****

Applied Analysis I - (Advanced PDE I)
(Math 940, Fall 2014)

by

Baisheng Yan

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

Contents

Chapter 1

Preliminaries

1.1. Banach Spaces

1.1.1. Vector 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. If, on the other hand, $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$ does not imply $\alpha_i = 0$ for each i, the elements $x_1, \ldots, x_n$ 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 $\{u_1, \ldots, u_n\}$ is a **basis** for X; in all other cases $\dim X = \infty$.

1.1.2. Normed Spaces. 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 $x \neq 0$.

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

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

1.1.3. Completeness and Banach Spaces. 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 $\{x \in \Omega : f(x) \neq 0\}$ and is denoted by $\mathrm{supp}(f)$.) The latter two spaces are clearly subspaces of $C(\Omega)$.

For each n-tuple $\alpha = (\alpha_1, \ldots, \alpha_n)$ 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 $L^p(\Omega), 1 \le p \le \infty$, 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.

Finally 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

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

Example 1.3. 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.

1.1.4. Hilbert 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

$$\|x\| = (x, x)^{1/2}. \tag{1.1}$$

Theorem 1.4. *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 (1.1). □

Furthermore, by (d), equation (1.1) 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 1.5. 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 1.6. *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

$$\|x_n - x_m\|^2 \leq 2(\|x_n\|^2 - d^2) + 2(\|x_m\|^2 - d^2) \tag{1.2}$$

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 (1.2), 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 1.7. (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 1.6 implies that there exists a $y \in M$ such that $\|x - y\| \le \|x - w\|$ 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

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

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

1.2.1. Operators on a Banach Space. 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

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

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

$$\|Tx\| \le m\|x\| \quad \text{for all} \quad x \in X. \tag{1.3}$$

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

$$\|T\| = \sup_{x \in X,\, \|x\|=1} \|Tx\| = \sup_{x \in X,\, \|x\| \le 1} \|Tx\|. \tag{1.4}$$

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

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

Theorem 1.8. *If X and Y are normed spaces, then $\mathcal{B}(X, Y)$ is a normed space with norm defined by equation* (1.4)*. 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

$$\|T_n x - T_m x\| \le \|T_n - T_m\|\|x\| \tag{1.5}$$

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 (1.5) is smaller than $\varepsilon\|x\|$ provided that m and n are large enough. Thus, (letting $n \to \infty$)

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

Hence, $\|Tx\| \le (\|T_m\| + \varepsilon)\|x\|$, 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 in linear functional analysis; see, e.g., [**?**].

Theorem 1.9. (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 1.10. **(Bounded Inverse)** *If X and Y are Banach spaces and if $T \in \mathcal{B}(X,Y)$ is one-to-one and onto, then $T^{-1} \in \mathcal{B}(Y,X)$.*

1.2.2. Dual Spaces and Reflexivity. 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 $|\langle f, x\rangle| \le \|f\|\,\|x\|$ for all $f \in X^*$, $x \in X$.

Example 1.11. 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 can be found in [**?**].

Theorem 1.12. **(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 1.13. *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 $F_x \in X^{**}$ by $F_x(f) = \langle f, x\rangle$ for all $f \in X^*$. From Corollary 1.13, 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 1.14. *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.*

1.2.3. Bounded Linear Functionals on a Hilbert Space.

Theorem 1.15. **(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 $y = \overline{f(z)}z$. 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 $|f(x)| \le \|x\|\|y\|$, which yields $\|f\| \le \|y\|$. The reverse inequality follows by choosing $x = y$ in the representation. □

Corollary 1.16. *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 1.17. *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, $\|T^*y\| = \|z\| = \|f\| \le \|T\|\|y\|$, i.e., $T^* \in \mathcal{B}(H)$ and $\|T^*\| \le \|T\|$. The reverse inequality follows easily from $\|Tx\|^2 = (Tx, Tx) = (x, T^*Tx) \le \|Tx\|\|T^*\|\|x\|$. □

1.3. Weak Convergence and Compact Operators

1.3.1. Weak Convergence. 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 1.18. *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 $\|x\| \le \liminf \|x_n\|$.*

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 1.13, $z = 0$. 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 $\|x_n\| = \|Jx_n\| \le c$ which implies $\{x_n\}$ is bounded. Finally, for $f \in X^*$

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

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

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

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

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

The main result of this section is given by:

Theorem 1.19. *If X is a reflexive Banach space, then the closed unit ball is* **weakly compact***, i.e., the sequence $\{x_n\}$, with $\|x_n\| \leq 1$ has a subsequence which converges weakly to an x with $\|x\| \leq 1$.*

1.3.2. Compact Operators. 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 1.20. (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 1.21. *Let X and Y be Banach spaces. If $T_n : X \to Y$ are linear and compact for $n \geq 1$ and if $\lim_{n\to\infty} \|T_n - T\| = 0$, then T is compact. Thus, compact operators form a closed, but not a dense, 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 $k \geq 2$ 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 $k \geq 2$. Choose $\varepsilon > 0$. Let k be such that $\|T - T_k\|M < \varepsilon/4$ and note that

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

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

Theorem 1.22. *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 1.23. *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 Compact Linear Operators

1.4.1. Fredholm Alternative.

Theorem 1.24. (Fredholm Alternative) *Let $T : H \to H$ be a compact linear 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.*

1.4.2. Spectrum of Compact Operators. 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 $(x, \phi) \ne 0$ 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 1.25. *Let $T : X \to X$ be a compact linear operator on the normed space X. Then for each $r > 0$ there exist at most finitely many $\lambda \in \sigma_p(T)$ for which $|\lambda| > r$.*

1.4.3. Symmetric Compact Operators. 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 1.26. *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 1.27. *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. Nonlinear Functional Analysis

In this final preliminary section, we list some useful results in **nonlinear functional analysis**. Lots of the proof 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

$$\|Tx - Ty\| \leq k\|x - y\| \quad \text{for all} \quad x, y \in X. \tag{1.6}$$

Theorem 1.28. (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

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

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

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

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

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

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. □

Our main existence result will be based upon the following so-called **method of continuity** or **continuation method**.

Theorem 1.29. *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 $x \in X$

$$\|x\|_X \le c\|T_t x\|_Y. \tag{1.7}$$

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

$$[0,s] \subset S \quad \text{implies} \quad [0, s+\tau] \subset S. \tag{1.8}$$

(Note that any smaller τ works.) Since τ can be chosen independently of s, (1.8) 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

$$T_s x = f + \tau T_0 x - \tau T_1 x. \tag{1.9}$$

By (1.7), $T_s^{-1} : Y \to X$ exists and $\|T_s^{-1}\| \le c$. Hence (1.9) is equivalent to

$$x = T_s^{-1}(f + \tau T_0 x - \tau T_1 x) \equiv Ax \tag{1.10}$$

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

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

By the contraction mapping theorem, (1.10) 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 $f : \Omega \times \mathbb{R} \to \mathbb{R}$ be a given function. Assume

(i) for every $\xi \in \mathbb{R}$, $f(x,\xi)$ (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 $\mathbb{R}$

(iii) for all $(x,\xi) \in \Omega \times \mathbb{R}$

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

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

$$Nu(x) = f(x, u(x)),\ x \in \Omega.$$

We have the following result needed later.

Lemma 1.30. *$N : L^p(\Omega) \to L^q(\Omega)$ is continuous and bounded with*

$$\|Nu\|_q \le const\ (\|a\|_q + \|u\|_p^{p/q}) \quad \textit{for all} \quad u \in L^p(\Omega) \tag{1.11}$$

and

$$\langle Nu, v\rangle = \int_\Omega f(x,u(x))v(x)dx \quad \textit{for all} \quad u, v \in L^p(\Omega). \tag{1.12}$$

Here $\langle\cdot,\cdot\rangle$ denotes the duality pairing between $L^p(\Omega)$ and $L^q(\Omega)$.

Proof. Since $u \in L^p(\Omega)$, the function $u(x)$ is measurable on Ω and thus, by (i) and (ii), the function $f(x, u(x))$ is also measurable on Ω. From the inequality $(\sum_{i=1}^n \xi_i)^r \leq c \sum_{i=1}^n \xi_i^r$ and (iii) we get

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

Integrating over Ω and applying the above inequality once more yields (1.11), which shows that N is bounded.

To show that N is continuous, let $u_n \to u$ in $L^p(\Omega)$. 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. Since $Nu \in (L^p(\Omega))^*$, the integral representation (1.12) is clear. □

Remarks. (a) The following remarkable statement can be proved: If f satisfies (i) and (ii) above and if the corresponding Nemytskii operator is such that $N : L^p(\Omega) \to L^q(\Omega)$, then N is continuous, bounded and (iii) holds.

(b) If (iii) is replaced by

(iii)′ for all $(x, \xi) \in \Omega \times \mathbb{R}$

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

where b is a fixed nonnegative number, $a \in L^r(\Omega)$ is nonnegative and $1 < p, r < \infty$, then the above results are valid with q replaced by r. (i.e., $N : L^p(\Omega) \to L^r(\Omega)$.)

(c) We say that f satisfies the **Caratheodory property**, written $f \in$ **Car**, if (i) and (ii) above are met. If in addition (iii) is met, then we write $f \in$ **Car(p)**.

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 1.31. (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 2.11, 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 1.32. *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 $\lambda \in \mathcal{M}$*

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

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

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

Therefore

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

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

Lemma 1.33. *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*

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

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 1.34. *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 1.35. (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 $S = \{[u, v] : \|u - u_0\| \le a,\ \|v - v_0\| \le b\}$. 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

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

for $\|u - u_0\| \le \delta$, $\|v - v_0\| \le \delta$. By virtue of Lemma 1.33, $B(u, \cdot)$ 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 $0 < \varepsilon \le \delta$ such that

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

for $\|u - u_0\| \le \varepsilon$. The existence of a unique fixed point in the closed ball $\overline{B(v_0, \delta)}$ follows from Lemma 1.34 and the continuity from Lemma 1.32. □

Example 1.36. 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 1.18, 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 1.37. *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 $f'(u_0) = 0$.

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 $v \in X$. □

The point $u_0 \in X$ is called a **critical point** of the functional f defined on X if $f'(u_0)v = 0$ for every $v \in X$.

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.

1.5.5.1. *Convex sets.* A set C in the 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 (see, e.g., [**?**]).

Theorem 1.38. *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 on the real Banach space X.

A is **monotone** if

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

A is **strongly monotone** if

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

where $c > 0$ and $p > 1$.

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

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

In the following we set

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

for fixed u and v.

Lemma 1.39. *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

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

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

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

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 $\varphi'(1) \ge \varphi'(0)$ we obtain

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

which implies (c).

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

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

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

Theorem 1.40. *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

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

It follows from (a) that C_r is closed and convex for all r, and thus is weakly closed (cf. Theorem 1.38). 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 $f(u_{n'}) \le r$ (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 1.39, $\varphi : [0,1] \to \mathbb{R}$ is convex and φ' is monotone. By the classical mean value theorem,

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

i.e.,

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

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 an **extreme** of f with respect to M_c if there exists a neighborhood of u_0, $U(u_0) \subset X$, such that

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

or

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

In the first case we say that f is (local) **maximal** at u_0 with respect to M_c, while in the second case f is (local) **minimal** 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 its F-derivative $g'(u_0) \neq 0$.

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 extremum 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 extremum of f with respect to X, then we can choose $\lambda = 0$, which implies the usual result.

Lemma 1.41. *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

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

Then

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

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

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

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

(iii) *u_0 is a local extremum 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 1.43. (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, $i = 1, \dots, n$

(iii) u_0 *is an extremum of* g_0 *with respect to the constraint* C*:*

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

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 1.44. *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 1.37, 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, which is frequently referred to as **Ehrling's inequality**, will be needed in the next result.

Theorem 1.45. *Let X, Y, Z be three Banach spaces such that*

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

Assume that the imbedding $X \subset Y$ is compact and that the imbedding $Y \subset Z$ 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

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

Without loss of generality we can assume $\|u_n\|_X = 1$. Since the imbedding $X \subset Y$ 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. □

Chapter 2

Sobolev Spaces

This chapter is devoted to the study of the necessary Sobolev function spaces which permit a modern approach to partial 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

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

where $v = D^\alpha u$. Motivated by (2.1), we now enlarge the class of functions for which the notion of derivative can introduced.

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 the weak derivative is 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 as we will see in a forthcoming exercise. However, it can be shown that if all weak derivatives exist of a certain order, then all lower order weak derivatives exist.

EXAMPLE 2.1. (a) The function $u(x) = |x_1|$ has in the ball $\Omega = B(0,1)$ weak derivatives $u_{x_1} = \text{sgn } x_1, u_{x_i} = 0, i = 2, \ldots, n$. 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

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

Hence $|x_1|_{x_1} = \text{sgn } x_1$. 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.

(b) By the above computation, the function $u(x) = |x|$ has a weak derivative $u'(x) = \text{sgn } x$ on the interval $\Omega = (-1,1)$. On the other hand, sgn x does not have a weak derivative on Ω due to the discontinuity at $x = 0$.

(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.

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

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

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 let

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

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 on $W^{k,p}(\Omega)$ is introduced by defining

$$\|u\|_{k,p} = \|u\|_{W^{k,p}(\Omega)} = \begin{cases} (\int_\Omega \sum_{|\alpha| \leq k} |D^\alpha u|^p dx)^{1/p} & \text{if } 1 \leq p < \infty, \\ \sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^\infty(\Omega)} & \text{if } p = \infty. \end{cases} \tag{2.3}$$

The space $W^{k,p}(\Omega)$ is known as a **Sobolev space of order k and power p.**

We define the space $W_0^{k,p}(\Omega)$ 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 2.1. The spaces $W^{k,2}(\Omega)$ and $W_0^{k,2}(\Omega)$ are special since they become a Hilbert space 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:

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

Theorem 2.2. *$W^{k,p}(\Omega)$ is a Banach space under the norm (2.3). If $1 < p < \infty$, it is reflexive, and if $1 \leq p < \infty$, it is separable.*

Proof. 1. We first prove that $W^{k,p}(\Omega)$ is complete with respect to the norm (2.3). We prove this for $1 \leq p < \infty$; the case $p = \infty$ is similar. 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 α, $|\alpha| \leq k$, 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 $n \to \infty$. This proves the completeness of $W^{k,p}(\Omega)$; hence it is a Banach space.

2. 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 $v = (v_1, \dots, v_{n+1}) \in (L^p(\Omega))^{n+1}$, then T is a (linear) isometry. Now $(L^p(\Omega))^{n+1}$ is reflexive for $1 < p < \infty$ and separable for $1 \leq p < \infty$. 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 $W^{1,p}(\Omega)$ (see Theorem 1.14). Similarly, we can handle the case $k \geq 2$. □

The following result is of independent importance. We omit the proof.

Theorem 2.3. *$u_n \rightharpoonup u$ in $W^{k,p}(\Omega)$ if and only if $D^\alpha u_n \rightharpoonup D^\alpha u$ in $L^p(\Omega)$ for all $|\alpha| \leq k$.*

EXAMPLE 2.4. Let Ω be a bounded open connected set in $\mathbb{R}^n$. Divide Ω into N open disjoint subsets $\Omega_1, \Omega_2, \dots, \Omega_N$. 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 $\Omega_1, \Omega_2, \dots, \Omega_N$, and can be extended continuously to $\bar{\Omega}_1, \bar{\Omega}_2, \dots, \bar{\Omega}_N$, 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 2.2. 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.

Exercise 2.3. (a) Consider the function $u(x) = \text{sgn} x_1 + \text{sgn} x_2$ in the ball $B(0,1) \subset \mathbb{R}^2$. Show that the weak derivative u_{x_1} does not exist, yet the weak derivative $u_{x_1 x_2}$ does exist.

(b) Let Ω be the hemisphere of radius $R < 1$ in $\mathbb{R}^n$:

$$r^2 \equiv \sum_{i=1}^n x_i^2 \leq R^2, \ \ x_n \geq 0 \ \ (n \geq 3).$$

Show that $u = (r^{(n/2-1)} \ln r)^{-1} \in H^1(\Omega)$.

(c) Let $B = B(0,1)$ be the open unit ball in $\mathbb{R}^n$, and let

$$u(x) = |x|^{-\alpha}, \; x \in B.$$

For what values of α, n, p does u belong to $W^{1,p}(B)$?

2.2. Approximations and Extensions

2.2.1. Mollifiers. 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

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

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

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

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

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.

Remark 2.4. 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 2.5. *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. □

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 2.6. *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*

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

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

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

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

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. $\square$

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

We need the following well-known result.

Theorem 2.7. (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 $x \in \Omega$.*

2.2.2. Approximation Theorems.

Lemma 2.8. *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 $x \in B(0,2a) \setminus B(0,a)$. 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 function $\bar{U}$ is uniformly continuous in $B(0,2a)$, there is a $0 < \delta < a$ 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 2.9. *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) $\|u_h - u\|_p \to 0 \quad as \quad h \to 0.$

If $u \in C^k(\bar{\Omega})$ and $\bar{\Omega}$ is compact, then, for all $\Omega' \subset\subset \Omega$,

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

Proof. 1. 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

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

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

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

which implies (a).

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

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

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

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

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

and so by Fubini's Theorem

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

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

3. 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. □

Remark 2.5. In (c) of the theorem above, 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 2.10. *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 2.5(b) and 2.9(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 2.11. *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 there exists a sequence of $C^{|\alpha|}(\Omega)$ functions $\{u_h\}$ with $\|u_h - u\|_{L^1(S)} \to 0$, $\|D^\alpha u_h - v\|_{L^1(S)} \to 0$ as $h \to 0$, for all compact sets $S \subset \Omega$.*

Proof. 1. (Necessity) Suppose $v = D^\alpha u$. Let $S \subset \Omega$, 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 2.9 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 2.5 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, $\|D^\alpha u_h - v\|_{L^1(S)} \to 0$.

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

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

which is the claim. □

Theorem 2.12. *Let Ω be a domain 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 2.11, $D^\alpha u_h(x) = (D^\alpha u)_h(x) = 0$ 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

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

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. (See Meyers and Serrin $H = W$. The proofs make use of a partition of unity argument.)

Theorem 2.13. *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 2.14. *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 $W^{k,p}(\Omega)$.

Exercise 2.6. 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 $L^q(\Omega)$ $(p > 1, 1/p + 1/q = 1)$.

Exercise 2.7. (a) If $u \in W_0^{k,p}(\Omega)$ and $v \in C^k(\bar{\Omega})$, prove that $uv \in W_0^{k,p}(\Omega)$.

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

2.2.3. Chain Rules.

Theorem 2.15. *Let Ω be an open set in $\mathbb{R}^n$. Let $f \in C^1(\mathbb{R})$, $|f'(s)| \leq M$ 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. 1. According to Theorem 2.11, 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 2.11 again.

2. If $f(0) = 0$, 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 $\varphi \in C_0^1(\Omega)$

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

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

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

which establishes the desired result for u^+. □

The next result extends the result on $|u|, u^+$ and u^-.

Theorem 2.17. *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

$$\begin{aligned} |v_n(x) - v_n(y)| &= |f(u_n(x)) - f(u_n(y))| \\ &\leq c|u_n(x) - u_n(y)| \leq c_n|x - y|. \end{aligned}$$

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 $|\partial v_n/\partial x_i| \leq c_n$ for $1 \leq i \leq n$, 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

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

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

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

and so

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

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 $1 \leq i \leq n$. 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, $f \circ u \in W_0^{1,p}(\Omega)$. □

Corollary 2.18. *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^+, $u^- \in W_0^{1,p}(\Omega)$. □

2.2.4. Extensions. If $\Omega \subset \Omega'$, then any function $u(x) \in C_0^k(\Omega)$ has an obvious extension $U(x) \in C_0^k(\Omega')$. 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.) However, if $u \in W^{k,p}(\Omega)$ has compact support in Ω, then $u \in W_0^{k,p}(\Omega)$ and thus the obvious extension belongs to $W_0^{k,p}(\Omega')$.

We now consider a more general extension result.

Theorem 2.19. *Let Ω be a bounded open set in $\mathbb{R}^n$ with $\Omega \subset\subset \Omega'$ and assume $k \geq 1$.*

(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,*

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

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,*

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

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*

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

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

Proof. 1. 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 $B = B(x^0, r)$ 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 $x \in B \cap \Omega$.)

