A** (2000), 611–631.

[24] L. Tartar, *The compensated compactness method applied to systems of conservation laws,* in "Systems of Nonlinear Partial Differential Equations", NATO ASI Series, Vol. **103**, D. Reidel, Dordrecht, 1983.

[25] R. Temam, "Navier-Stokes Equations: Theory and Numerical Analysis." Revised Edition. Elsevier Science Publishers, Amsterdam, 1984.

[26] K. Yosida, "Functional Analysis," 6th edition, Springer, New York, 1980.

[27] L. C. Young, "Lectures on the Calculus of Variations and Optimal Control Theory," Saunders, 1969 (reprint by Chelsea, 1980).

[28] E. Zeidler, "Nonlinear Functional Analysis and its Applications," Vol.1-4, Springer.

Index

H^2-Regularity Theorem, 76
i-th difference quotient, **53**
p-Laplace equations, **78**
Šverák Theorem, 111

Acerbi & Fusco Theorem, 105
adjugate matrix, **112**
admissible class, **2**
Ambrosetti-Rabinowitz Theorem, 90
Ascoli Theorem, 14

Banach space
  reflexive, 12
  separable, 8
Banach-Steinhaus Theorem, 11
Bessel's inequality, **15**
biharmonic, **61**
bilinear form, **2**, **58**
  bounded, 58
  strongly positive, 58
Bounded Inverse Theorem, 11

Cartesian product, **9**
Cauchy sequence, **8**
Cauchy's inequality with ϵ, **59**
Cauchy-Schwarz inequality, **10**
Chain Rule Theorem, 36
characteristic function, **120**
coercive, **58**
coercivity, **59**
compensated compactness, **4**
continuation method, **17**
Contraction Mapping Theorem, 17
convex, **23**, **73**
convex hull, **127**
convexification, **124**, **128**
convexity, **2**
critical point, **23**, **25**
critical points, **71**

Deformation Lemma Lemma, 90
determinant, **112**
Difference Quotients and Weak Derivatives Theorem, 53
differential operator
  uniformly elliptic, 55
direct method, **1**, **78**
Dirichlet class, **75**
dual space, **12**
  second dual space, 12

Ehrling's inequality, **27**, **80**
eigenfunctions, **16**, **69**
eigenspace, **69**
eigenvalue, **69**
Eigenvalue Theorem Theorem, 69
eigenvalues, **16**
eigenvectors, **16**
ellipticity condition, **2**
ellipticity conditions, **57**
energy method, **2**
equicontinuous, **14**
Estimates for $W^{1,p}$, $n < p \leq \infty$ Theorem, 49
Euler-Lagrange equations, **72**
Existence in Nonlinear Elasticity Theorem, 121
Existence of Minimizers Theorem, 78
Existence of minimizers Theorem, 106
extended-valued, **121**
extremal point, **25**

first variation, **71, 72**
fixed point, **17**
Fourier series, **15**
Fourier transform, **54**, **62**
Frechet differentiable, **19**
Fredholm Alternative Theorem, 15
function
  cut-off function, 32
  mollified function, 32
  p-mean continuous, 33
  support, 8
  trace, 41, 42
functional
  convex, 24
  weakly coercive, 22, 77
  weakly continuous, 22
  weakly lower semicontinuous, 22

Gårding's estimates, **3**, **59**
Gagliardo-Nirenberg-Sobolev Inequality Lemma, 45
Gateaux differentiable, **19**
Gateaux-differentiable, **74**
generalized Hölder's inequality, **9**
generalized Hölder inequality, **46**
Global H^2-Regularity Theorem, 66
Global Higher Regularity Theorem, 67
Green's 1st identity, **43**
Green's 2nd identity, **43**

Hölder's inequality, **9**
Hölder's inequality, **9**
Hahn-Banach Theorem, 12
harmonic map equation, **82**
harmonic maps, **83**
Higher Trace Theorem, 43
Hilbert space, **10**
hyperelastic material, **120**

Implicit Function Theorem, 21
inner product, **9**
Interior H^2-Regularity Theorem, 65
Interior Higher Regularity Theorem, 67
inverse Fourier tranform, **54**