2. 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

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

3. 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 Ω'. 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 Ω' and has the required properties. Thus (b) is established.

4. We now prove (a). If $u \in W^{k,p}(\Omega)$, then by Theorem 2.14, 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.

5. We now prove (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. □

2.2.5. Trace Theorem. 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

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

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

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

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

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

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

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

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

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

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

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

An integration with respect to t yields

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

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

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

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). □

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 2.21. (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 $\overline{R(\gamma_0)} = L^p(\partial\Omega)$.*

Proof. 1. Suppose $u \in C(\bar{\Omega}) \cap W^{1,p}(\Omega)$. Then by Theorem 2.19, 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, $\gamma_0 u = u|_{\partial\Omega}$.

2. 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.

3. 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.

4. To see that $\overline{R(\gamma_0)} = L^p(\partial\Omega)$, 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 $U \in W^{1,p}(\Omega)$. □

Remark 2.8. We 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

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

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 $k \geq 2$, $\partial u/\partial n$ 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 2.22. (Higher-order traces) *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.

2.2.6. Green's Identities. In this section we assume that $p = 2$ and we continue to assume Ω is a bounded domain.

Theorem 2.23. (Integration by Parts) *Let $u, v \in H^1(\Omega)$ and let $\partial\Omega \in C^1$. Then for any $i = 1, \ldots, n$*

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

($D_i u, D_i v$ are weak derivatives.)

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 2.24. *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. □

Exercise 2.9. 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 embedding 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 general embedding theorem can be stated as follows.

Theorem 2.25. (Sobolev Inequalities) *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 embedding; that is,

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

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

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

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

and

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

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 embedding; that is,

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

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*

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

is a continuous embedding for $d \leq 1/q$.

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

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.

2.3.1. Gagliardo-Nirenberg-Sobolev Inequality. Suppose $1 \le p < n$. Do there exist constants $C > 0$ and $1 \le q < \infty$ such that

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

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

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

and so

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

But then if $1 - n/p + n/q > 0$ (or < 0), we can upon sending λ to 0 (or ∞) in (2.13) obtain a contradiction ($u = 0$). Thus we must have $1 - n/p + n/q = 0$; that is, $q = p^*$, 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 2.26. (Gagliardo-Nirenberg-Sobolev Inequality) *Assume $1 \le p < n$. Then there is a constant C, depending only on p and n, such that*

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

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

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

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

and so

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

Consequently

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

Integrate this inequality with respect to x_1:

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

the last inequality resulting from the extended Hölder inequality in the appendix.

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

$$\begin{aligned}
\int_{\mathbb{R}^n} |u(x)|^{\frac{n}{n-1}} dx &\leq \prod_{i=1}^{n} \left(\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |\nabla u| dx_1 \dots dy_i \dots dx_n \right)^{\frac{1}{n-1}} \\
&= \left(\int_{\mathbb{R}^n} |\nabla u| dx \right)^{\frac{n}{n-1}}
\end{aligned}$$

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. First note that

$$(D_i|u|^\gamma)^2 = \begin{cases} (\gamma u^{\gamma-1} D_i u)^2 & \text{if } u \geq 0 \\ (-\gamma(-u)^{\gamma-1} D_i u)^2 & \text{if } u \leq 0 \end{cases} = (\gamma |u|^{\gamma-1} D_i u)^2.$$

Thus $v \in C_0^1(\mathbb{R}^n)$, and

$$\begin{aligned}
\left(\int_{\mathbb{R}^n} |u(x)|^{\frac{\gamma n}{n-1}} dx \right)^{\frac{n-1}{n}} &\leq \int_{\mathbb{R}^n} |\nabla |u|^\gamma| dx \\
&= \gamma \int_{\mathbb{R}^n} |u|^{\gamma-1} |\nabla u| dx \\
&\leq \gamma \left(\int_{\mathbb{R}^n} |u|^{\frac{p(\gamma-1)}{p-1}} dx \right)^{\frac{p-1}{p}} \left(\int_{\mathbb{R}^n} |\nabla u|^p dx \right)^{\frac{1}{p}}.
\end{aligned}$$

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

$$\left(\int_{\mathbb{R}^n} |u|^{p^*} dx \right)^{\frac{1}{p^*}} \leq C \left(\int_{\mathbb{R}^n} |\nabla u|^p dx \right)^{\frac{1}{p}}.$$

□

Theorem 2.27. *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 2.26,

$$\|u_m - u_l\|_{L^{p^*}(\mathbb{R}^n)} \le C\|\nabla u_m - \nabla u_l\|_{L^p(\mathbb{R}^n)}$$

for all $l, m \ge 1$; whence $u_m \to U$ in $L^{p^*}(\mathbb{R}^n)$ as well. Since Lemma 2.26 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 2.28. *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 $\mathbb{R}^n \backslash \bar{\Omega}$ and apply Lemma 2.26 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 $q \in [1, p^*]$. □

2.3.2. Morrey's Inequality. 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 2.29. (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 $u \in C^1(\mathbb{R}^n)$,

$$\int_{B(x,r)} |u(y) - u(x)|\, dy \le \frac{r^n}{n} \int_{B(x,r)} \frac{|\nabla u(y)|}{|x-y|^{n-1}}\, dy. \tag{2.21}$$

To prove this, note that, for any w with $|w| = 1$ and $0 < s < r$,

$$\begin{aligned} |u(x+sw) - u(x)| &= \left| \int_0^s \frac{d}{dt} u(x+tw) dt \right| \\ &= \left| \int_0^s \nabla u(x+tw) \cdot w dt \right| \\ &\leq \int_0^s |\nabla u(x+sw)| \, dt. \end{aligned}$$

Now we integrate w over $\partial B(0,1)$ to obtain

$$\begin{aligned} \int_{\partial B(0,1)} |u(x+sw) - u(x)| \, dS &\leq \int_0^s \int_{\partial B(0,1)} |\nabla u(x+sw)| \, dS dt \\ &= \int_{B(x,s)} \frac{|\nabla u(y)|}{|x-y|^{n-1}} \, dy \\ &\leq \int_{B(x,r)} \frac{|\nabla u(y)|}{|x-y|^{n-1}} \, dy. \end{aligned}$$

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 $x \in \mathbb{R}^n$,

$$\begin{aligned} |u(x)| &\leq \frac{1}{|B(x,1)|} \int_{B(x,1)} |u(y) - u(x)| \, dy + \frac{1}{|B(x,1)|} \int_{B(x,1)} |u(y)| \, dy \\ &\leq C \left(\int_{\mathbb{R}^n} |\nabla u(y)|^p \, dy \right)^{1/p} \left(\int_{B(x,1)} |y-x|^{\frac{(1-n)p}{p-1}} \, dy \right)^{\frac{p-1}{p}} + C\|u\|_{L^p(\mathbb{R}^n)} \\ &\leq C\|u\|_{W^{1,p}(\mathbb{R}^n)}. \end{aligned}$$

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

$$|u(x) - u(y)| \leq \frac{1}{|W|} \int_W |u(x) - u(z)| \, dz + \frac{1}{|W|} \int_W |u(y) - u(z)| \, dz. \tag{2.22}$$

Note that $|W| = \beta r^n$, $r = |x-y|$ and $\int_W \leq \min\{\int_{B(x,r)}, \int_{B(y,r)}\}$. Hence, using (2.21), by Hölder's inequality, we obtain

$$\begin{aligned} \int_W |u(x) - u(z)| \, dz &\leq \int_{B(x,r)} |u(x) - u(z)| \, dz \leq \frac{r^n}{n} \int_{B(x,r)} |Du(z)||x-z|^{1-n} \, dz \\ &\leq \frac{r^n}{n} \left(\int_{B(x,r)} |\nabla u(z)|^p \, dz \right)^{1/p} \left(\int_{B(x,r)} |z-x|^{\frac{(1-n)p}{p-1}} \, dz \right)^{\frac{p-1}{p}} \\ &\leq C\, r^n \, \|\nabla u\|_{L^p(\mathbb{R}^n)} \left(\int_0^r s^{\frac{(1-n)p}{p-1}} s^{n-1} \, ds \right)^{\frac{p-1}{p}} \\ &\leq C\, r^{n+\gamma} \|\nabla u\|_{L^p(\mathbb{R}^n)}, \end{aligned}$$

where $\gamma = 1 - \frac{n}{p}$; similarly,

$$\int_W |u(y) - u(z)| \, dz \leq C\, r^{n+\gamma} \|\nabla u\|_{L^p(\mathbb{R}^n)}.$$

Hence, by (2.22),

$$|u(x) - u(y)| \leq C\, |x-y|^\gamma \|\nabla u\|_{L^p(\mathbb{R}^n)}.$$

This inequality and the bound on $\|u\|_{C^0}$ above complete the proof. □

Theorem 2.30. (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, after possibly redefining u on a null set, $u \in C^{0,1-\frac{n}{p}}(\bar{\Omega})$ 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. □

2.3.3. General Cases. We can now concatenate the above estimates to obtain more complicated 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.

2.4. Compactness

We now consider the compactness of the embeddings. Note that if X and Y are Banach spaces with $X \subset Y$ then we say that X is **compactly embedded** in Y, written $X \subset\subset Y$, provided

(i) $\|u\|_Y \leq C\|u\|_X$ $(u \in X)$ for some constant C; that is, the embedding is continuous;

(ii) each bounded sequence in X has a convergent subsequence in Y.

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 2.31. (Rellich-Kondrachov) *Let $\Omega \subset \mathbb{R}^n$ be bounded and open. Then for $1 \leq p < n$:*

(a) *The embedding $W_0^{1,p}(\Omega) \subset L^q(\Omega)$ is compact for each $1 \leq q < np/(n-p)$.*

(b) *Assuming $\partial\Omega \in C^1$, the embedding $W^{1,p}(\Omega) \subset L^q(\Omega)$ is compact for each $1 \leq q < np/(n-p)$.*

(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 embedding $W^{1,p}(\Omega) \subset C^{0,\alpha}(\bar{\Omega})$ is compact for each $0 \leq \alpha < 1-(n/p)$.*

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 embedding $C_0^1(\Omega) \subset L^2(\Omega)$ is compact. Thus, let $S = \{u \in C_0^1(\Omega) : \|u\|_{1,2} \leq 1\}$. 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

$$|u_h(x)| \leq \int_{B(0,h)} \omega_h(z)|u(x-z)|dz \leq (\sup\omega_h)\|u\|_1 \leq c_1(\sup\omega_h)\|u\|_{1,2}$$

and

$$|D_i u_h(x)| \leq c_2 \sup|D_i\omega_h|\|u\|_{1,2}, \quad i = 1,\ldots,n$$

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

$$\|u_h - u\|_2^2 \leq \int_{B(0,h)} \omega_h(z)\left(\int_\Omega |u(x-z)-u(x)|^2 dx\right) dz$$

and

$$\begin{aligned}
\int_\Omega |u(x-z)-u(x)|^2dx &= \int_\Omega \left| \int_0^1 \frac{du(x-tz)}{dt}dt \right|^2 dx \\
&= \int_\Omega \left| \int_0^1 (-\nabla u(x-tz)\cdot z)dt \right|^2 dx \\
&\le \int_\Omega |z|^2 \left(\int_0^1 |\nabla u(x-tz)|^2 dt \right) dx \le |z|^2 \|u\|_{1,2}^2.
\end{aligned}$$

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 $m, n \ge N$, $\|u_m - u_n\|_2 < \varepsilon$. 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 embedding is continuous if $\alpha = 1 - (n/p)$. Now use the fact that $C^{0,\beta}$ is compact in $C^{0,\alpha}$ if $\alpha < \beta$. □

Remarks. (a) When $p = n$, we can easily show that the embedding in (a) is compact for all $1 \le q < \infty$. Hence, it follows that the embedding $W_0^{1,p}(\Omega) \subset L^p(\Omega)$ is compact for all $p \ge 1$. However, when $p = n$, we do not have embedding $W^{1,n}(\Omega) \subset L^\infty(\Omega)$. For example, $u = \ln\ln(1 + \frac{1}{|x|}) \in W^{1,n}(B(0,1))$ but not to $L^\infty(B(0,1))$ if $n \ge 2$.

(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 $1 \le q \le \infty$. Further, if

$$\|f\|_{L^q(\mathbb{R})} = \|f\|_{L^q(I)} = a > 0$$

then for any $j \ne k$ 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 embeddings $W^{1,p}(\mathbb{R}) \subset L^q(\mathbb{R})$ can be compact. This example generalizes to n dimensional space and to open sets like a half-space.

2.5. Additional Topics

2.5.1. Equivalent Norms of $W^{1,p}(\Omega)$. 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

$$\|x\| \le c_1|x| \le c_2\|x\| \quad \text{for all } x \in X.$$

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 2.32. *Let $\partial\Omega \in C^1$ and let $1 \le p < \infty$. 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 $u \in W^{1,p}(\Omega)$*

$$q(u) \le d\|u\|_{1,p}.$$

(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

$$\|u\|_{1,p} \le c\|u\| \quad \text{for all} \quad u \in W^{1,p}(\Omega). \tag{2.25}$$

Suppose (2.25) is false. Then there exist $v_n(x) \in W^{1,p}(\Omega)$ such that $\|v_n\|_{1,p} > n\|v_n\|$. Set $u_n = v_n/\|v_n\|_{1,p}$. So

$$\|u_n\|_{1,p} = 1 \quad \text{and} \quad 1 > n\|u_n\|. \tag{2.26}$$

According to Theorem 2.31, 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 $q(u_n) \to 0$. From $u_n \to u$, $\nabla u_n \to 0$ both in $L^p(\Omega)$, we have $\nabla u = 0$ a.e. in Ω and hence $u = C$, a constant. This also implies $u_n \to C$ in $W^{1,p}(\Omega)$ and, by $\|u_n\|_{1,p} = 1$, $C \ne 0$. By continuity, $q(C) = 0$, which implies $C = 0$ by (ii). We thus derive a contradiction. ☐

Example 2.33. Let $\partial\Omega \in C^1$. Assume $a(x) \in C(\overline{\Omega}), \sigma(x) \in C(\partial\Omega)$ with $a \ge 0$ ($\not\equiv 0$), $\sigma \ge 0$ ($\not\equiv 0$). Then the following norms are equivalent to $\|\cdot\|_{1,p}$ on $W^{1,p}(\Omega)$:

$$\|u\| = \left(\int_\Omega \sum_{i=1}^n |D_i u|^p dx + \left|\int_\Omega u dx\right|^p\right)^{1/p} \quad \text{with } q(u) = \left|\textstyle\int_\Omega u dx\right|. \tag{2.27}$$

$$\|u\| = \left(\int_\Omega \sum_{i=1}^n |D_i u|^p dx + \left|\int_{\partial\Omega} \gamma_0 u dS\right|^p\right)^{1/p} \quad \text{with } q(u) = \left|\textstyle\int_{\partial\Omega} \gamma_0 u dS\right|. \tag{2.28}$$

$$\|u\| = \left(\int_\Omega \sum_{i=1}^n |D_i u|^p dx + \int_{\partial\Omega} \sigma|\gamma_0 u|^p dS\right)^{1/p} \quad \text{with } q(u) = \left(\textstyle\int_{\partial\Omega} \sigma|\gamma_0 u|^p dS\right)^{1/p}. \tag{2.29}$$

(2.30) $$\|u\| = \left(\int_\Omega \sum_{i=1}^n |D_i u|^p dx + \int_\Omega a|u|^p dx\right)^{1/p} \qquad \text{with } q(u) = \left(\int_\Omega a|u|^p dx\right)^{1/p}.$$

Clearly property (ii) of Theorem 2.32 is satisfied for these semi-norms $q(u)$. In order to verify condition (i), one uses the trace theorem in (2.28) and (2.29).

2.5.2. Poincaré's Inequalities. Using $u - (u)_\Omega$, where $(u)_\Omega = ⨍_\Omega u dx$, in the equivalent norm (2.27) we obtain that

(2.31) $$\int_\Omega |u(x) - (u)_\Omega|^p dx \le c \int_\Omega \sum_{i=1}^n |D_i u|^p dx, \quad u \in W^{1,p}(\Omega)$$

where the constant $c > 0$ is independent of u. This inequality is often referred to as **Poincaré's inequality**.

We also note that if $u \in W_0^{1,p}(\Omega)$ then the equivalent norm (2.28) implies

(2.32) $$\int_\Omega |u(x)|^p dx \le c \int_\Omega \sum_{i=1}^n |D_i u|^p dx, \quad u \in W_0^{1,p}(\Omega),$$

where the constant $c > 0$ is independent of u. This is also called a **Poincaré's inequality**. Therefore

$$\|u\|_{1,p,0} = \left(\int_\Omega \sum_{i=1}^n |D_i u|^p dx\right)^{1/p}$$

defines an equivalent norm on $W_0^{1,p}(\Omega)$.

2.5.3. Difference Quotients. For later regularity theory, we will be forced to study the difference quotient approximations to weak derivatives.

Assume $u\colon \Omega \to \mathbb{R}^N$ is locally integrable. Let $\{e_1, \cdots, e_n\}$ be the standard basis of $\mathbb{R}^n$. Define the ***i*th-difference quotient of size** h of u by

$$D_i^h u(x) = \frac{u(x + he_i) - u(x)}{h}, \quad h \ne 0.$$

Then $D_i^h u$ is defined on $\Omega_{h,i} = \{x \in \Omega \mid x + he_i \in \Omega\}$. Note that

$$\Omega_h = \{x \in \Omega \mid \operatorname{dist}(x; \partial\Omega) > |h|\} \subset \Omega_{h,i}.$$

We have the following properties of $D_i^h u$.

1) If $u \in W^{1,p}(\Omega; \mathbb{R}^N)$ then $D_i^h u \in W^{1,p}(\Omega_{h,i}; \mathbb{R}^N)$ and

$$D(D_i^h u) = D_i^h (Du) \quad \text{on } \Omega_{h,i}.$$

2) If either u or v has compact support $\Omega' \subset\subset \Omega$ then the **integration-by-parts formula** for difference quotient holds:

$$\int_\Omega u \cdot D_i^h v\, dx = -\int_\Omega v \cdot D_i^{-h} u\, dx \quad \forall\, |h| < \operatorname{dist}(\Omega'; \partial\Omega).$$

3) $D_i^h(\phi\, u)(x) = \phi(x)\, D_i^h u(x) + u(x + he_i)\, D_i^h \phi(x)$.

Theorem 2.34. (Difference quotient and weak derivatives) *(a) Let $u \in W^{1,p}(\Omega)$. Then $D_i^h u \in L^p(\Omega')$ for any $\Omega' \subset\subset \Omega$ satisfying $|h| < \operatorname{dist}(\Omega'; \partial\Omega)$. Moreover, we have*

$$\|D_i^h u\|_{L^p(\Omega')} \le \|D_i u\|_{L^p(\Omega)}.$$

(*b*) *Let* $u \in L^p(\Omega)$, $1 < p < \infty$, *and* $\Omega' \subset\subset \Omega$. *If there exists a constant* $K > 0$ *such that*

$$\liminf_{h\to 0} \|D_i^h u\|_{L^p(\Omega')} \le K,$$

then the weak derivative $D_i u$ *exists and satisfies* $\|D_i u\|_{L^p(\Omega')} \le K$.

Proof. (*a*) Let us suppose initially that $u \in C^1(\Omega) \cap W^{1,p}(\Omega)$. Then, for $h > 0$,

$$D_i^h u(x) = \frac{1}{h}\int_0^h D_i u(x + the_i)\,dt,$$

so that by Hölder's inequality

$$|D_i^h u(x)|^p \le \frac{1}{h}\int_0^h |D_i u(x + the_i)|^p\,dt,$$

and hence

$$\int_{\Omega'} |D_i^h u(x)|^p\,dx \le \frac{1}{h}\int_0^h \int_{B_h(\Omega')} |D_i u|^p\,dx\,dt \le \int_\Omega |D_i u|^p\,dx,$$

where $B_h(\Omega') = \{x \in \Omega \,|\, \mathrm{dist}(x;\Omega') < h\}$. The extension of this inequality to arbitrary functions in $W^{1,p}(\Omega)$ follows by a straight-forward approximation argument.

(*b*) Since $1 < p < \infty$, there exists a sequence $\{h_m\}$ tending to zero and a function $v \in L^p(\Omega')$ with $\|v\|_p \le K$ such that $D_i^{h_m} u \rightharpoonup v$ in $L^p(\Omega')$ as $m \to \infty$. This means for all $\phi \in C_0^\infty(\Omega')$

$$\lim_{m\to\infty} \int_{\Omega'} \phi\, D_i^{h_m} u\,dx = \int_{\Omega'} \phi\, v\,dx.$$

Now for $|h_m| < \mathrm{dist}(\mathrm{supp}\,\phi; \partial\Omega')$, we have

$$\int_{\Omega'} \phi\ D_i^{h_m} u\,dx = -\int_{\Omega'} u\ D_i^{-h_m}\phi\,dx \to -\int_{\Omega'} u\, D_i\phi\,dx.$$

Hence

$$\int_{\Omega'} \phi\, v\,dx = -\int_{\Omega'} u\, D_i\phi\,dx,$$

which shows $v = D_i u \in L^p(\Omega')$ and $\|D_i u\|_{L^p(\Omega')} \le K$. □

Remark 2.10. Variants of Theorem 2.34 can be valid even if it is not the case $\Omega' \subset\subset \Omega$. For example if Ω is the open half-ball $B(0,1) \cap \{x_n > 0\}$, $\Omega' = B(0,1/2) \cap \{x_n > 0\}$, and if $u \in W^{1,p}(\Omega)$, then we have the bound

$$\|D_i^h u\|_{L^p(\Omega')} \le \|D_i u\|_{L^p(\Omega)}$$

for $i = 1, 2, \cdots, n-1$ and $0 < |h| < \mathrm{dist}(\Omega', \partial\Omega)$.

We will need this remark for boundary regularity later.

2.5.4. Fourier Transform Methods. 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 transform** 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 2.35. (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 2.36. (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 2.37. *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 embedding, which is known valid for integers s by the previous Sobolev embedding theorem,

$$H^s(\mathbb{R}^n) \subset L^\infty(\mathbb{R}^n) \quad \text{if } s > \tfrac{n}{2}.$$

2.6. Spaces of Functions Involving Time

We study spaces of functions mapping time into Banach spaces. These will be essential for the study of weak solutions to evolution equations later.

2.6.1. Calculus of Abstract Functions. Let X be a real Banach space with norm $\|\cdot\|$ and let I be any interval of the real line $\mathbb{R}$.

Definition 2.11. (i) A function $u : I \to X$ is called an **abstract function**.

(ii) An abstract function $u : I \to X$ is said to be **continuous** at a point $t_0 \in I$ if

$$\lim_{t\to t_0} \|u(t) - u(t_0)\| = 0.$$

(If t_0 is an end point of I, the continuity at t_0 is defined through the one-sided limit.) If $u(t)$ is continuous at each point of I, then we write $u \in C(I; X)$.

(iii) Abstract function $u : I \to X$ is said to be **differentiable** at the point $t_0 \in I$ if there exists an element $l = u'(t_0) \in X$ such that

$$\lim_{h\to 0} \|[u(t_0 + h) - u(t_0)]/h - u'(t_0)\| = 0.$$

We say $u(t)$ is **differentiable on** I if it is differentiable at each point of I.

Remark 2.12. (i) If $u : I \to X$ is continuous at $t_0 \in I$ then real-valued function $\|u(t)\|$ is continuous at t_0.

(ii) If $I = [a, b]$ is a compact interval, then $C([a, b]; X)$ becomes a Banach space with norm

$$\|u\|_{C([a,b];X)} := \max_{t\in[a,b]} \|u(t)\|.$$

Theorem 2.38. (Mean Value Theorem) *Let $u(t) \in C([a, b]; X)$ and suppose $u'(t)$ exists for every $t \in (a, b)$. Then*

$$\|u(a) - u(b)\| \le (b - a) \sup_{a<t<b} \|u'(t)\|.$$

Proof. We use a standard device which reduces the problem to the classical case. Namely, consider the real-valued function $\phi(t) = f(u(t))$, where $f \in X^*$. Since f is continuous and linear we have

$$\phi'(t) = \lim_{h\to 0} f\left(\frac{u(t + h) - u(t)}{h}\right) = f(u'(t)).$$

Now apply the classical mean value theorem to $\phi(\cdot)$ to get

$$\begin{aligned}\|u(b) - u(a)\| &= \sup_{\|f\|=1} f(u(b) - u(a)) \\ &= \sup_{\|f\|=1} (\phi(b) - \phi(a)) \\ &= \sup_{\|f\|=1} (b - a)f(u'(t_0)) \quad t_0 \in (a, b) \\ &\le (b - a) \sup_{a<t<b} \|u'(t)\|.\end{aligned}$$

□

Definition 2.13. Let $u : [a, b] \to X$ be an abstract function and define the partial sums

$$S_Z = \sum_{i=1}^{n} u(\bar{t_i})(t_i - t_{i-1}), \quad t_{i-1} \le \bar{t_i} \le t_i,$$

where Z is the partition $a = t_0 < t_1 < \cdots < t_n = b$ of $[a, b]$ and $\Delta Z = \max_i(t_i - t_{i-1})$ is the mesh of the partition. We define the **Riemann integral**

$$\int_a^b u(t)dt = \lim_{k\to\infty} S_{Z_k}$$

if such a common limit value exists for all sequences of partitions $\{Z_k\}$ for which $\Delta Z_k \to 0$ as $k \to \infty$.

Theorem 2.39. *If $u(t) \in C([a,b];X)$, then the Riemann integral exists.*

Proof. As in the classical proof, one uses the uniform continuity of $u(t)$ together with the completeness of X. We shall omit the details. □

Theorem 2.40. *Let $u(t) : [a,b] \to X$ be continuous. Then the following hold:*

(a)
$$\| \int_a^b u(t)dt\| \le \int_a^b \|u(t)\|dt$$

(b)
$$f\left(\int_a^b u(t)dt\right) = \int_a^b f(u(t))dt \quad \textit{for all} \quad f \in X^*$$

(c)
$$\frac{d}{dt}\int_a^t u(s)ds = u(t) \quad \textit{for all} \quad a \le t \le b$$

(d) *If $u'(t) \in C((a,b);X)$, then for any $\alpha, \beta \in (a,b)$*
$$\int_\alpha^\beta u'(s)ds = u(\beta) - u(\alpha)$$

Proof. (a) and (b) follow by passing to the limit in the corresponding relations of Riemann sums.

(c) Set $v(t) = \int_a^t u(s)ds$. Since $u(t)$ is uniformly continuous on $[a,b]$, we have

$$\begin{aligned} \|[v(t+h) - v(t)]/h - u(t)\| &= \|h^{-1}\int_t^{t+h}[u(s) - u(t)]ds\| \\ &\le \max_{|s-t|\le|h|}\|u(s) - u(t)\| \to 0 \quad \text{as} \quad h \to 0. \end{aligned}$$

(d) Let $\phi(t) = f(u(t))$, where $f \in X^*$. Then by using (b), we obtain

$$f\left(u(\beta) - u(\alpha) - \int_\alpha^\beta u'(t)dt\right) = 0 \quad \text{for all} \quad f \in X^*.$$

The result now follows easily. □

2.6.2. Measurable Functions and Sobolev Spaces. We now extend the continuous functions to measurable functions. In what follows, we assume I is a bounded interval in $\mathbb{R}$ and X is a Banach space.

Definition 2.14. (i) A function $s\colon I \to X$ is called **simple** if it has the form

$$s(t) = \sum_{i=1}^m \chi_{E_i}(t)u_i \quad (t \in I),$$

where E_i is a Lebesgue measurable subset of I and $u_i \in X$ for $i = 1, 2, \cdots, m$. In this case, we define

$$\int_I s(t)\,dt = \sum_{i=1}^m |E_i|u_i \in X.$$

(ii) A function $f\colon I \to X$ is **strongly measurable** if there exist simple functions $s_k\colon I \to X$ such that

$$s_k(t) \to f(t) \quad \forall \text{ a.e. } t \in I.$$

(iii) A function $f\colon I \to X$ is **weakly measurable** if for each $u^* \in X^*$ the function $t \mapsto \langle u^*, f(t)\rangle$ is Lebesgue measurable on I.

(iv) A function $f\colon I \to X$ is **almost separably valued** if there exists a subset $N \subset I$ with $|N| = 0$ such that the set $\{f(t) \mid t \in I \setminus N\}$ is separable (that is, has a countable dense subset).

(v) A strongly measurable function $f\colon I \to X$ is **(Bochner) integrable** if there exists a sequence of simple functions $s_k\colon I \to X$ such that

$$\lim_{k\to\infty} \int_I \|s_k(t) - f(t)\|\, dt = 0.$$

In this case, we define

$$\int_I f(t)\, dt = \lim_{k\to\infty} \int_I s_k(t)\, dt \in X.$$

Theorem 2.41. (Bochner-Pettis Theorem) *(i) A function $f\colon I \to X$ is strongly measurable if and only if f is weakly measurable and almost separably valued.*

(ii) A strongly measurable function $f\colon I \to X$ is Bochner integrable if and only if $\|f(t)\|$ is Lebesgue integrable; that is,

$$\|f\|_{L^1(I;X)} := \int_I \|f(t)\|\, dt < \infty.$$

Remark 2.15. (i) The space $L^p(I;X)$ consists of all **strongly measurable** functions $u\colon I \to X$ with

$$\|u\|_{L^p(I;X)} := \left(\int_I \|u(t)\|^p\, dt\right)^{1/p} < \infty$$

if $1 \le p < \infty$ and

$$\|u\|_{L^\infty(I;X)} := \operatorname{esssup}_{t\in I} \|u(t)\| < \infty$$

if $p = \infty$. As in the usual Lebesgue space cases, we identify functions that are almost everywhere equal. Then $L^p(I;X)$ becomes a Banach space for all $1 \le p \le \infty$.

(ii) If X is reflexive, then we have

$$(L^p(I;X))^* \approx L^q(I;X^*) \quad (1 < p < \infty,\ q = \frac{p}{p-1}).$$

However, usually, $(L^1(I;X))^* \not\approx L^\infty(I;X^*)$. In fact, if X is a separable Banach space, then

$$(L^1(I;X))^* \approx L^\infty_w(I;X^*),$$

where $L^\infty_w(I;X^*)$ consists of functions $g\colon I \to X^*$ such that for each $u \in X$ the function $t \mapsto \langle g(t), u\rangle$ is Lebesgue measurable and essentially bounded on I with the norm

$$\|g\|_w := \sup_{u\in X, \|u\|\le 1} \|\langle g(t), u\rangle\|_{L^\infty(I)} < \infty.$$

(iii) The space $C(\bar I;X)$ consists of all continuous functions $u\colon \bar I \to X$ with

$$\|u\|_{C(\bar I;X)} := \max_{t\in\bar I} \|u(t)\| < \infty.$$

Definition 2.16. (i) Let $u, v \in L^1(I;X)$. We say v is the **weak derivative** of u, written $u' = v$, provided

$$\int_I \phi'(t)u(t)\,dt = -\int_I \phi(t)v(t)\,dt$$

holds in X for all scalar test functions $\phi \in C_c^\infty(I)$.

(ii) The Sobolev space $W^{1,p}(I;X)$ consists of all functions $u \in L^p(I;X)$ such that weak derivative u' exists and belongs to $L^p(I;X)$. The norm is defined by

$$\|u\|_{W^{1,p}(I;X)} = \begin{cases} \left(\int_I (\|u(t)\|^p + \|u'(t)\|^p)dt\right)^{1/p} & (1 \le p < \infty), \\ \operatorname{esssup}_{t\in I}(\|u(t)\| + \|u'(t)\|) & (p = \infty). \end{cases}$$

We write $H^1(I;X) = W^{1,2}(I;X)$.

Theorem 2.42. (Calculus in $W^{1,p}(I;X)$) *Let $u \in W^{1,p}(I;X)$ for some $1 \le p \le \infty$. Then*

(i) $u \in C(\bar{I};X)$ (after being redefined on a null set of time), with

$$\|u\|_{C(\bar{I};X)} \le C\|u\|_{W^{1,p}(I;X)}$$

for a constant C depending on I.

(ii) $u(t) = u(s) + \int_s^t u'(\tau)\,d\tau$ for all $s \le t$ in I.

Proof. Extend u outside of I by 0 on $t \in \mathbb{R}$, and then set $u^\varepsilon = \omega_\varepsilon \star u$, ω_ε denoting the usual mollifier on $\mathbb{R}$. We have $(u^\varepsilon)' = \omega_\varepsilon \star u'$ on $I_\varepsilon := (a+\varepsilon, b-\varepsilon)$ if $I = (a,b)$. Then the proof can be completed by standard approximation method upon $\varepsilon \to 0$. ☐

Theorem 2.43. (More calculus) *Suppose $u \in L^2(I;H_0^1(\Omega))$, with $u' \in L^2(I;H^{-1}(\Omega))$. Then*

(i) $u \in C(\bar{I};L^2(\Omega))$ (after being redefined on a null set of time), with

$$\|u\|_{C(\bar{I};L^2(\Omega))} \le C(\|u\|_{L^2(I;H_0^1(\Omega))} + \|u'\|_{L^2(I;H^{-1}(\Omega))})$$

for a constant C depending on I.

(ii) The mapping $t \mapsto \|u(t)\|^2_{L^2(\Omega)}$ is absolutely continuous on $\bar{I}$, with

$$\frac{d}{dt}\|u(t)\|^2_{L^2(\Omega)} = 2\langle u'(t), u(t)\rangle$$

for a.e. $t \in I$, where "$\langle\cdot\rangle$" is the pairing in $H^{-1}(\Omega) \times H_0^1(\Omega)$.

For use later in the regularity study, we will need an extension of Theorem 2.43.

Theorem 2.44. (Mapping into better spaces) *Let Ω be a bounded domain with smooth $\partial\Omega$ and m a nonnegative integer. Suppose $u \in L^2(I;H^{m+2}(\Omega))$, with $u' \in L^2(I;H^m(\Omega))$. Then $u \in C(\bar{I};H^{m+1}(\Omega))$ (after being redefined on a null set of time), with*

$$\|u\|_{C(\bar{I};H^{m+1}(\Omega))} \le C(\|u\|_{L^2(I;H^{m+2}(\Omega))} + \|u'\|_{L^2(I;H^m(\Omega))})$$

for a constant C depending on I, Ω and m.

Remark 2.17. In most of the study of evolution equations later, the interval $I = (0,T)$. In this case, we write $L^p(I;X)$ as $L^p(0,T;X)$; other spaces are to be denoted similarly.

Second-Order Linear Elliptic Equations

3.1. Differential Equations in Divergence Form

Henceforth, $\Omega \subset \mathbb{R}^n$ denotes a bounded domain with boundary $\partial\Omega \in C^1$.

3.1.1. Linear Elliptic Equations. We study 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 $H^1_0(\Omega)$) and L is a **second-order differential operator** having either the **divergence form**

$$Lu \equiv -\sum_{i,j=1}^{n} D_i\left(a_{ij}(x)D_j u\right) + \sum_{i=1}^{n} b_i(x)D_i u + c(x)u \tag{3.2}$$

or else

$$Lu \equiv -\sum_{i,j=1}^{n} a_{ij}(x)D_{ij}u + \sum_{i=1}^{n} b_i(x)D_i u + c(x)u$$

with given real coefficients $a_{ij}(x),\ b_i(x)$ and $c(x)$. We also assume

$$a_{ij}(x) = a_{ji}(x) \quad (i, j = 1, \ldots, n).$$

Definition 3.1. The partial differential operator L is said to be **uniformly elliptic** in Ω if there exists a number $\theta > 0$ such that for every $x \in \Omega$ and every real vector $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$

$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq \theta \sum_{i=1}^{n} |\xi_i|^2. \tag{3.3}$$

We will assume $a_{ij}, b_i, c \in L^\infty(\Omega)$. Define the **bilinear form**

$$B_1[u,v] \equiv \int_\Omega \big[\sum_{i,j=1}^{n} a_{ij}D_j u D_i v + (\sum_{i=1}^{n} b_i D_i u + cu)v\big]\,dx$$

Definition 3.2. Let $f \in L^2(\Omega)$. A function $u \in H_0^1(\Omega)$ is called a **weak solution** of (3.1) with L given by (3.2) if $B_1[u,v] = (f,v)_{L^2}$ for all $v \in H_0^1(\Omega)$; that is, the following holds

$$\int_\Omega \big[\sum_{i,j=1}^n a_{ij} D_j u D_i v + (\sum_{i=1}^n b_i D_i u + cu) v\big] dx = \int_\Omega f v dx \quad \forall\, v \in H_0^1(\Omega). \tag{3.4}$$

If $f \in H^{-1}(\Omega) = (H_0^1(\Omega))^*$, the dual space of $H_0^1(\Omega)$, then weak solutions are defined by replacing the right-hand side by $\langle f, u\rangle$, $\langle , \rangle$ being the dual pairing on $H^{-1}(\Omega) \times H_0^1(\Omega)$.

Exercise 3.3. 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 Systems in Divergence Form.

For N unknown functions, $u^1, \cdots, u^N$, we can write $u = (u^1, \cdots, u^N)$ and 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 learned. If $u \in W^{1,p}(\Omega;\mathbb{R}^N)$ then we use Du to denote the $N \times n$ Jacobi matrix

$$Du = (\partial u^k / \partial x_i)_{1\le k\le N, 1\le i\le n}.$$

The (Dirichlet) BVP for a most general **system of second-order (quasilinear) partial differential equations in divergence form** can be written as follows:

$$-\operatorname{div} A(x,u,Du) + b(x,u,Du) = F \ \text{ in } \Omega, \quad u = 0 \ \text{ on } \partial\Omega, \tag{3.5}$$

where $A(x,s,\xi) = (A_i^k(x,u,\xi))$, $1 \le i \le n$, $1 \le k \le N$, and $b(x,s,\xi) = (b^k(x,u,\xi))$, $1 \le k \le N$, are given functions of $(x,u,\xi) \in \Omega \times \mathbb{R}^N \times \mathbb{M}^{N\times n}$, 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 coefficients A, b usually satisfy certain **structural conditions** that 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 $p' = \frac{p}{p-1}$. In such cases, a function $u \in W_0^{1,p}(\Omega;\mathbb{R}^N)$ is called a **weak solution** of (3.5) if the following holds

$$\int_\Omega \left[\sum_{i=1}^n A_i^k(x,u,Du) D_i\varphi + b^k(x,u,Du)\varphi\right] dx = \langle f^k, \varphi\rangle \tag{3.6}$$

for all $\varphi \in W_0^{1,p}(\Omega)$ and each $k = 1, 2, \cdots, N$.