Jensen's inequality Lemma, 108

Lagrange multiplier, **25**
Lagrange multipliers, **79**
Lagrange Theorem, 26
Lax-Milgram Theorem, **58**
Lax-Milgram theorem, **3**
Lax-Milgram Theorem Theorem, 58
Legendre ellipticity condition, **57**
Legendre-Hadamard, **110**
Legendre-Hadamard condition, **57**, **73**
linear functionals
  bounded linear functionals, 12
  continuous linear functionals, 12
Lipschitz convergence, **101**
liquid crystals, **82**
Ljusternik Theorem, 26
lower semi-continuous, **4**

maximum, **25**
mean value property, **88**
method of continuity, **17**
method of moving planes, **96**
Meyers-Serrin Theorem, 36
minimization method, **71**
minimizer, **78**
minimizing sequence, **78**
minimum, **25**
Morrey Theorem, 101, 104
Morrey's Inequality Theorem, 47
mountain pass method, **3**

Nemystkii operator, **80**
Neumann boundary, **60**
nonexistence, **81**
nonlinear eigenvalue problem, **79**
nonlinear functional analysis, **17**
norm, **8**
  equivalent, 51

null-Lagrangian, **5**, **82**, **115**
null-Lagrangians, **4**, **110**

operator
  bounded, 11
  compact, 14
  contraction, 17
  Hilbert space adjoint, 13
  Laplace operator, 43
  linear operator, 11
  Nemytskii operator, 18
  operator norm, 11
  resolvent, 15
  strongly continuous, 14
  weakly continuous, 14
ordinary point, **25**
orientation-preserving, **121**
orthogonal, **10**
orthogonal complement, **10**
orthogonality condition, **69**
orthonormal, **15**
  complete, 15
orthonormal basis, **69**
Oseen-Frank energy, **82**
Oseen-Frank model, **4**

Palais-Smale condition, **3**
Palais-Smale condition (PS), **90**
Parallelogram law, **10**
Partition of Unity Theorem, 33
Plancherel's Theorem Theorem, 54
Pohozaev Theorem, 93
Poincaré's inequalities, **52**
Poincaré's inequality, **52**
point spectrum, **16**
polyconvex, **112**, **114**
polyconvexity, **4**
principal eigenvalue, **69**
projection, **10**
Projection Theorem, 10
Property of Fourier Tranforms Theorem, 54
pyramid function, **127**

quasiconvex, **103**
quasiconvex envelope, **124**
quasiconvex functions, **103**
quasiconvexication, **124**
quasiconvexity, **4**, **101**

rank-one convex, **74**, **108**
regularity, **2**
relaxation, **125, 126**
Relaxation Principle Theorem, 126
Rellich-Kondrachov Theorem, 50
Representation by Local Functions Theorem, 86
resolvent set, **15**
Riesz Representation Theorem, 12
Riesz representation theorem, **3**

second variation, **72**
second-order differential operator in divergence form, **55**
semilinear, **78**
seminorm, **51**
sequence

weakly convergent, 13
Serrin's Maximum Principle, **96**
Serrin's Maximum Principle Lemma, 96
set
convex, 7
totally bounded, 49
weakly closed, 22, 77
simplified Maxwell equation, **83**
Sobolev conjugate, **45**
Sobolev space of order k, **30**
Sobolev spaces, **2**
Sobolev-Galiardo-Nirenberg inequality, **89**
Sobolev-Rellich-Kondrachov Theorem, 44
spectrum, **16**
star-shaped, **95**
stored energy density function, **120**
structural conditions, **74**
subdifferential, **127**
subgradient, **127**
symmetric, **16**

Tonelli theorem, **122**
Tonelli's theorem, **82**
total stored energy, **120**
Trace Theorem, 42
Triangle inequality, **8**, **10**

variational problems, **71**
vector space, **7**
Banach space, 8
basis, 8
complete, 8
dimension, 7
inner product space, 9
normed space, 8
subspace, 7
Vitali covering, **106**

weak convergence, **1**
weak derivative, **28**
weak lower semicontinuous envelope, **125**
weak solution, **55**, **65**, **75, 76**
weak solutions, **2**
weakly compact, **14**
weakly lower semicontinuous, **77**, **99**
Weierstrass theorem, **1**