Definition 3.4. The system (3.5) is said to be **linear** if both A and b are linear in the variables (u,ξ); that is,

$$\begin{aligned} A_i^k(x,u,Du) &= \sum_{1\le l\le N,\, 1\le j\le n} a_{ij}^{kl}(x)\, D_j u^l + \sum_{l=1}^N d_i^{kl}(x)\, u^l, \\ b^k(x,u,Du) &= \sum_{1\le j\le n,\, 1\le l\le N} b_j^{kl}(x)\, D_j u^l + \sum_{l=1}^N c^{kl}(x)\, u^l. \end{aligned} \tag{3.7}$$

For linear systems, the suitable space is Hilbert space $H_0^1(\Omega;\mathbb{R}^N)$ equipped with 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

$$\langle F, u\rangle = \sum_{k=1}^N \langle f^k, u^k\rangle \quad \text{if } F = (f^k) \text{ and } u = (u^k).$$

The bilinear form in this case is defined by

$$B_2[u,v] \equiv \int_\Omega \Big(a_{ij}^{kl} D_j u^l D_i v^k + d_i^{kl} u^l D_i v^k + b_j^{kl} D_j u^l v^k + c^{kl} u^l v^k \Big)\, dx;$$

here the conventional summation notation is used. With this $B_2[u,v]$, a weak solution to the Dirichlet problem of linear system (3.5) is then a function $u \in H_0^1(\Omega;\mathbb{R}^N)$ such that

$$B_2[u,v] = \langle F, v\rangle \quad \forall\, v \in H_0^1(\Omega;\mathbb{R}^N). \tag{3.8}$$

Ellipticity Conditions. There are several *ellipticity* conditions for the system (3.5) in terms of the leading coefficients $A(x,u,\xi)$. Assume A is smooth on ξ and define

$$A_{ij}^{kl}(x,u,\xi) = \frac{\partial A_i^k(x,u,\xi)}{\partial \xi_j^l}, \quad \xi = (\xi_j^l).$$

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

$$\sum_{i,j=1}^n \sum_{k,l=1}^N A_{ij}^{kl}(x,s,\xi)\, \eta_i^k\, \eta_j^l \ge \nu\, |\eta|^2 \quad \text{for all } N\times n \text{ matrix } \eta = (\eta_i^k). \tag{3.9}$$

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 (uniform) **Legendre-Hadamard condition**:

$$\sum_{i,j=1}^n \sum_{k,l=1}^N A_{ij}^{kl}(x,s,\xi)\, q^k q^l p_i p_j \ge \nu\, |p|^2\, |q|^2 \quad \forall\, p\in\mathbb{R}^n,\; q\in\mathbb{R}^N. \tag{3.10}$$

For systems with *linear leading terms* A given by (3.7), the Legendre condition and Legendre-Hadamard condition become, respectively,

$$\sum_{i,j=1}^n \sum_{k,l=1}^N a_{ij}^{kl}(x)\, \eta_i^k\, \eta_j^l \ge \nu\, |\eta|^2 \quad \forall\, \eta; \tag{3.11}$$

$$\sum_{i,j=1}^n \sum_{k,l=1}^N a_{ij}^{kl}(x)\, q^k q^l p_i p_j \ge \nu\, |p|^2\, |q|^2 \quad \forall\, p,\; q. \tag{3.12}$$

Exercise 3.5. If $N > 1$, 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

$$\sum_{i,j,k,l=1}^2 a_{ij}^{kl}\, \xi_i^k\, \xi_j^l \equiv \det \xi + \varepsilon\, |\xi|^2.$$

Show that the Legendre-Hadamard condition holds for all $\varepsilon > 0$. But, the Legendre condition holds for this system if and only if $\varepsilon > 1/2$.

Exercise 3.6. Let $u = (v, w)$ and $x = (x_1, x_2) = (x, y) \in \mathbb{R}^2$. 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 $\Delta f = f_{xx} + f_{yy}$):

$$\epsilon^2\Delta^2 v - v_{xxyy} = 0, \quad \epsilon^2\Delta^2 w + w_{xxyy} = 0.$$

We can easily see that both equations are **elliptic** if and only if $\epsilon > 1/2$.

Exercise 3.7. Formulate the *biharmonic equation* $\Delta^2 u = f$ as a linear system and find the appropriate bilinear form $B[u, v]$ in the definition of weak solutions.

3.2. The Lax-Milgram Theorem

Let H denote a real Hilbert space with inner product $(\cdot, \cdot)$ and norm $\|\cdot\|$.

A map $B : H \times H \to \mathbb{R}$ 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 3.1. (Lax-Milgram Theorem) *Let $B\colon H \to H$ be a bilinear form. Assume*

(i) B is **bounded***; i.e., $|B[u, v]| \leq \alpha\|u\|\|v\|$, and*

(ii) B is **strongly positive***; i.e., $B[u, u] \geq \beta\|u\|^2$,*

where α, β are positive constants. Let $f \in H^*$. Then there exists a unique element $u \in H$ such that

$$B[u, v] = \langle f, v\rangle, \quad \forall\ v \in H. \tag{3.13}$$

Moreover, the solution u satisfies $\|u\| \leq \frac{1}{\beta}\|f\|$.

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

$$B[u, v] = (w, v) \quad \forall\ v \in H.$$

It can be easily shown that $A : H \to H$ is linear. From (i), $\|Au\|^2 = B[u, Au] \leq \alpha\|u\|\|Au\|$, and hence $\|Au\| \leq \alpha\|u\|$ for all $u \in H$; that is, A is bounded. Furthermore, by (ii), $\beta\|u\|^2 \leq B[u, u] = (Au, u) \leq \|Au\|\|u\|$ and hence $\|Au\| \geq \beta\|u\|$ 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. There are many different proofs for this, and 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 will show 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 $u \in H$

$$\begin{aligned} \|(I - tA)u\|^2 &= \|u\|^2 + t^2\|Au\|^2 - 2t(Au, u) \\ &\leq \|u\|^2(1 + t^2\alpha^2 - 2\beta t). \end{aligned}$$

We now choose t such that $0 < t < \frac{2\beta}{\alpha^2}$. Then the expression in parentheses is positive and less than 1. Thus the map $T\colon H \to H$ is a contraction on H and therefore has a fixed point. This fixed point u solves $Au = w_0$ and thus is the unique solution of (3.13); moreover, we have $\|f\| = \|w_0\| = \|Au\| \geq \beta\|u\|$ and hence $\|u\| \leq \frac{1}{\beta}\|f\|$. The proof is complete. □

3.3. Energy Estimates and Existence Theory

We study the bilinear forms B_1, B_2 defined above. In the following, we assume all coefficients involved in the problems are in $L^\infty(\Omega)$. One can easily show the boundedness:

$$|B_j[u,v]| \leq \alpha\|u\|\|v\|$$

for all u, v in the respective Hilbert spaces $H = H_0^1(\Omega)$ or $H = H_0^1(\Omega; \mathbb{R}^N)$ for $j = 1, 2$.

The strong positivity (also called **coercivity**) for both B_1 and B_2 is not always guaranteed and involves estimating on the quadratic form $B_j[u,u]$, usually called **Gårding's estimates**. We will derive these estimates for both of them and state the corresponding existence theorems below.

3.3.1. Gårding's estimate for $B_1[u,u]$.

Theorem 3.2. *Assume the ellipticity condition* (3.3) *holds. Then, there are constants* $\beta > 0$ *and* $\gamma \geq 0$ *such that*

$$B_1[u,u] \geq \beta\|u\|^2 - \gamma\|u\|^2_{L^2(\Omega)} \quad \forall\, u \in H = H_0^1(\Omega). \tag{3.14}$$

Proof. Note that, by the ellipticity,

$$B_1[u,u] - \int_\Omega \Big(\sum_{i=1}^n b_i D_i u + cu\Big) u\, dx \geq \theta \int_\Omega \sum_{i=1}^n |D_i u|^2 dx.$$

Let $m = \max\{\|b_i\|_{L^\infty(\Omega)} \,|\, 1 \leq i \leq n\}$ and $k_0 = \|c\|_{L^\infty(\Omega)}$. Then

$$\begin{aligned} |(b_i D_i u, u)_2| &\leq m\|D_i u\|_2 \|u\|_2 \\ &\leq (m/2)(\varepsilon\|D_i u\|_2^2 + (1/\varepsilon)\|u\|_2^2) \end{aligned}$$

where in the last step we used the arithmetic-geometric inequality $|\alpha\beta| \leq (\varepsilon/2)\alpha^2 + (1/2\varepsilon)\beta^2$. Combining the estimates we find

$$B_1[u,u] \geq (\theta - m\varepsilon/2)\|Du\|^2_{L^2(\Omega)} - (k_0 + mn/2\varepsilon)\|u\|^2_{L^2(\Omega)}.$$

By choosing $\varepsilon > 0$ so that $\theta - m\varepsilon/2 > 0$ we arrive at the desired inequality, using the Poincare inequality: $\|u\|_{H^1(\Omega)} \leq C\|Du\|_{L^2(\Omega)}$ for all $u \in H_0^1(\Omega)$. □

Theorem 3.3. (First Existence Theorem for weak solutions) *There is a number* $\gamma \geq 0$ *such that for each* $\lambda \geq \gamma$ *and for each function* $f \in L^2(\Omega)$*, the boundary value problem*

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

has a unique weak solution $u \in H = H_0^1(\Omega)$ *which satisfies*

$$\|u\|_H \leq c\|f\|_{L^2(\Omega)}$$

where the positive constant c is independent of f*. Then result also holds for all* $f \in H^{-1}(\Omega)$*, with* $\|f\|_{L^2(\Omega)}$ *replaced by* $\|f\|_{H^{-1}(\Omega)}$*.*

Proof. Take γ from (3.14), let $\lambda \geq \gamma$ and 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. □

Example 3.4. Consider the **Neumann boundary** value problem

$$-\Delta u(x) = f(x) \ \text{ in } \Omega, \quad \frac{\partial u}{\partial \nu} = 0 \ \text{ on } \partial\Omega. \tag{3.15}$$

A function $u \in H^1(\Omega)$ is said to be a **weak solution** to (3.15 if

$$\int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega fv \, dx, \quad \forall \ v \in H^1(\Omega). \tag{3.16}$$

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

$$H = \left\{ u \in H^1(\Omega) \mid \int_\Omega u(x)\, dx = 0 \right\}$$

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

$$(u,w)_H = (f,w)_{L^2(\Omega)}, \quad \forall \ w \in H. \tag{3.17}$$

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 (3.17) and obtain (3.16) using $\int_\Omega f dx = 0$.

Example 3.5. Let us consider the nonhomogeneous Dirichlet boundary value problem

$$-\Delta u = f \ \text{ in } \ \Omega, \quad u|_{\partial\Omega} = \varphi \tag{3.18}$$

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 (3.18) 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 (3.18) and set $u = z + w$. Then $z \in H_0^1(\Omega)$ satisfies

$$\int_\Omega \nabla z \cdot \nabla v dx = \int_\Omega (fv - \nabla v \cdot \nabla w) dx \quad \text{for all} \quad v \in H_0^1(\Omega). \tag{3.19}$$

Since the right hand side belongs to the dual space $H^{-1}(\Omega) = H_0^1(\Omega)^*$, the Lax-Milgram theorem yields the existence of a unique $z \in H_0^1(\Omega)$ which satisfies (3.19). Hence (3.18) has a unique weak solution u.

EXAMPLE 3.6. 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, it can be shown that $\|\Delta u\|_2$ defines a norm on $H_0^2(\Omega)$ which is equivalent to the usual norm on $H^2(\Omega)$. (**Exercise!**) Hence

$$B[u,u] = \|\Delta u\|_2^2 \ge c\|u\|_{2,2}^2$$

and so, by the Lax-Milgram theorem (in fact, just the Riesz Representation Theorem), there exists a unique weak solution $u \in H$.

Exercise 3.8. Denote by H_c^1 the space

$$H_c^1 = \{u \in H^1(\Omega) : \gamma_0 u = \text{const}\}.$$

Note that the constant may be different for different u's.

(a) Prove that H_c^1 is complete.

(b) Let $f \in C(\bar\Omega)$. Prove existence of a unique $u \in H_c^1$ satisfying

$$\int_\Omega (\nabla u \cdot \nabla v + uv) dx = \int_\Omega f v dx \quad \forall\, v \in H_c^1.$$

(c) If $u \in C^2(\bar\Omega)$ satisfies the equation in (b), find the underlying BVP.

Exercise 3.9. Let $\Omega = (1, +\infty)$. Show that the BVP $-u'' = f \in L^2(\Omega),\ u \in H_0^1(\Omega)$ does not have a weak solution.

3.3.2. Gårding's estimate for $B_2[u,u]$.

We will derive the Gårding estimate for $B_2[u,u]$. For simplicity, let $H = H_0^1(\Omega;\mathbb{R}^N)$ and let $(u,v)_H$ and $\|u\|_H$ be the equivalent inner product and norm defined above on H. Define the bilinear form of the leading terms by

$$A[u,v] = \sum_{i,j=1}^n \sum_{k,l=1}^N \int_\Omega a_{ij}^{kl}(x)\, D_j u^l\, D_i v^k\, dx.$$

Theorem 3.7. *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. In the first case, the conclusion follows easily from the Legendre condition. We prove the second case when A_{ij}^{kl} are constants satisfying the Legendre-Hadamard condition

$$\sum_{i,j=1}^{n}\sum_{k,l=1}^{N} a_{ij}^{kl}\, q^k\, q^l\, p_i p_j \geq \nu\, |p|^2 |q|^2, \quad \forall p \in \mathbb{R}^n,\ q \in \mathbb{R}^N.$$

We prove

$$A[u,u] = \sum_{i,j=1}^{n}\sum_{k,l=1}^{N} \int_\Omega a_{ij}^{kl}\, D_j u^l\, D_i u^k\, dx \geq \nu \int_\Omega |Du|^2\, dx$$

for all $u \in C_0^\infty(\Omega; \mathbb{R}^N)$. For these test functions u, we extend them onto $\mathbb{R}^n$ by zero outside Ω and thus consider them as functions in $C_0^\infty(\mathbb{R}^n; \mathbb{R}^N)$. Define the **Fourier transforms** for such functions 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 $u,\, v \in C_0^\infty(\mathbb{R}^n; \mathbb{R}^N)$,

$$\int_{\mathbb{R}^n} u(x)\cdot v(x)\, dx = \int_{\mathbb{R}^n} \hat{u}(y)\cdot \overline{\hat{v}(y)}\, dy,$$

$$\widehat{D_j u^k}(y) = i\, y_j \widehat{u^k}(y);$$

the last 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,

$$\mathrm{Re} \sum_{i,j=1}^{n}\sum_{k,l=1}^{N} \left(a_{ij}^{kl}\, y_i\, y_j\, \widehat{u^k}(y)\, \overline{\widehat{u^l}(y)}\right) \geq \nu\, |y|^2\, (|\eta|^2 + |\xi|^2) = \nu\, |y|^2\, |\hat{u}(y)|^2.$$

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 3.8. (Gårding's estimate for system) *Let $B_2[u,v]$ be defined by* (3.8). *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. 1. By uniform continuity, we can choose a small $\epsilon > 0$ such that

$$|a_{ij}^{kl}(x) - a_{ij}^{kl}(y)| \leq \frac{\nu}{2}, \quad \forall\, x,\, y \in \bar{\Omega},\ |x-y| \leq \epsilon.$$

We claim

$$\int_\Omega a_{ij}^{kl}(x)\, D_i u^k\, D_j u^l\, dx \geq \frac{\nu}{2} \int_\Omega |Du(x)|^2\, dx = \frac{\nu}{2} \sum_{\substack{1\leq i\leq n \\ 1\leq k\leq N}} \int_\Omega |D_i u^k(x)|^2\, dx \tag{3.20}$$

for all test functions $u \in C_0^\infty(\Omega;\mathbb{R}^N)$ with $\mathrm{diam}(\mathrm{supp}\, u) \leq \epsilon$. To see this, we choose any point $x_0 \in \mathrm{supp}\, u$. Then

$$\begin{aligned}\int_\Omega a_{ij}^{kl}(x)\, D_i u^k\, D_j u^l\, dx &= \int_\Omega a_{ij}^{kl}(x_0)\, D_i u^k\, D_j u^l\, dx \\ &\quad + \int_{\mathrm{supp}\, u} \big(a_{ij}^{kl}(x) - a_{ij}^{kl}(x_0)\big)\, D_i u^k\, D_j u^l\, dx \\ &\geq \nu \int_\Omega |Du(x)|^2\, dx - \frac{\nu}{2} \int_\Omega |Du(x)|^2\, dx,\end{aligned}$$

which proves (3.20).

2. Now assume $u \in C_0^\infty(\Omega;\mathbb{R}^N)$, with arbitrary compact support. We 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

$$\varphi_m(x) = \frac{\zeta_m(x)}{\left(\sum_{j=1}^M \zeta_j^2(x)\right)^{1/2}}, \quad m = 1, 2, ..., M.$$

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

$$\begin{aligned}a_{ij}^{kl}(x)\, D_i u^k\, D_j u^l &= \sum_{m=1}^M \Big(a_{ij}^{kl}(x)\, \varphi_m^2\, D_i u^k\, D_j u^l\Big) \\ &= \sum_{m=1}^M a_{ij}^{kl}(x)\, D_i(\varphi_m\, u^k)\, D_j(\varphi_m\, u^l) \qquad (3.21) \\ &\quad - \sum_{m=1}^M a_{ij}^{kl}(x)\, \Big(\varphi_m\, u^l\, D_i\varphi_m\, D_i u^k + \varphi_m\, u^k\, D_i\varphi_m\, D_j u^l + u^k\, u^l\, D_i\varphi_m\, D_j\varphi_m\Big).\end{aligned}$$

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 (3.20)

$$\begin{aligned}
\int_\Omega a_{ij}^{kl}(x)\, D_i(\varphi_m\, u^k) D_j(\varphi_m\, u^l)\, dx &\ge \frac{\nu}{2} \sum_{\substack{1\le i\le n\\ 1\le k\le N}} \int_\Omega |D_i(\varphi_m\, u^k)|^2\, dx \\
&= \frac{\nu}{2} \sum_{\substack{1\le i\le n\\ 1\le k\le N}} \int_\Omega \Big(\varphi_m^2\, |D_i u^k|^2 + |D_i\varphi_m|^2\, |u^k|^2 + 2\varphi_m\, u^k\, D_i\varphi_m\, D_i u^k\Big) dx \\
&\ge \frac{\nu}{2} \sum_{\substack{1\le i\le n\\ 1\le k\le N}} \int_\Omega \Big(\varphi_m^2\, |D_i u^k|^2\, dx + 2\varphi_m\, u^k\, D_i\varphi_m\, D_i u^k\Big)\, dx \\
&\ge \frac{\nu}{4} \sum_{\substack{1\le i\le n\\ 1\le k\le N}} \int_\Omega \varphi_m^2\, |D_i u^k|^2\, dx - C\|u\|^2_{L^2(\Omega)} = \frac{\nu}{4}\int_\Omega \varphi_m^2\, |Du|^2\, dx - C\|u\|^2_{L^2(\Omega)},
\end{aligned}$$

where we have used the Cauchy inequality with ϵ. Then by (3.21) and the fact that $\sum_{m=1}^M \varphi_m^2 = 1$ on Ω,

$$\int_\Omega a_{ij}^{kl}(x)\, D_i u^k\, D_j u^l\, dx \ge \frac{\nu}{4}\int_\Omega |Du|^2\, dx - CM\, \|u\|^2_{L^2(\Omega)} - C_1\, \|u\|_{L^2(\Omega)}\, \|Du\|_{L^2(\Omega)}.$$

The terms in $B_2[u,u]$ involving b, c and d can all be estimated by

$$C_2(\|u\|_{L^2(\Omega)}\, \|Du\|_{L^2(\Omega)} + \|u\|^2_{L^2(\Omega)}).$$

Finally, using the Cauchy inequality with ϵ again, we have

$$B_2[u,u] \ge \frac{\nu}{8}\|u\|^2_{H_0^1(\Omega)} - C_3\|u\|^2_{L^2(\Omega)} \quad \forall\, u \in C_0^\infty(\Omega; \mathbb{R}^N)$$

and, by density, 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 3.9. *Under the hypotheses of the previous theorem, for $\lambda \ge \lambda_1$, the Dirichlet problem for the system* (3.5) *with linear coefficients* (3.7) *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 coefficients.*

Corollary 3.10. *Given $\lambda \ge \lambda_1$ as in the theorem, then 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 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 $H_0^1(\Omega; \mathbb{R}^N)$, which, by the compact embedding theorem, is compactly embedded in $L^2(\Omega; \mathbb{R}^N)$. Hence, as a linear operator from $L^2(\Omega; \mathbb{R}^N)$ to $L^2(\Omega; \mathbb{R}^N)$, $\mathcal{K}$ is compact. □

3.4. Fredholm Alternatives

We study the general linear system Lu whose bilinear form is given by $B_2[u,v]$ defined above on $H = H_0^1(\Omega; \mathbb{R}^N)$. We need some necessary results on the spectral theory of compact linear operators given in §1.4 of Chapter 1.

Definition 3.10. The **adjoint bilinear form** B_2^* of B_2 is defined by

$$B_2^*[u,v] = B_2[v,u] \quad \forall\, u,\, v \in H = H_0^1(\Omega;\mathbb{R}^N).$$

This bilinear form $B_2^*[u,v]$ is associated to the **formal adjoint** of Lu of the form

$$L^*u = -\operatorname{div} A^*(x,u,Du) + b^*(x,u,Du), \tag{3.22}$$

with linear coefficients $A^*(x,u,Du) = (\tilde{A}_{ij}^{kl})$ and $b^*(x,u,Du) = (\tilde{b}^k)$ given by

$$\begin{aligned} \tilde{A}_i^k(x,u,Du) &= \sum_{1\le l\le N,\, 1\le j\le n} \tilde{a}_{ij}^{kl}(x)\, D_j u^l + \sum_{l=1}^N \tilde{d}_i^{kl}(x)\, u^l, \\ \tilde{b}^k(x,u,Du) &= \sum_{1\le j\le n,\, 1\le l\le N} \tilde{b}_j^{kl}(x)\, D_j u^l + \sum_{l=1}^N \tilde{c}^{kl}(x)\, u^l \end{aligned} \tag{3.23}$$

where

$$\tilde{a}_{ij}^{kl} = a_{ji}^{lk}, \quad \tilde{d}_i^{kl} = b_i^{lk}, \quad \tilde{b}_j^{kl} = d_i^{lk}, \quad \tilde{c}^{kl} = c^{lk} \quad (1 \le i,j \le n,\ 1 \le k,l \le N).$$

The ellipticity condition of L^*u is the same as that of Lu, and also $B_2^*[u,u] = B_2[u,u]$.

Theorem 3.11. (Second Existence Theorem for weak solutions) *Assume the ellipticity and boundedness of the coefficients of Lu.*

(i) *Precisely one of the following statements holds:*
either

$$\begin{cases} \text{for each } F \in L^2(\Omega;\mathbb{R}^N) \text{ there exists a unique} \\ \text{weak solution } u \in H_0^1(\Omega;\mathbb{R}^N) \text{ of } Lu = F, \end{cases} \tag{3.24}$$

or else

$$\text{there exists a weak solution } u \neq 0 \text{ in } H_0^1(\Omega;\mathbb{R}^N) \text{ of } Lu = 0. \tag{3.25}$$

(ii) *Furthermore, should case* (3.25) *hold, the dimension of the subspace $\mathcal{N} \subset H_0^1(\Omega;\mathbb{R}^N)$ of weak solutions of $Lu = 0$ is finite and equals the dimension of the subspace $\mathcal{N}^* \subset H_0^1(\Omega;\mathbb{R}^N)$ of weak solutions of adjoint problem $L^*u = 0$.*

(iii) *Finally, the problem $Lu = F$ has a weak solution if and only if*

$$(F,v)_{L^2(\Omega;\mathbb{R}^N)} = 0 \quad \forall\, v \in \mathcal{N}^*.$$

The dichotomy (3.24), (3.25) *is called the* **Fredholm alternatives**.

Proof. We assume the Gårding inequality holds (see Theorem 3.8 for sufficient conditions):

$$B_2[u,u] \ge \sigma \|u\|_{H_0^1}^2 - \mu \|u\|_{L^2}^2, \quad \forall\, u \in H_0^1(\Omega;\mathbb{R}^N), \tag{3.26}$$

where $\sigma > 0$ and $\mu \in \mathbb{R}$ are constants. We also assume $\mu > 0$. For each $F \in L^2(\Omega;\mathbb{R}^N)$, define $u = \mathcal{K}F$ to be the unique weak solution in $H_0^1(\Omega;\mathbb{R}^N)$ of the BVP

$$Lu + \mu u = F \quad \text{in } \Omega, \quad u|_{\partial\Omega} = 0.$$

By Theorem 3.9 and Corollary 3.10, this $\mathcal{K}$ is well defined and is a **compact linear operator** on $L^2(\Omega;\mathbb{R}^N)$. 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 $H_0^1(\Omega;\mathbb{R}^N)$ into $L^2(\Omega;\mathbb{R}^N)$. Furthermore, given $F \in L^2(\Omega;\mathbb{R}^N)$, $u \in H_0^1(\Omega;\mathbb{R}^N)$ is a weak solution of $Lu = F$ if and only if $Lu + \mu u = F + \mu u$, which is equivalent to the equation $u = \mathcal{K}(F + \mu u) = \mathcal{K}F + \mu\mathcal{K}u$; that is, $(I - \mu\mathcal{K})u = \mathcal{K}F$. Hence, we have $\mathcal{N} = \mathcal{N}(I - \mu\mathcal{K})$ and similarly, $\mathcal{N}^* = \mathcal{N}(I - \mu\mathcal{K}^*)$; moreover, $Lu = F$ if and only if $\mathcal{K}F \in \mathcal{R}(I-\mu\mathcal{K}) = (\mathcal{N}(I-\mu\mathcal{K}^*))^\perp = (\mathcal{N}^*)^\perp$. The proof of (iii) follows as, for all $v \in \mathcal{N}^*$, $v = \mu\mathcal{K}^* v$ and so

$$(F, v) = (F, \mu\mathcal{K}^* v) = \mu(\mathcal{K}F, v);$$

hence $\mathcal{K}F \in (\mathcal{N}^*)^\perp$ if and only if $F \in (\mathcal{N}^*)^\perp$. □

3.4.1. Symmetric Elliptic Operators. In what follows, we assume Ω is a bounded domain. We consider the operator

$$Lu = -\operatorname{div} A(x, Du) + c(x)u, \quad u \in H_0^1(\Omega; \mathbb{R}^N),$$

where $A(x, Du)$ is a linear system defined with $A(x,\xi)$, $\xi \in \mathbb{M}^{N\times n}$, given by

$$A_i^k(x,\xi) = \sum_{1\le l\le N,\, 1\le j\le n} a_{ij}^{kl}(x)\xi_j^l.$$

Here $a_{ij}^{kl}(x)$ and $c(x)$ are given functions in $L^\infty(\Omega)$. The bilinear form associated to L is

$$B[u,v] = \int_\Omega \left(\sum_{k,l=1}^N \sum_{i,j=1}^n a_{ij}^{kl}(x) D_j u^l D_i v^k + c(x) u\cdot v \right) dx, \quad u,\, v \in H_0^1(\Omega;\mathbb{R}^N). \tag{3.27}$$

We assume that B is *symmetric* on $H_0^1(\Omega;\mathbb{R}^N)$, that is,

$$B[u,v] = B[v,u] \quad \forall\, u, v \in H_0^1(\Omega;\mathbb{R}^N).$$

In this case $B^* = B$ and Lu is **self-adjoint**: $L^*u = Lu$. This condition is equivalent to the following symmetry condition:

$$a_{ij}^{kl}(x) = a_{ji}^{lk}(x), \quad \forall\, i,\, j = 1,2,\cdots,n;\;\; k,\, l = 1,2,\cdots,N. \tag{3.28}$$

We also assume the Gårding inequality holds (see Theorem 3.8 for sufficient conditions):

$$B[u,u] \ge \sigma\|u\|_{H_0^1}^2 - \mu\|u\|_{L^2}^2, \quad \forall\, u \in H_0^1(\Omega;\mathbb{R}^N), \tag{3.29}$$

where $\sigma > 0$ and $\mu \in \mathbb{R}$ are constants.

For each $F \in L^2(\Omega;\mathbb{R}^N)$, define $u = \mathcal{K}F$ to be the unique weak solution in $H_0^1(\Omega;\mathbb{R}^N)$ of the BVP

$$Lu + \mu u = F \quad \text{in } \Omega, \quad u|_{\partial\Omega} = 0.$$

By Theorem 3.9 and Corollary 3.10, 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 $H_0^1(\Omega;\mathbb{R}^N)$ into $L^2(\Omega;\mathbb{R}^N)$.

Theorem 3.12. $\mathcal{K}\colon L^2(\Omega;\mathbb{R}^N) \to L^2(\Omega;\mathbb{R}^N)$ *is symmetric and positive; that is,*

$$(\mathcal{K}F, G)_{L^2} = (\mathcal{K}G, F)_{L^2}, \quad (\mathcal{K}F, F)_{L^2} \ge 0, \quad \forall\, F,\, G \in L^2(\Omega;\mathbb{R}^N).$$

Furthermore, given $\lambda \in \mathbb{R}$ and $F \in L^2(\Omega;\mathbb{R}^N)$, $u \in H_0^1(\Omega;\mathbb{R}^N)$ 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 (3.29),

$$(\mathcal{K}F, F)_{L^2} = (u, F)_{L^2} = B[u,u] + \mu\|u\|_{L^2}^2 \ge \sigma\|u\|_{H_0^1}^2 = \sigma\|\mathcal{K}F\|_{H_0^1}^2 \ge 0. \tag{3.30}$$

Finally, $u \in H_0^1(\Omega;\mathbb{R}^N)$ 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, $[I - (\lambda+\mu)\mathcal{K}]u = \mathcal{K}F$. □

3.4.2. 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.

Theorem 3.13. (Eigenvalue Theorem) *Assume* (3.28) *and* (3.29). *Then the eigenvalues of L consist of a countable set $\Sigma = \{\lambda_k\}_{k=1}^{\infty}$, where*

$$-\mu < \lambda_1 \le \lambda_2 \le \lambda_3 \le \cdots$$

are listed repeatedly the same times as the multiplicity, and

$$\lim_{k\to\infty} \lambda_k = \infty.$$

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 $L^2(\Omega;\mathbb{R}^N)$.*

The first (smallest) eigenvalue λ_1, which is called the (Dirichlet) **principal eigenvalue** *of L, is characterized by*

$$\lambda_1 = \min_{\substack{u\in H_0^1(\Omega;\mathbb{R}^N)\\ \|u\|_{L^2(\Omega)}=1}} B[u,u]. \tag{3.31}$$

Moreover, if $u \in H_0^1(\Omega;\mathbb{R}^N)$, $u \not\equiv 0$, then u is an eigenfunction corresponding to λ_1 if and only if

$$B[u,u] = \lambda_1 \|u\|_{L^2(\Omega)}^2.$$

Proof. 1. From Theorem 3.12, 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 (3.30), $\mathcal{K}$ 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

$$-\mu < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_j \to \infty.$$

2. We now prove the second statement. If u is an eigenfunction corresponding to λ_1 with $\|u\|_{L^2(\Omega)} = 1$, then easily $B[u,u] = \lambda_1(u,u) = \lambda_1\|u\|_{L^2(\Omega)}^2 = \lambda_1$. We now assume

$$u \in H_0^1(\Omega), \quad \|u\|_{L^2(\Omega)} = 1.$$

Let $\{w_k\}$ be the orthonormal basis of $L^2(\Omega)$ consisting of eigenfunctions. Then

$$\begin{cases} B[w_k, w_l] = \lambda_k(w_k, w_l) = 0 \ \ (k \neq l), \\ B[w_k, w_k] = \lambda_k(w_k, w_k) = \lambda_k. \end{cases}$$

Set $\tilde{w}_k = (\lambda_k + \mu)^{-1/2} w_k$, and consider the inner product on $H = H_0^1(\Omega;\mathbb{R}^N)$ defined by

$$((u,v)) := B_\mu[u,v] = B[u,v] + \mu(u,v)_{L^2(\Omega)} \quad (u,v \in H).$$

Then $((\tilde{w}_k, \tilde{w}_l)) = \delta_{kl}$. Let $d_k = (u, w_k)_{L^2(\Omega)}$. We have

$$\sum_{k=1}^{\infty} d_k^2 = \|u\|^2_{L^2(\Omega)} = 1, \quad u = \sum_{k=1}^{\infty} d_k w_k = \sum_{k=1}^{\infty} \tilde{d}_k \tilde{w}_k, \tag{3.32}$$

where $\tilde{d}_k = d_k\sqrt{\lambda_k + \mu}$, in the sense of norm-convergence in $L^2(\Omega;\mathbb{R}^N)$. We claim that the series for u converges also in the norm defined by the inner product $((\ ,\ ))$ in H. To prove this claim, for $m = 1, 2, \cdots$, define

$$u_m = \sum_{k=1}^{m} d_k w_k = \sum_{k=1}^{m} \tilde{d}_k \tilde{w}_k \in H.$$

From $((\tilde{w}_k, u)) = B[\tilde{w}_k, u] + \mu(\tilde{w}_k, u) = (\lambda_k + \mu)(\tilde{w}_k, u) = \tilde{d}_k$, we have

$$((u_m, u)) = \sum_{k=1}^{m} \tilde{d}_k^2 = ((u_m, u_m)) \quad (m = 1, 2, \cdots).$$

This implies $((u_m, u_m)) \le ((u, u))$ for all $m = 1, 2, \cdots$. Hence, $\{u_m\}$ is bounded in H and so, by a subsequence, $u_m \rightharpoonup \tilde{u}$ in H as $m \to \infty$. Since $u_m \to u$ in L^2, we must have $\tilde{u} = u$ and so

$$((u, u)) \le \liminf_{m\to\infty} ((u_m, u_m)),$$

which, combined with $((u - u_m, u - u_m)) = ((u, u)) + ((u_m, u_m)) - 2((u, u_m)) = ((u, u)) - ((u_m, u_m))$, implies that $u_m \to u$ in H, and the claim is proved.

3. Now, by (3.32), we have

$$B[u, u] = \sum_{k=1}^{\infty} d_k B[w_k, u] = \sum_{k=1}^{\infty} d_k^2 \lambda_k \ge \sum_{k=1}^{\infty} d_k^2 \lambda_1 = \lambda_1.$$

Hence (3.31) is proved. Moreover, if in addition $B[u, u] = \lambda_1$, then we have

$$\sum_{k=1}^{\infty} (\lambda_k - \lambda_1) d_k^2 = 0; \quad \text{so } d_k = 0 \text{ if } \lambda_k > \lambda_1.$$

Assume λ_1 has multiplicity m, with $Lw_k = \lambda_1 w_k$ $(k = 1, 2, \cdots, m)$. Then $u = \sum_{k=1}^m d_k w_k$, and so $Lu = \lambda_1 u$; that is, u is an eigenfunction corresponding to λ_1. ☐

We consider a special case when $N = 1$ and the operator Lu is given by

$$Lu = -\sum_{i,j=1}^{n} D_j(a_{ij}(x) D_i u) + c(x)u,$$

where the uniform ellipticity condition is satisfied, $\partial\Omega$ is smooth, and a_{ij}, c are smooth functions satisfying

$$a_{ij}(x) = a_{ji}(x), \quad c(x) \ge 0 \quad (x \in \bar{\Omega}).$$

Theorem 3.14. *The principal eigenvalue $\lambda_1 > 0$. Let w_1 be an eigenfunction corresponding to the principal eigenvalue λ_1 of L above. 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. 1. Since in this case the bilinear form B is positive: $B[u, u] \ge \sigma \|u\|^2_{H_0^1(\Omega)}$, we have $\lambda_1 > 0$. Let w_1 be an eigenfunction corresponding to λ_1 with $\|w_1\|_{L^2(\Omega)} = 1$, and set

$$w_1^+ = \max\{0, w_1\}, \quad w_1^- = \min\{0, w_1\}.$$

Then $w_1^\pm \in H_0^1(\Omega)$, $w_1 = w_1^+ + w_1^-$, $\|w_1^+\|^2_{L^2(\Omega)} + \|w_1^-\|^2_{L^2(\Omega)} = \|w_1\|^2_{L^2(\Omega)} = 1$, and

$$\nabla w_1^+ = \chi_{\{w_1\geq 0\}}\nabla w_1, \quad \nabla w_1^- = \chi_{\{w_1\leq 0\}}\nabla w_1.$$

Hence we have $B[w_1^+, w_1^-] = 0$, and

$$\lambda_1 = B[w_1, w_1] = B[w_1^+, w_1^+] + B[w_1^-, w_1^-] \geq \lambda_1\|w_1^+\|^2_{L^2(\Omega)} + \lambda_1\|w_1^-\|^2_{L^2(\Omega)} = \lambda_1.$$

But then the inequality must be equality. So

$$B[w_1^+, w_1^+] = \lambda_1\|w_1^+\|^2_{L^2(\Omega)}, \quad B[w_1^-, w_1^-] = \lambda_1\|w_1^-\|^2_{L^2(\Omega)}.$$

Therefore, $u = w_1^\pm$ are both solutions to the elliptic equation

$$\begin{cases} Lu = \lambda_1 u & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

2. Since the coefficients of L and Ω are smooth, $u = w_1^\pm$ are smooth solutions. (See Theorem 3.19 below.) Note that $Lw_1^+ = \lambda_1 w_1^+ \geq 0$ in Ω. By **Strong Maximum Principle**, either $w_1^+ \equiv 0$ or else $w_1^+ > 0$ in Ω; similarly, either $w_1^- \equiv 0$ or else $w_1^- < 0$ in Ω. This proves that either $w_1 < 0$ in Ω or else $w_1 > 0$ in Ω.

3. We now prove the eigenspace of λ_1 is one-dimensional. 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 $u \not\equiv 0$. 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 $x \in \Omega$; hence, in either case, $\int_\Omega u(x)dx \neq 0$, which is a contradiction. □

Remark 3.11. Let $L = -\Delta$. Then, there exists an orthonormal basis $\{w_k\}_{k=1}^\infty$ of $L^2(\Omega)$ consisting eigenfunctions w_k of $-\Delta$ in $H_0^1(\Omega)$. We can see that $\{w_k\}$ is also orthogonal in $H_0^1(\Omega)$; in fact,

$$\int_\Omega \nabla w_k(x)\cdot\nabla w_l(x)\,dx = B[w_k, w_l] = \lambda_k(w_k, w_l)_{L^2(\Omega)} = \lambda_k\delta_{kl} \quad (k, l = 1, 2, \cdots).$$

Furthermore, $w_k \in C^\infty(\Omega)$. If $\partial\Omega$ is smooth, then each w_k is smooth on $\bar\Omega$. See Theorem 3.19 below.

3.5. Regularity

We now address the question as to whether a weak solution u of the PDE

$$Lu = f \quad \text{in } \Omega$$

is smooth or not. This is the **regularity** problem for weak solutions.

We first study second-order linear differential equations of the divergence form

$$Lu \equiv -\sum_{i,j=1}^n D_i\left(a_{ij}(x)D_j u\right) + \sum_{i=1}^n b_i(x)D_i u + c(x)u \tag{3.33}$$

To see that there is some hope that a weak solution may be better than a typical function in $H_0^1(\Omega)$, let us consider the model problem $-\Delta u = f$ in $\mathbb{R}^n$. Assume u is smooth enough to justify the following calculations.

$$\begin{aligned}\int_{\mathbb{R}^n} f^2 dx = \int_{\mathbb{R}^n} (\Delta u)^2 dx &= \sum_{i,j=1}^n \int_{\mathbb{R}^n} u_{x_i x_i} u_{x_j x_j} dx \\ &= -\sum_{i,j=1}^n \int_{\mathbb{R}^n} u_{x_i x_i x_j} u_{x_j} dx \\ &= \sum_{i,j=1}^n \int_{\mathbb{R}^n} u_{x_i x_j} u_{x_i x_j} dx = \int_{\mathbb{R}^n} |D^2 u|^2 dx.\end{aligned}$$

Thus the L^2 norm of the second derivatives of u can be estimated by the L^2 norm of f. Similarly, if we differentiate the PDE with respect to x_k, we see that the L^2 norm of the third derivatives of u can be estimated by the L^2 norm of the first derivatives of f, etc. This suggests that we can expect a weak solution $u \in H_0^1(\Omega)$ to belong to $H^{m+2}(\Omega)$ whenever $f \in H^m(\Omega)$.

The above calculations do not really constitute a proof, since we assumed that u was smooth in order to carry out the calculation. If we merely start with a weak solution in $H_0^1(\Omega)$, we cannot justify the above computations.

3.5.1. Difference Quotient Method. One can instead rely upon an analysis of certain difference quotients to obatin higher regularity of weak solutions in $H^1(\Omega)$. Our first regularity result provides the interior H^2-regularity for weak solutions of the equation $Lu = f$ based on the difference quotient method.

Theorem 3.15. (Interior H^2-regularity) *Let L be uniformly elliptic, with $a_{ij} \in C^1(\Omega)$, b_i and $c \in L^\infty(\Omega)$. Let $f \in L^2(\Omega)$. If $u \in H^1(\Omega)$ is a weak solution of* (3.33), *then for any $\Omega' \subset\subset \Omega$ we have $u \in H^2(\Omega')$, and*

$$\|u\|_{H^2(\Omega')} \leq C\left(\|u\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}\right) \tag{3.34}$$

where the constant C depends only on n, Ω', Ω and the coefficients of L.

Proof. Set $q = f - \sum_{i=1}^n b_i D_i u - cu$. Since u is a weak solution of (3.33), (by the similar definition as above), this means that

$$\int_\Omega \sum_{i,j=1}^n a_{ij} D_i u D_j \varphi dx = \int_\Omega q\varphi dx \quad \forall\, \varphi \in H_0^1(\Omega),\ \operatorname{supp}\varphi \subset\subset \Omega. \tag{3.35}$$

Step 1: (Interior H^1-estimate). Take any $\Omega'' \subset\subset \Omega$. Choose a cutoff function $\zeta \in C_0^\infty(\Omega)$ with $0 \leq \zeta \leq 1$ and $\zeta|_{\Omega''} = 1$. We take $\varphi = \zeta^2 u$ in (3.35) and perform elementary calculations using the ellipticity condition, to discover

$$\int_\Omega \zeta^2 |\nabla u|^2 dx \leq C \int_\Omega (f^2 + u^2) dx.$$

Thus

$$\|u\|_{H^1(\Omega'')} \leq C(\|f\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)}), \tag{3.36}$$

where the constant C depends on Ω''.

Step 2: (Interior H^2-estimate). Take $\Omega' \subset\subset \Omega_1 \subset\subset \Omega_2 \subset\subset \Omega'' \subset\subset \Omega$. Let $v \in H_0^1(\Omega)$ be any function with $\mathrm{supp}\,(v) \subset\subset \Omega_1$. Let

$$\delta = \frac{1}{2} \min \Big\{ \mathrm{dist}\,(\mathrm{supp}\,(v), \partial\Omega_1), \mathrm{dist}\,(\Omega_1, \partial\Omega_2), \mathrm{dist}\,(\Omega_2, \partial\Omega'') \Big\}.$$

Let D_k^h be the difference quotient operator defined above. For $0 < |h| < \delta$, we choose the test function $\varphi = D_k^{-h} v$ in (3.35) and obtain, using integration by parts for difference quotient,

$$\int_\Omega D_k^h (\sum_{i,j=1}^n a_{ij} D_i u) D_j v dx = -\int_\Omega q D_k^{-h} v dx.$$

Notice that the integrals are in fact over domain Ω_1. Henceforth, we omit the $\sum$ sign. Using the definition of q and the equality

$$D_k^h (a_{ij} D_i u) = a_{ij}^h D_k^h D_i u + D_i u D_k^h a_{ij},$$

where $a_{ij}^h(x) = a_{ij}(x + he_k)$, we get

$$\begin{aligned} \int_\Omega a_{ij}^h D_i D_k^h u D_j v dx &= -\int_\Omega \left(D_k^h a_{ij} D_i u D_j v + q D_k^{-h} v \right) dx \\ &\leq C \left(\|u\|_{H^1(\Omega_1)} + \|f\|_{L^2(\Omega_1)} \right) \|\nabla v\|_{L^2(\Omega_2)}. \end{aligned}$$

Take $\eta \in C_0^\infty(\Omega_1)$ such that $\eta(x) = 1$ for $x \in \Omega'$ and choose $v = \eta^2 D_k^h u$. Then

$$\begin{aligned} &\int_\Omega \eta^2 a_{ij}^h D_i D_k^h u D_j D_k^h u dx \leq -2 \int_\Omega \eta a_{ij}^h D_i D_k^h u (D_j \eta) D_k^h u dx \\ &+C \left(\|u\|_{H^1(\Omega_1)} + \|f\|_{L^2(\Omega_1)} \right) \left(\|\eta \nabla D_k^h u\|_{L^2(\Omega_2)} + 2\|D_k^h u \nabla \eta\|_{L^2(\Omega_2)} \right). \end{aligned}$$

Using the ellipticity condition and Cauchy's inequality, we obtain

$$\frac{\theta}{2} \int_\Omega |\eta D_k^h \nabla u|^2 dx \leq C \int_\Omega |\nabla \eta|^2 |D_k^h u|^2 dx + C \left(\|u\|^2_{H^1(\Omega'')} + \|f\|^2_{L^2(\Omega'')} \right).$$

Hence

$$\|\eta D_k^h \nabla u\|^2_{L^2(\Omega)} \leq C \left(\|u\|^2_{H^1(\Omega'')} + \|f\|^2_{L^2(\Omega'')} \right).$$

Since $\eta = 1$ on Ω', by using Theorem 2.34, we derive that $D_k \nabla u \in L^2(\Omega')$. This proves that $u \in H^2(\Omega')$ and

$$\|u\|_{H^2(\Omega')} \leq C \left(\|u\|_{H^1(\Omega'')} + \|f\|_{L^2(\Omega'')} \right), \tag{3.37}$$

where C depends on Ω'. Combining with (3.36) it follows u satisfies (3.34). □

Remark 3.12. (i) The result holds if the coefficients a_{ij} are only (*locally*) Lipschitz continuous in Ω, since the proof above only used the fact that $D_k^h a_{ij}$ is bounded.

(ii) The proof shows that $D_k \nabla u \in L^2(\Omega')$ as long as the function $\varphi = D_k^{-h}(\eta^2 D_k^h u)$ is a function in $H^1(\Omega)$ with compact support in Ω even when $\Omega' \cap \partial\Omega \neq \emptyset$. This is used in the boundary regularity theory later.

By using an induction argument, we can also get higher regularity for the solution.

Theorem 3.16. (Higher interior regularity) *Let L be uniformly elliptic, with $a_{ij} \in C^{k+1}(\Omega)$, $b_i, c \in C^k(\Omega)$, and $f \in H^k(\Omega)$. If $u \in H^1(\Omega)$ is a weak solution of $Lu = f$, then for any $\Omega' \subset\subset \Omega$ we have $u \in H^{k+2}(\Omega')$ and*

$$\|u\|_{H^{k+2}(\Omega')} \leq C \left(\|u\|_{L^2(\Omega)} + \|f\|_{H^k(\Omega)} \right) \tag{3.38}$$

where the constant C depends only on n, Ω', Ω and the coefficients of L.

Proof. Suppose we have proved this theorem for k. Now assume $a_{ij} \in C^{k+2}(\Omega),\ b_i, c \in C^{k+1}(\Omega),\ f \in H^{k+1}(\Omega)$ and $u \in H^1(\Omega)$ is a weak solution of $Lu = f$. Then, by the induction assumption, $u \in H^{k+2}_{loc}(\Omega)$, with the estimate (3.38). We want to show $u \in H^{k+3}_{loc}(\Omega)$. Fix $\Omega' \subset\subset \Omega'' \subset\subset \Omega$ and a multiindex α with $|\alpha| = k+1$. Let

$$\tilde{u} = D^\alpha u \in H^1(\Omega'').$$

Given any $\tilde{v} \in C_0^\infty(\Omega'')$, let $\varphi = (-1)^{|\alpha|} D^\alpha \tilde{v}$ be put into the identity $B_1[u,\varphi] = (f,\varphi)_{L^2(\Omega)}$ and perform some elementary integration by parts, and eventually we discover

$$B_1[\tilde{u},\tilde{v}] = (\tilde{f},\tilde{v})_{L^2(\Omega)},$$

where

$$\tilde{f} := D^\alpha f - \sum_{\beta\le\alpha,\beta\ne\alpha} \binom{\alpha}{\beta} \left[-\sum_{i,j=1}^n (D^{\alpha-\beta} a_{ij} D^\beta u_{x_i})_{x_j} + \sum_{i=1}^n D^{\alpha-\beta} b_i D^\beta u_{x_i} + D^{\alpha-\beta} c D^\beta u \right].$$

That is, $\tilde{u} \in H^1(\Omega'')$ is a weak solution of $L\tilde{u} = \tilde{f}$ on Ω''. (This is equivalent to differentiating the equation $Lu = f$ with D^α-operator.) We have $\tilde{f} \in L^2(\Omega'')$, with, in light of the induction assumption on the $H^{k+2}(\Omega'')$-estimate of u,

$$\|\tilde{f}\|_{L^2(\Omega'')} \le C(\|f\|_{H^{k+1}(\Omega'')} + \|u\|_{H^{k+2}(\Omega'')}) \le C(\|f\|_{H^{k+1}(\Omega)} + \|u\|_{L^2(\Omega)}).$$

Therefore, by Theorem 3.15, $\tilde{u} \in H^2(\Omega')$, with the estimate

$$\|\tilde{u}\|_{H^2(\Omega')} \le C(\|\tilde{f}\|_{L^2(\Omega'')} + \|\tilde{u}\|_{L^2(\Omega'')}) \le C(\|f\|_{H^{k+1}(\Omega)} + \|u\|_{L^2(\Omega)}).$$

This exactly proves $u \in H^{k+3}(\Omega')$ and the corresponding estimate (3.38) with $k+1$. □

3.5.2. Boundary Regularity. We now study the regularity up to the boundary. For this purpose we need certain smoothness of the boundary. We have the following result.

Theorem 3.17. (Global H^2-regularity) *Assume in addition to the assumptions of Theorem 3.15 that $a_{ij} \in C^1(\bar{\Omega})$ and $\partial\Omega \in C^2$. If $u \in H_0^1(\Omega)$ is a weak solution to $Lu = f$, then $u \in H^2(\Omega)$, and*

$$\|u\|_{H^2(\Omega)} \le C(\|u\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}) \tag{3.39}$$

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. 1. First investigate the special case that Ω is a half ball

$$\Omega = B(0,r) \cap \{x_n > 0\}.$$

and $u \equiv 0$ along plane $\{x_n = 0\}$ in the sense of trace. Set $\Omega' = B(0,s) \cap \{x_n > 0\}$, where $0 < s < r$. Then select a smooth cutoff function $\zeta \in C_0^\infty(B(0,r))$ with

$$0 \le \zeta \le 1,\ \zeta|_{B(0,s)} \equiv 1.$$

So $\zeta \equiv 1$ on Ω' and $\zeta = 0$ near the curved part of $\partial\Omega$.

2. Now fix $k \in \{1, 2, \cdots, n-1\}$. For $h > 0$ sufficiently small, let

$$\varphi = D_k^{-h} v, \quad v = \zeta^2 D_k^h u.$$

Let us note carefully that if $x \in \Omega$ then

$$\varphi(x) = \frac{\zeta^2(x - he_k)[u(x) - u(x - he_k)] - \zeta^2(x)[u(x + he_k) - u(x)]}{h^2}.$$

Since $u = 0$ along $\{x_n = 0\}$ and $\zeta = 0$ near the curved portion of $\partial\Omega$, we see $\varphi \in H_0^1(\Omega)$. Therefore, we can use this φ as a test function in (3.35) as we did in the Step 2 above. The end result is that

$$D_k u \in H^1(\Omega') \quad (k = 1, 2, \cdots, n-1)$$

with the estimate

$$\sum_{k,l=1,k+l<2n}^{n} \|D_{kl}u\|_{L^2(\Omega')} \le C(\|f\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)}). \tag{3.40}$$

3. We must estimate $\|D_{nn}u\|_{L^2(\Omega')}$. We now use the fact that the assumption $a_{ij} \in C^1$ and the interior regularity imply the equation $Lu = f$ is satisfied almost everywhere in Ω; in this case we say $u \in H^2_{loc}(\Omega)$ is a **strong solution**. Since the ellipticity implies $a_{nn}(x) \ge \theta > 0$, we can actually solve $D_{nn}u$ from the equation $Lu = f$ in terms of $D_{ij}u$ and $D_i u$ with $i + j < 2n$, $i, j = 1, 2, \cdots, n$. This way we deduce the pointwise estimate

$$|D_{nn}u| \le C \left(\sum_{i,j=1,i+j<2n}^{n} |D_{ij}u| + |\nabla u| + |u| + |f| \right).$$

Hence, by (3.40),

$$\|u\|_{H^2(\Omega')} \le C(\|f\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)}).$$

4. Now it is standard to treat smooth domains by locally flattening the boundary. Since $\partial\Omega$ is C^2, at each point $x^0 \in \partial\Omega$, we have a small ball $B(x^0, r)$ and a C^2 map $y = \Phi(x)$, with $\Phi(x^0) = 0$, that maps $B(x^0, r)$ *bijectively* onto a domain in the y space such that

$$\Phi(\Omega \cap B(x^0, r)) \subset \{y \in \mathbb{R}^n \mid y_n > 0\}.$$

We assume the inverse of this map is $x = \Psi(y)$. Both Ψ and Φ are C^2. Choose $s > 0$ so small that the half-ball $V := B(0, s) \cap \{y_n > 0\}$ lies in $\Phi(\Omega \cap B(x^0, r))$. Set $V' = B(0, s/2) \cap \{y_n > 0\}$. Finally define

$$v(y) = u(\Psi(y)) \quad (y \in V).$$

Then $v \in H^1(V)$ and $v = 0$ on $\partial V \cap \{y_n = 0\}$ (in the sense of trace). Moreover, $u(x) = v(\Phi(x))$ and hence

$$u_{x_i}(x) = \sum_{k=1}^{n} v_{y_k}(\Phi(x)) \Phi^k_{x_i}(x) \ \ (i = 1, 2, \cdots, n).$$

5. We now show that v is a weak solution of a linear PDE $Mv = g$ in V. To find this PDE, let $I(y) = \det \frac{\partial \Psi(y)}{\partial y}$ be the Jacobi matrix of $x = \Psi(y)$; since $I(y) \ne 0$ and $\Psi \in C^2$, we have $|I|, |I|^{-1} \in C^1(\bar{V})$. Let $\zeta \in H^1(V)$ with $\operatorname{supp} \zeta \subset\subset V$ and let $\varphi = \zeta/|I|$. Then $\varphi \in H^1(V)$ with $\operatorname{supp} \varphi \subset\subset V$. Let $w(x) = \varphi(\Phi(x))$ for $x \in \Omega' = \Psi(V)$. Then $w \in H^1(\Omega')$ and $\operatorname{supp} w \subset\subset \Omega'$. We use the weak formulation of $Lu = f$: $B_1[u, w] = (f, w)_{L^2(\Omega)}$ and the change of variable $x = \Psi(y)$ to compute

$$(f, w)_{L^2(\Omega)} = \int_{\Omega'} f(x) w(x) dx = \int_V f(\Psi(y)) \varphi(y) |I(y)| dy := (g, \zeta)_{L^2(V)}, \tag{3.41}$$

where for $g(y) = f(\Psi(y))$. We also compute

$$B_1[u, w] = \int_{\Omega'} \big(a_{ij}(x) u_{x_i}(x) w_{x_j}(x) + b_i(x) u_{x_i} w(x) + c(x) u(x) w(x)\big)\, dx$$

$$= \int_{\Omega'} \Big(a_{ij}(x) v_{y_k}(\Phi(x)) \Phi^k_{x_i}(x) \varphi_{y_l}(\Phi(x)) \Phi^l_{x_j}(x) + b_i(x) v_{y_k}(\Phi(x)) \Phi^k_{x_i}(x) w(x) + c(x) u(x) w(x)\Big)\, dx$$

$$= \int_V \Big(a_{ij}(\Psi(y)) v_{y_k}(y) \Phi^k_{x_i}(\Psi(y)) \varphi_{y_l}(y) \Phi^l_{x_j}(\Psi(y)) \\ + b_i(\Psi(y)) v_{y_k}(y) \Phi^k_{x_i}(\Psi(y)) \varphi(y) + c(\Psi(y)) v(y) \varphi(y) \Big) |I(y)| dy.$$

Since $\varphi_{y_l}|I| = \zeta_{y_l} - \frac{|I|_{y_l}}{|I|}\zeta$, we have

$$B_1[u,w] = \int_V \Big(\tilde{a}_{ij}(y) v_{y_k}(y) \zeta_{y_k}(y) + \tilde{b}_i(y) v_{y_k}(y) \zeta(y) + \tilde{c}(y) v(y) \zeta(y) \Big) dy := \tilde{B}[v,\zeta], \tag{3.42}$$

where $\tilde{c}(y) = c(\Psi(y))$,

$$\tilde{a}_{kl}(y) = \sum_{i,j=1}^n a_{ij}(\Psi(y)) \Phi^k_{x_i}(\Psi(y)) \Phi^l_{x_j}(\Psi(y)) \ \ (k,l = 1,2,\cdots,n),$$

and

$$\tilde{b}_k(y) = \sum_{i=1}^n b_i(x) \Phi^k_{x_i}(x) - \sum_{i,j,l=1}^n a_{ij}(\Psi(y)) \Phi^k_{x_i}(\Psi(y)) \Phi^l_{x_j}(\Psi(y)) \frac{|I|_{y_l}}{|I|} \ \ (k = 1,2,\cdots,n).$$

By (3.41), (3.42), it follows that

$$\tilde{B}[v,\zeta] = (g,\zeta)_{L^2(V)} \text{ for all } \zeta \in H^1(V) \text{ with } \operatorname{supp}\zeta \subset\subset V;$$

hence, $v \in H^1(V)$ is a weak solution of $Mv = g$ in V, where

$$Mv := -\sum_{k,l=1}^n D_{y_l}(\tilde{a}_{kl}(y) D_{y_k} v) + \sum_{k=1}^n \tilde{b}_k(y) D_{y_k} v + \tilde{c}(y) v.$$

6. We easily have that $\tilde{a}_{kl} \in C^1(\bar{V})$, $\tilde{b}_k$, $\tilde{c} \in L^\infty(V)$. We now check that the operator M is uniformly elliptic in V. Indeed, if $y \in V$ and $\xi \in \mathbb{R}^n$, then, again with $x = \Psi(y)$,

$$\sum_{k,l=1}^n \tilde{a}_{kl}(y)\xi_k\xi_l = \sum_{r,s=1}^n \sum_{k,l=1}^n a_{rs}(x) \Phi^k_{x_r} \Phi^l_{x_s} \xi_k \xi_l = \sum_{r,s=1}^n a_{rs}(x) \eta_r(x) \eta_s(x) \ge \theta |\eta(x)|^2,$$

where $\eta(x) = (\eta_1(x), \cdots, \eta_n(x))$, with

$$\eta_r(x) = \sum_{k=1}^n \Phi^k_{x_r}(x) \xi_k \quad (r = 1,2,\cdots,n).$$

That is, $\eta(x) = \xi D\Phi(x)$. Hence $\xi = \eta(x) D\Psi(y)$ with $y = \Phi(x)$. So $|\xi| \le C|\eta(x)|$ for some constant C. This shows that

$$\sum_{k,l=1}^n \tilde{a}_{kl}(y)\xi_k\xi_l \ge \theta|\eta(x)|^2 \ge \theta'|\xi|^2$$

for some constant $\theta' > 0$ and all $y \in V$ and $\xi \in \mathbb{R}^n$. By the result with flat boundary in Step 1, we have

$$\|v\|_{H^2(V')} \le C(\|g\|_{L^2(V)} + \|v\|_{L^2(V)}).$$

Consequently, with $O' = \Psi(V')$, using the fact Φ, Ψ are of C^2, we deduce

$$\|u\|_{H^2(O')} \le C(\|f\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)}). \tag{3.43}$$

Note that $x^0 \in \Psi(B(0,s/2)) := G'$, which is an open set containing open set O'.

7. Since $\partial\Omega$ is compact, we can find finitely many open sets $O_i' \subset G_i'$ $(i = 1, 2, \cdots, k)$ such that $\partial\Omega \subset \cup_{i=1}^k G_i'$. Then there exists a $\delta > 0$ such that

$$F := \{x \in \Omega \mid \operatorname{dist}(x, \partial\Omega) \leq \delta\} \subset \bigcup_{i=1}^{k} O_i'.$$

Then $U = (\Omega \setminus F) \subset\subset \Omega$. By (3.43), we have

$$\|u\|_{H^2(F)} \leq C(\|f\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)}).$$

By interior regularity,

$$\|u\|_{H^2(U)} \leq C(\|f\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)}).$$

Combining these two estimates, we deduce (3.39). □

Theorem 3.18. (Higher global regularity) *Let L be uniformly elliptic, with $a_{ij} \in C^{k+1}(\bar{\Omega})$, $b_i, c \in C^k(\bar{\Omega})$, $f \in H^k(\Omega)$, and $\partial\Omega \in C^{k+2}$. Then a weak solution u of $Lu = f$ satisfying $u \in H_0^1(\Omega)$ belongs to $H^{k+2}(\Omega)$, and*

$$\|u\|_{H^{k+2}(\Omega)} \leq C(\|u\|_{L^2(\Omega)} + \|f\|_{H^k(\Omega)}) \tag{3.44}$$

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, whenever $Lu = f$,

$$\|u\|_{H^{k+2}(\Omega)} \leq C\|f\|_{H^k(\Omega)} \tag{3.45}$$

where C is independent of u and f.

Proof. 1. As above, we first investigate the special case

$$\Omega = B(0, 1) \cap \{x_n > 0\}.$$

Set $\Omega_t = B(0, t) \cap \{x_n > 0\}$ for each $0 < t < 1$. We intend to show by induction on k that whenever $u = 0$ along $\{x_n = 0\}$ (always in the sense of trace), we have $u \in H^{k+2}(\Omega_t)$ and

$$\|u\|_{H^{k+2}(\Omega_t)} \leq C(\|u\|_{L^2(\Omega)} + \|f\|_{H^k(\Omega)}). \tag{3.46}$$

Suppose this is proved with k. Now assume $a_{ij} \in C^{k+2}(\bar{\Omega})$, $b_i, c \in C^{k+1}(\bar{\Omega})$, $f \in H^{k+1}(\Omega)$, and u is a weak solution of $Lu = f$ in Ω, which vanishes along $\{x_n = 0\}$. Fix any $0 < t < r < 1$. By induction assumption, $u \in H^{k+2}(\Omega_r)$, with

$$\|u\|_{H^{k+2}(\Omega_r)} \leq C(\|u\|_{L^2(\Omega)} + \|f\|_{H^k(\Omega)}). \tag{3.47}$$

Furthermore, according to the interior regularity, $u \in H^{k+3}_{loc}(\Omega)$.

2. Let α be any multiindex with $|\alpha| = k + 1$ and $\alpha_n = 0$. Then $\tilde{u} := D^\alpha u \in H^1(\Omega)$ and vanishes along $\{x_n = 0\}$. (For example, this can be shown by induction on $|\alpha|$ using the difference quotient operator D_j^h.) Furthermore, as in the proof of the interior higher regularity theorem, $\tilde{u}$ is a weak solution of $L\tilde{u} = \tilde{f}$ in Ω, with the same $\tilde{f}$ as above. This $\tilde{f}$ belongs to $L^2(\Omega_r)$ and

$$\|\tilde{f}\|_{L^2(\Omega_r)} \leq C(\|f\|_{H^{k+1}(\Omega)} + \|u\|_{L^2(\Omega)}).$$

Consequently, $\tilde{u} \in H^2(\Omega_t)$, with

$$\|\tilde{u}\|_{H^2(\Omega_t)} \leq C(\|\tilde{f}\|_{L^2(\Omega_r)} + \|\tilde{u}\|_{L^2(\Omega_r)}) \leq C(\|f\|_{H^{k+1}(\Omega)} + \|u\|_{L^2(\Omega)}).$$

This proves

$$\|D^\beta u\|_{L^2(\Omega_t)} \leq C(\|f\|_{H^{k+1}(\Omega)} + \|u\|_{L^2(\Omega)}) \tag{3.48}$$

for all β with $|\beta| = k + 3$ and $\beta_n = 0, 1, 2$.

3. We need to extend (3.48) to all β with $|\beta| = k+3$. Fix k, we prove (3.48) for all β by induction on $j = 0, 1, \cdots, k+2$ with $\beta_n \le j$. We have already shown it for $j = 0, 1, 2$. Assume we have shown it for j. Now assume β with $|\beta| = k+3$ and $\beta_n = j+1$. Let us write $\beta = \gamma + \delta$, for $\delta = (0, \cdots, 0, 2)$ and so $|\gamma| = k+1$. Since $u \in H^{k+3}_{loc}(\Omega)$ and $Lu = f$ in Ω, we have $D^\gamma Lu = D^\gamma f$ a.e. in Ω. Now

$$D^\gamma Lu = a_{nn} D^\beta u + R,$$

where R is the sum of terms involving at most j derivatives of u with respect to x_n and at most $k+3$ derivatives in all. Since $a_{nn} \ge \theta$, we can solve $D^\beta u$ in terms of R and $D^\gamma f$; hence,

$$\|D^\beta u\|_{L^2(\Omega_t)} \le C(\|f\|_{H^{k+1}(\Omega)} + \|u\|_{L^2(\Omega)}).$$

By induction, we deduce (3.48), which proves

$$\|u\|_{H^{k+3}(\Omega_t)} \le C(\|f\|_{H^{k+1}(\Omega)} + \|u\|_{L^2(\Omega)}).$$

This estimate in turn completes the induction process on k, begun in step 2. This proves (3.46).

4. As above, we can cover the domain $\bar\Omega$ by finitely many small balls and use the method of flattening the boundary to eventually deduce (3.44).

5. We prove the last statement (3.45) for $k = 0$; the more general case is similar. In view of (3.44), 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 $u_n \in H^2(\Omega) \cap H^1_0(\Omega)$ and $f_n \in L^2(\Omega)$ for which $\|u_n\|_2 = 1$ and $\|f_n\|_2 \to 0$. By (3.44) 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 $v \in H^1_0(\Omega)$, the functional $l(u) = B[u, v] \in (H^2(\Omega))^*$, and so by passing to the limit in

$$l(u_n) = B[u_n, v] = \int_\Omega f_n v dx \quad \text{for all} \quad v \in H^1_0(\Omega)$$

we see that $B[u, v] = 0$ for all $v \in H^1_0(\Omega)$ and thus u is a weak solution of $Lu = 0$. Hence, $u \equiv 0$ by weak uniqueness. This contradicts $\|u\|_2 = \lim_{n\to\infty} \|u_n\|_2 = 1$. □

Finally, we iterate this regualrity theorem to obtain

Theorem 3.19. (Infinite smoothness) *Let L be uniformly elliptic, with a_{ij}, b_i, c, $f \in C^\infty(\bar\Omega)$, and $\partial\Omega \in C^\infty$. Then a weak solution $u \in H^1_0(\Omega)$ of $Lu = f$ belongs to $C^\infty(\bar\Omega)$.*

3.6. Regularity for Linear Systems*

In this extra section, we study the regularity problem for weak solutions of elliptic second-order linear systems. The methods include some direct generalization of linear equations and some new techniques.

Let $A(x, \xi) = A(x)\,\xi$ be a linear matrix function of ξ given by

$$(A(x,\xi))^i_\alpha = (A(x)\xi)^i_\alpha = A^{\alpha\beta}_{ij}(x)\xi^j_\beta,$$

where $A^{\alpha\beta}_{ij} \in L^\infty(\Omega)$. Consider the linear partial differential system

$$-D_\alpha\big(A^{\alpha\beta}_{ij}(x)\, D_\beta u^j\big) = g^i - D_\alpha f^i_\alpha, \quad i = 1, 2, \cdots, N, \tag{3.49}$$

where we assume $g^i,\ f^i_\alpha \in L^2_{loc}(\Omega)$. We also write this system as

$$-\operatorname{div}(A(x)Du) = g - \operatorname{div} f,$$

where $g = (g^i)$, $f = (f^i_\alpha)$. Recall that $u \in H^1_{loc}(\Omega;\mathbb{R}^N)$ is a *weak solution* of (3.49) if

$$\int_\Omega A(x)Du \cdot D\phi(x)\,dx = \int_\Omega (g(x)\cdot\phi(x) + f(x)\cdot D\phi(x))\,dx \tag{3.50}$$

holds for all $\phi \in C^\infty_0(\Omega;\mathbb{R}^N)$. Since $A^{\alpha\beta}_{ij} \in L^\infty(\Omega)$, the test function ϕ in (3.50) can be chosen in $H^1_0(\Omega';\mathbb{R}^N)$ for any subdomain $\Omega' \subset\subset \Omega$.

The regularity for system (3.49) relies on certain *ellipticity* conditions. We shall assume one of the following conditions holds: with some constant $\nu > 0$,

$$(H1)\qquad A^{\alpha\beta}_{ij}(x)\,\xi^i_\alpha\xi^j_\beta \ge \nu\,|\xi|^2.$$

$$(H2)\qquad A^{\alpha\beta}_{ij} \text{ are constants,}\quad A^{\alpha\beta}_{ij}\,p_\alpha p_\beta\,q^i q^j \ge \nu\,|p|^2|q|^2.$$

$$(H3)\qquad A^{\alpha\beta}_{ij} \in C(\bar\Omega),\ \ A^{\alpha\beta}_{ij}(x)\,p_\alpha p_\beta\,q^i q^j \ge \nu\,|p|^2|q|^2.$$

Under such a condition, we shall have the following Gårding inequality holds:

$$\int_{B_R} A(x)D\psi\cdot D\psi \ge \nu_0\int_{B_R}|D\psi|^2 - \nu_1\int_{B_R}|\psi|^2,\quad \forall\,\psi \in H^1_0(B_R;\mathbb{R}^N), \tag{3.51}$$

where $\nu_0 > 0$, $\nu_1 \ge 0$ are constants. For, under the hypothesis (H1) or (H2) the Gårding inequality (3.51) holds with $\nu_0 = \nu$, $\nu_1 = 0$, and under (H3) the inequality (3.51) also holds (see Theorem 3.8).

3.6.1. Caccioppoli-type Estimates. Assume $u \in H^1_{loc}(\Omega;\mathbb{R}^N)$ is a weak solution of (3.49). Almost all the estimates pertaining to regularity of u are derived using test functions of the form $\phi = \eta\,(u-\lambda)$, where η is a *cut-off* function which belongs to $W^{1,\infty}_0(\Omega')$ for certain $\Omega' \subset\subset \Omega$. Let $B_\rho \subset\subset B_R \subset\subset \Omega$ be concentric balls with center $a \in \Omega$. Let

$$\theta(t) = \begin{cases} 1 & \text{if } 0 \le t \le \rho,\\ \frac{R-t}{R-\rho} & \text{if } \rho \le t \le R,\\ 0 & \text{if } t > R.\end{cases}$$

Let $\zeta = \zeta_{\rho,R}(x) = \theta(|x-a|)$. Then $\zeta \in W^{1,\infty}_0(\Omega)$ with $\operatorname{supp}\zeta \subseteq \bar B_R$ and

$$0 \le \zeta \le 1,\quad \zeta|_{B_\rho} \equiv 1,\quad |D\zeta| \le \frac{\chi_{\rho,R}}{R-\rho}, \tag{3.52}$$

where $\chi_{\rho,R} = \chi_{B_R\setminus B_\rho}(x)$ is the characteristic function of $B_R \setminus B_\rho$. Define

$$\psi = \zeta\,(u-\lambda),\quad \phi = \zeta^2(u-\lambda) = \zeta\,\psi.$$

Then $\psi,\ \phi \in H^1_0(B_R;\mathbb{R}^N)$ and

$$D\phi = \zeta D\psi + \psi\otimes D\zeta,\quad D\psi = \zeta Du + (u-\lambda)\otimes D\zeta.$$

Using ϕ as a test function in (3.50) yields

$$\begin{aligned}\int_{B_R}(g\cdot\phi + f\cdot D\phi) &= \int_{B_R} A(x)Du\cdot D\phi\\ &= \int_{B_R} A(x)Du\cdot\zeta D\psi + \int_{B_R} A(x)Du\cdot\psi\otimes D\zeta.\end{aligned}$$

Note that

$$\begin{aligned} A(x)D\psi \cdot D\psi &= A(x)Du \cdot \zeta D\psi + A(x)(u-\lambda)\otimes D\zeta \cdot D\psi, \\ A(x)Du \cdot \psi \otimes D\zeta &= A(x)\zeta Du \cdot (u-\lambda)\otimes D\zeta \\ &= A(x)D\psi \cdot (u-\lambda)\otimes D\zeta \\ &- A(x)(u-\lambda)\otimes D\zeta \cdot (u-\lambda)\otimes D\zeta. \end{aligned}$$

Therefore, it follows that

$$\begin{aligned} \int_{B_R} A(x)D\psi \cdot D\psi &= \int_{B_R} A(x)(u-\lambda)\otimes D\zeta \cdot D\psi \\ &+ \int_{B_R} (g\cdot\phi + f\cdot D\phi) - \int_{B_R} A(x)D\psi \cdot (u-\lambda)\otimes D\zeta \\ &+ \int_{B_R} A(x)(u-\lambda)\otimes D\zeta \cdot (u-\lambda)\otimes D\zeta. \end{aligned}$$

(Note that the first and third terms on the righthand side would cancel out if $A(x)\xi\cdot\eta$ is symmetric in ξ, η.) Then, by (3.51), it follows that

$$\begin{aligned} \nu_0 \int_{B_R} |D\psi|^2 &\le \int_{B_R} A(x)D\psi\cdot D\psi + \nu_1 \int_{B_R} |\psi|^2 \\ &\le \Big|\int_{B_R} g\cdot\phi\Big| + \int_{B_R} |f|\cdot|D\psi| + \int_{B_R} |f|\cdot \frac{\chi_{\rho,R}\,|u-\lambda|}{R-\rho} \\ &+ C\int_{B_R} |D\psi|\cdot \frac{\chi_{\rho,R}\,|u-\lambda|}{R-\rho} + C\int_{B_R} \frac{\chi_{\rho,R}\,|u-\lambda|^2}{(R-\rho)^2} + \nu_1\int_{B_R} |u-\lambda|^2. \end{aligned}$$

Using the Cauchy inequality with ϵ, we deduce

$$\int_{B_R} |D\psi|^2 \le C\left[\int_{B_R\setminus B_\rho} \frac{|u-\lambda|^2}{(R-\rho)^2} + \Big|\int_{B_R} g\cdot\phi\Big| + \int_{B_R} |f|^2 + \nu_1 \int_{B_R} |u-\lambda|^2\right]. \tag{3.53}$$

Since $\psi|_{B_\rho} = u-\lambda$, this last estimate (3.53) proves the following theorems.

Theorem 3.20. *Let $u \in H^1_{loc}(\Omega;\mathbb{R}^N)$ be a weak solution of* (3.49). *Assume either condition* $(H1)$ *or* $(H2)$ *holds. Then*

$$\int_{B_\rho} |Du|^2 \le C\left[\int_{B_R\setminus B_\rho} \frac{|u-\lambda|^2}{(R-\rho)^2} + \Big|\int_{B_R} g\cdot\zeta^2(u-\lambda)\Big| + \int_{B_R} |f|^2\right] \tag{3.54}$$

for all concentric balls $B_\rho \subset\subset B_R \subset\subset \Omega$ and constants $\lambda\in\mathbb{R}^N$, where $\zeta = \zeta_{\rho,R}$ and $C>0$ is a constant depending on the L^∞-norm of $A^{\alpha\beta}_{ij}$.

Theorem 3.21. *Assume condition* $(H3)$ *holds. Then*

$$\int_{B_\rho} |Du|^2 \le C\left[\int_{B_R\setminus B_\rho} \frac{|u-\lambda|^2}{(R-\rho)^2} + \Big|\int_{B_R} g\cdot\zeta^2(u-\lambda)\Big| + \int_{B_R} (|f|^2 + |u-\lambda|^2)\right] \tag{3.55}$$

for all concentric balls $B_\rho \subset\subset B_R \subset\subset \Omega$ and constants $\lambda\in\mathbb{R}^N$.

Corollary 3.22. *Let $u \in H^1_{loc}(\Omega;\mathbb{R}^N)$ be a weak solution of* (3.49). *Assume either condition* $(H1)$ *or* $(H2)$ *holds. Then*

$$\int_{B_{R/2}} |Du|^2 \le C\left[\int_{B_R\setminus B_{R/2}} \frac{|u-\lambda|^2}{R^2} + \int_{B_R} \Big(|u-\lambda|\cdot|g| + |f|^2\Big)\right] \tag{3.56}$$

for all balls $B_R \subset\subset \Omega$ and constants $\lambda \in \mathbb{R}^N$.

Remark 3.13. 1) The estimates (3.54), (3.55) and (3.56) are usually referred to as the *Caccioppoli-type* inequalities or Caccioppoli estimates.

2) In both (3.54), (3.55), we keep the term $\int_{B_R} g \cdot \zeta^2(u-\lambda)$ in the estimates. We shall see later that this term needs a special consideration when we deal with higher regularity for weak solutions, especially when g is of the form of quotient difference.

As an application of these Caccioppoli estimates, we prove the following results of **Liouville's theorem**.

Corollary 3.23. *Suppose $u \in H^1_{loc}(\mathbb{R}^n;\mathbb{R}^N)$ is a weak solution of*

$$-D_\alpha(A^{\alpha\beta}_{ij}(x)D_\beta u^j) = 0, \tag{3.57}$$

where coefficients $A^{\alpha\beta}_{ij}(x)$ satisfy $(H1)$ or $(H2)$. If $|Du| \in L^2(\mathbb{R}^n)$, then u is a constant.

Proof. By (3.56), it follows that

$$\int_{B_{R/2}} |Du|^2 \le \frac{C}{R^2}\int_{B_R\setminus B_{R/2}} |u-\lambda|^2.$$

We choose

$$\lambda = \frac{1}{|B_R \setminus B_{R/2}|}\int_{B_R\setminus B_{R/2}} u(x)\,dx.$$

Then the Poincaré inequality shows that

$$\int_{B_R\setminus B_{R/2}} |u-\lambda|^2 \le c(n)\,R^2\int_{B_R\setminus B_{R/2}} |Du|^2.$$

Therefore

$$\int_{B_{R/2}} |Du|^2 \le C\int_{B_R\setminus B_{R/2}} |Du|^2.$$

Adding $C\int_{B_{R/2}} |Du|^2$ to both sides of this inequality (a.k.a. the *hole-filling* technique of Widman), we obtain

$$\int_{B_{R/2}} |Du|^2 \le \frac{C}{C+1}\int_{B_R} |Du|^2.$$

Letting $R \to \infty$ we have

$$\int_{\mathbb{R}^n} |Du|^2\,dx \le \frac{C}{C+1}\int_{\mathbb{R}^n} |Du|^2\,dx.$$

Since $\frac{C}{C+1} < 1$ we have $\int_{\mathbb{R}^n} |Du|^2 = 0$ and thus $Du \equiv 0$; hence $u \equiv$ constant. □

Corollary 3.24. *Assume either condition $(H1)$ or $(H2)$ holds. Then any bounded weak solution $u \in H^1_{loc}(\mathbb{R}^2;\mathbb{R}^N)$ to (3.57) for $n = 2$ must be constant.*

Proof. Let $|u| \le M$; then by the Caccioppoli inequality (3.56) with $\lambda = 0$ we have

$$\int_{B_{R/2}} |Du|^2\,dx \le CM < \infty, \quad \forall\, R > 0.$$

This implies $|Du| \in L^2(\mathbf{R}^2)$; hence by the result above, u is a constant. □

3.6.2. Regularity – Difference Quotient Method. This is completely similar to the scalar case studied above.

Assume $u \in H^1_{loc}(\Omega;\mathbb{R}^N)$ is a weak solution of linear system

$$-\operatorname{div}(A(x)\,Du) = g - \operatorname{div} f.$$

This means

$$\int_\Omega A(x)\,Du(x)\cdot D\phi(x)\,dx = \int_\Omega g(x)\cdot\phi(x)\,dx + \int_\Omega f(x)\cdot D\phi(x)\,dx$$

holds for all $\phi \in W^{1,2}_0(\Omega';\mathbb{R}^N)$. If $0<|h|<\operatorname{dist}(\Omega';\partial\Omega)$ we have

$$\int_\Omega A(x+he_s)\,Du(x+he_s)\cdot D\phi(x)\,dx$$
$$= \int_\Omega g(x+he_s)\cdot\phi(x)\,dx + \int_\Omega f(x+he_s)\cdot D\phi(x)\,dx.$$

Subtract two equations and divide by h to get

$$\begin{aligned}\int_\Omega A(x+he_s)\,DD^h_s u\cdot D\phi &= \int_\Omega D^h_s g(x)\cdot\phi(x)\,dx \\ &+ \int_\Omega D^h_s f(x)\cdot D\phi(x)\,dx - \int_\Omega D^h_s A(x)\,Du(x)\cdot D\phi(x)\,dx.\end{aligned}$$

This shows that $v = D^h_s u$ is a weak solution of system

$$-\operatorname{div}(A(x+he_s)\,Dv) = D^h_s g - \operatorname{div}(D^h_s f) + \operatorname{div}(D^h_s A\,Du) \quad \text{on } \Omega'. \tag{3.58}$$

Assume that Gårding's inequality (3.51) holds. Then we can invoke the estimate (3.53) with $\lambda=0$, $\rho = R/2$ to obtain

$$\begin{aligned}\int_{B_R}|D\psi|^2 \le C\Big[\int_{B_R}\frac{1}{R^2}\,|D^h_s u|^2 + \Big|\int_{B_R} D^h_s g\cdot\phi\Big| \\ + \int_{B_R}\big(|D^h_s f|^2 + \nu_1|D^h_s u|^2 + |D^h_s A|^2|Du|^2\big)\Big],\end{aligned} \tag{3.59}$$

where $\psi = \zeta\, D^h_s u$, $\phi = \zeta^2\, D^h_s u$ and $\zeta = \zeta_{R/2,R}$ is defined as before. Note that

$$\begin{aligned}-\int_{B_R} D^h_s g\cdot\phi &= \int_\Omega D^h_s g\cdot\phi = \int_\Omega g\cdot D^{-h}_s\phi \\ &= \int_\Omega g\cdot\zeta(x-he_s)\,D^{-h}_s\psi + \int_\Omega g\cdot\psi\,D^{-h}_s\zeta \\ &\equiv I + II.\end{aligned}$$

We estimate I, II as follows.

$$\begin{aligned}|I| &\le \int_{\Omega'}|g|\cdot|D^{-h}_s\psi| \le \epsilon\int_{\Omega'}|D^{-h}_s\psi|^2 + C_\epsilon\int_{\Omega'}|g|^2 \\ &\le \epsilon\int_\Omega|D_s\psi|^2 + C_\epsilon\int_{\Omega'}|g|^2 \le \epsilon\int_{B_R}|D\psi|^2 + C_\epsilon\int_{\Omega'}|g|^2,\end{aligned}$$

where $\Omega' \subset\subset \Omega$ is a domain containing $\bar B_R$.

$$|II| \le \int_{B_R}|g|\,|D^h_s u|\,|D^{-h}_s\zeta| \le \frac{C}{R^2}\int_{\Omega'}|g|^2 + C\int_{B_R}|D^h_s u|^2.$$

Combining these estimates with (3.59) yields

$$\int_{B_R} |D\psi|^2 \le C(R) \int_{\Omega'} \Big(|D_s^h u|^2 + |g|^2 + |D_s^h f|^2 + |D_s^h A|^2 |Du|^2\Big).$$

Since $\psi = D_s^h u$ and $D\psi = D_s^h Du$ on $B_{R/2}$ we have

$$\int_{B_{R/2}} |D_s^h Du|^2 \le C(R) \int_{\Omega'} \Big(|D_s^h u|^2 + |g|^2 + |D_s^h f|^2 + |D_s^h A|^2 |Du|^2\Big). \tag{3.60}$$

Finally, if we assume $f \in H^1_{loc}(\Omega; \mathbb{M}^{N\times n})$ and $A(x)$ is Lipschitz continuous with Lipschitz constant K then

$$\int_{\Omega'} |D_s^h f|^2 \le \int_{\Omega''} |Df|^2, \quad |D_s^h A(x)| \le K,$$

where $\Omega' \subset\subset \Omega'' \subset\subset \Omega$, and hence by (3.60) we have

$$\int_{B_{R/2}} |D_s^h Du|^2(x)\, dx \le M < \infty \quad \forall\ |h| << 1,$$

and thus $D_s Du$ exists and belongs to $L^2(B_{R/2}; \mathbb{M}^{N\times n})$ for all $s = 1, 2, \cdots, n$. This implies $u \in H^2_{loc}(\Omega; \mathbb{R}^N)$. Therefore, we have proved the following theorem.

Theorem 3.25. *Suppose $A \in C(\bar\Omega)$ is Lipschitz continuous and the Gårding inequality (3.51) holds. If $g \in L^2_{loc}(\Omega; \mathbb{R}^N)$, $f \in H^1_{loc}(\Omega; \mathbb{M}^{N\times n})$ and $u \in H^1_{loc}(\Omega; \mathbb{R}^N)$ is a weak solution of the system*

$$-\operatorname{div}(A(x)\, Du) = g - \operatorname{div} f$$

then $u \in H^2_{loc}(\Omega; \mathbb{R}^N)$.

The following higher regularity result can be proved by the standard bootstrap method.

Theorem 3.26. *Suppose $u \in H^1_{loc}(\Omega; \mathbb{R}^N)$ is a weak solution of the system*

$$-\operatorname{div}(A(x)\, Du) = g - \operatorname{div} f$$

with $A \in C^{k,1}(\bar\Omega)$ (that is, $D^k A$ is Lipschitz continuous) satisfying the Gårding inequality (3.51) and $g \in H^k_{loc}(\Omega; \mathbb{R}^N)$, $f \in H^{k+1}_{loc}(\Omega; \mathbb{M}^{N\times n})$. Then $u \in H^{k+2}_{loc}(\Omega; \mathbb{R}^N)$.

Proof. Let $\psi \in C_0^\infty(\Omega; \mathbb{R}^N)$; then we use $\phi = D_s\psi$ as a test function for the system to obtain

$$\int_\Omega D_s(A(x)\, Du) \cdot D\psi\, dx = \int_\Omega D_s g \cdot \psi + \int_\Omega D_s f \cdot D\psi.$$

Since $D_s(A(x)\, Du) = (D_s A)\, Du + A(x)\, DD_s u$ we thus have

$$\int_\Omega A(x)\, D(D_s u) \cdot D\psi = \int_\Omega D_s g \cdot \psi + \int_\Omega \big(D_s f - (D_s A)\, Du\big) \cdot D\psi.$$

This shows $v = D_s u \in H^1_{loc}(\Omega; \mathbb{R}^N)$ is a weak solution of

$$-\operatorname{div}(A(x)\, Dv) = D_s g - \operatorname{div}(D_s f - (D_s A)\, Du),$$

and hence $v \in H^2_{loc}(\Omega; \mathbb{R}^N)$; that is, $u \in H^3_{loc}(\Omega; \mathbb{R}^N)$. The result for general k then follows from induction. □

Remark 3.14. Note that if A, g, f are all of C^∞ then any weak solution $u \in H^1_{loc}(\Omega; \mathbb{R}^N)$ must be in $C^\infty(\Omega; \mathbb{R}^N)$. We also have the following result.

Theorem 3.27. *Assume* $(H2)$ *holds. Let* $u \in H^1_{loc}(\Omega;\mathbb{R}^N)$ *be a weak solution of*

$$-D_\alpha(A^{\alpha\beta}_{ij}\, D_\beta u^j) = 0, \quad i = 1,2,...,N. \tag{3.61}$$

Then $u \in H^k_{loc}(\Omega;\mathbb{R}^N)$ *for all* $k = 1,2,...$ *and*

$$\|u\|_{H^k(B_{R/2};\mathbb{R}^N)} \le C(k,R)\,\|u\|_{L^2(B_R;\mathbb{R}^N)}$$

for any ball $B_R \subset\subset \Omega$.

Proof. By the Caccioppoli-type inequality, we have for any weak solution u of system (3.61)

$$\int_{B_\rho} |Du|^2\,dx \le \frac{C}{(R-\rho)^2}\int_{B_R}|u|^2\,dx.$$

The regularity result shows that $u \in W^{k,2}_{loc}(\Omega;\mathbb{R}^N)$ for all k and then it follows that any derivative $D^k u$ is also a weak solution of (3.61). Therefore, the conclusion will follow from a successive use of the above Caccioppoli inequality with a finite number of $R/2 = \rho_1 < \rho_2 < \cdots < \rho_K = R$. □

In exactly the same way as the scalar equations, we obtain the global H^2-regularity.

Theorem 3.28. (Global H^2-regularity) *Let* $\partial\Omega$ *be* C^2. *Suppose* $A \in C^1(\bar\Omega)$ *and the Gårding inequality* (3.51) *holds. If* $g \in L^2(\Omega;\mathbb{R}^N)$, $f \in H^1(\Omega;\mathbb{M}^{N\times n})$ *and* $u \in H^1(\Omega;\mathbb{R}^N)$ *is a weak solution of the system*

$$-\operatorname{div}(A(x)\,Du) = g - \operatorname{div} f$$

then $u \in H^2(\Omega;\mathbb{R}^N)$ *and we have*

$$\|u\|_{H^2(\Omega)} \le C(\|g\|_{L^2(\Omega)} + \|f\|_{H^1(\Omega)} + \|u\|_{L^2(\Omega)}), \tag{3.62}$$

where the constant C *depends only on* Ω *and the coefficients.*

3.6.3. Morrey and Campanato Spaces. Let $\Omega \subset \mathbb{R}^n$ be a bounded open domain. For $x \in \mathbb{R}^n$, $\rho > 0$ let

$$\Omega(x,\rho) = \{y \in \Omega \,|\, |y-x| < \rho\}.$$

Definition 3.15. For $1 \le p < \infty$, $\lambda \ge 0$ we define the **Morrey space** $L^{p,\lambda}(\Omega;\mathbb{R}^N)$ by

$$L^{p,\lambda}(\Omega;\mathbb{R}^N) = \left\{ u \in L^p(\Omega;\mathbb{R}^N) \,\middle|\, \sup_{\substack{a\in\Omega \\ 0<\rho<\operatorname{diam}\Omega}} \rho^{-\lambda}\int_{\Omega(a,\rho)} |u(x)|^p\,dx < \infty \right\}.$$

We define a norm by

$$\|u\|_{L^{p,\lambda}(\Omega;\mathbb{R}^N)} = \sup_{\substack{a\in\Omega \\ 0<\rho<\operatorname{diam}\Omega}} \left(\rho^{-\lambda}\int_{\Omega(a,\rho)}|u(x)|^p\,dx\right)^{1/p}.$$

Theorem 3.29. $L^{p,\lambda}(\Omega;\mathbb{R}^N)$ *is a Banach space.*

Lemma 3.30. (Lebesgue Differentiation Theorem) *If* $v \in L^1_{loc}(\Omega)$ *then*

$$\lim_{\rho\to 0} \fint_{B_\rho(x)} |v(x) - v(y)|\,dy = 0$$

for almost every $x \in \Omega$.

Theorem 3.31. *(a) If* $\lambda > n$ *then* $L^{p,\lambda}(\Omega;\mathbb{R}^N) = \{0\}$.

(b) $L^{p,0}(\Omega;\mathbb{R}^N) \cong L^p(\Omega;\mathbb{R}^N)$; $L^{p,n}(\Omega;\mathbb{R}^N) \cong L^\infty(\Omega;\mathbb{R}^N)$.

(c) If $1 \le p \le q < \infty$, $\frac{n-\lambda}{p} \ge \frac{n-\mu}{q}$, *then* $L^{q,\mu}(\Omega;\mathbb{R}^N) \subset L^{p,\lambda}(\Omega;\mathbb{R}^N)$.

Proof. (a) By Lebesgue's differentiation theorem,

$$|u(a)| = \lim_{\rho\to 0} ⨍_{\Omega(a,\rho)} |u(x)|\,dx, \quad \forall\ a.e.\ a \in \Omega. \tag{3.63}$$

Now, by Hölder's inequality,

$$⨍_{\Omega(a,\rho)} |u(x)|\,dx \le \left(⨍_{\Omega(a,\rho)} |u(x)|^p\,dx \right)^{1/p} \le C\,\rho^{\frac{\lambda-n}{p}}\,\|u\|_{L^{p,n}(\Omega;\mathbb{R}^N)}. \tag{3.64}$$

If $\lambda > n$, letting $\rho \to 0$ we have $u(a) = 0$ for almost every $a \in \Omega$; thus $u \equiv 0$.

(b) That $L^{p,0}(\Omega;\mathbb{R}^N) \cong L^p(\Omega;\mathbb{R}^N)$ easily follows from the definition. We prove $L^{p,n}(\Omega;\mathbb{R}^N) \cong L^\infty(\Omega;\mathbb{R}^N)$. If $u \in L^\infty(\Omega;\mathbb{R}^N)$, then

$$\rho^{-n} \int_{\Omega(a,\rho)} |u(x)|^p\,dx \le C\,\|u\|_\infty^p$$

so that $\|u\|_{L^{p,n}} \le C\,\|u\|_\infty$. Suppose now $u \in L^{p,n}(\Omega;\mathbb{R}^N)$. Then by (3.63), (3.64)

$$|u(a)| = \lim_{\rho\to 0} ⨍_{\Omega(a,\rho)} |u| \le C\,\|u\|_{L^{p,n}(\Omega;\mathbb{R}^N)}.$$

Hence $\|u\|_{L^\infty(\Omega;\mathbb{R}^N)} \le C\,\|u\|_{L^{p,n}(\Omega;\mathbb{R}^N)}$. Therefore $L^{p,n}(\Omega;\mathbb{R}^N) \cong L^\infty(\Omega;\mathbb{R}^N)$.

(c) We first note that $u \in L^{p,\lambda}(\Omega;\mathbb{R}^N)$ if and only if $\int_{\Omega(a,\rho)} |u(x)|^p\,dx \le C\,\rho^\lambda$ for all $a \in \Omega$ and $0 < \rho < \rho_0 = \min\{1, \operatorname{diam}\Omega\}$. Suppose $u \in L^{q,\mu}(\Omega;\mathbb{R}^N)$. Then, by Hölder's inequality, for all $a \in \Omega$, $0 < \rho < \rho_0 < 1$,

$$\begin{aligned}
\int_{\Omega(a,\rho)} |u|^p\,dx &\le |\Omega(a,\rho)|^{1-\frac{p}{q}} \left(\int_{\Omega(a,\rho)} |u|^q\,dx \right)^{\frac{p}{q}} \\
&\le C\,\rho^{n-\frac{np}{q}}\,\left(\|u\|^q_{L^{q,\mu}(\Omega;\mathbb{R}^N)}\,\rho^\mu\right)^{\frac{p}{q}} \\
&\le C\,\rho^{\frac{\mu p}{q}+n-\frac{np}{q}}\,\|u\|^p_{L^{q,\mu}(\Omega;\mathbb{R}^N)} \\
&\le C\,\rho^\lambda\,\|u\|^p_{L^{q,\mu}(\Omega;\mathbb{R}^N)},
\end{aligned}$$

where we have used the assumption $\frac{\mu p}{q} + n - \frac{np}{q} \ge \lambda$ and the fact $0 < \rho < 1$. Therefore, $u \in L^{p,\lambda}(\Omega;\mathbb{R}^N)$. □

Definition 3.16. For $1 \le p < \infty$, $\lambda \ge 0$ we define the **Campanato space** $\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)$ by

$$\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N) = \left\{ u \in L^p(\Omega;\mathbb{R}^N) \;\middle|\; \sup_{\substack{a\in\Omega \\ 0<\rho<\operatorname{diam}\Omega}} \rho^{-\lambda} \int_{\Omega(a,\rho)} |u - u_{a,\rho}|^p\,dx < \infty \right\},$$

where $u_{a,\rho}$ is the average of u on $\Omega(a,\rho)$. Define the seminorm and norm by

$$\begin{aligned}
[u]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)} &= \sup_{\substack{a\in\Omega \\ 0<\rho<\operatorname{diam}\Omega}} \left(\rho^{-\lambda} \int_{\Omega(a,\rho)} |u - u_{a,\rho}|^p\,dx \right)^{1/p}, \\
\|u\|_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)} &= \|u\|_{L^p(\Omega;\mathbb{R}^N)} + [u]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)}.
\end{aligned}$$

Remark 3.17. For the estimates on linear elliptic systems of second-order partial differential equations, only spaces $L^{2,\lambda}(\Omega;\mathbb{R}^N)$ and $\mathcal{L}^{2,\lambda}(\Omega;\mathbb{R}^N)$ are needed. The spaces with other $p \geq 1$ are useful for nonlinear problems; however, we shall not study the nonlinear problems in this course.

For $0 < \alpha \leq 1$, we define the **Hölder space** $C^{0,\alpha}(\bar\Omega;\mathbb{R}^N)$ by

$$C^{0,\alpha}(\bar\Omega;\mathbb{R}^N) = \left\{ v \in L^\infty(\Omega;\mathbb{R}^N) \,\middle|\, |v(x)-v(y)| \leq C\,|x-y|^\alpha,\ \forall\, x,\, y \in \Omega \right\}$$

and define the seminorm and norm by

$$\begin{aligned} [v]_{C^{0,\alpha}(\Omega;\mathbb{R}^N)} &= \sup_{\substack{x,\,y\in\Omega \\ x\neq y}} \frac{|v(x)-v(y)|}{|x-y|^\alpha}, \\ \|v\|_{C^{0,\alpha}(\Omega;\mathbb{R}^N)} &= \|v\|_{L^\infty(\Omega;\mathbb{R}^N)} + [v]_{C^{0,\alpha}(\Omega;\mathbb{R}^N)}. \end{aligned}$$

Theorem 3.32. *Both $\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)$ and $C^{0,\alpha}(\bar\Omega;\mathbb{R}^N)$ are Banach spaces.*

Theorem 3.33. (*a*) *For any $p \geq 1$, $\lambda \geq 0$, $L^{p,\lambda}(\Omega;\mathbb{R}^N) \subset \mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)$.*

(*b*) *For any $0 < \alpha \leq 1$, $C^{0,\alpha}(\bar\Omega;\mathbb{R}^N) \subset \mathcal{L}^{p,n+p\alpha}(\Omega;\mathbb{R}^N)$.*

Proof. (a) Note that

$$\left(\int_{\Omega(a,\rho)} |v(y) - v_{a,\rho}|^p \, dx \right)^{1/p} \leq \|v\|_{L^p(\Omega(a,\rho)} + |v_{a,\rho}| \cdot |\Omega(a,\rho)|^{1/p}.$$

It turns out that we can exactly estimate the two terms on the right-hand side by

$$\|v\|_{L^p(\Omega(a,\rho))} \leq \rho^{\lambda/p}\, \|v\|_{L^{p,\lambda}(\Omega;\mathbb{R}^N)},$$

$$|v_{a,\rho}| \cdot |\Omega(a,\rho)|^{1/p} \leq \rho^{\lambda/p}\, \|v\|_{L^{p,\lambda}(\Omega;\mathbb{R}^N)}$$

so that it follows that

$$[v]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)} \leq 2\,\|v\|_{L^{p,\lambda}(\Omega;\mathbb{R}^N)}.$$

Hence $L^{p,\lambda}(\Omega;\mathbb{R}^N) \subset \mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)$.

(b) Assume $v \in C^{0,\alpha}(\bar\Omega;\mathbb{R}^N)$. Then

$$\begin{aligned} |v(x) - v_{a,\rho}| &= |\fint_{\Omega(a,\rho)} (v(x) - v(y))\, dy| \\ &\leq [v]_{C^{0,\alpha}(\bar\Omega;\mathbb{R}^N)} \cdot \fint_{\Omega(a,\rho)} |x-y|^\alpha \, dy \\ &\leq [v]_{C^{0,\alpha}(\bar\Omega;\mathbb{R}^N)} \cdot (2\rho)^\alpha. \end{aligned}$$

Hence

$$\int_{\Omega(a,\rho)} |v(x) - v_{a,\rho}|^p \, dx \leq C\,[v]^p_{C^{0,\alpha}(\bar\Omega;\mathbb{R}^N)} \cdot \rho^{n+p\alpha}$$

and hence

$$[v]_{\mathcal{L}^{p,n+p\alpha}(\Omega;\mathbb{R}^N)} \leq C\,[v]_{C^{0,\alpha}(\bar\Omega;\mathbb{R}^N)}. \tag{3.65}$$

The proof is complete. □

In order to study the properties of Campanato functions, we need a condition on domain Ω introduced by Campanato.

Definition 3.18. We say that $\Omega \subset \mathbb{R}^n$ is **of type** A if there exists a constant $A > 0$ such that

$$|\Omega(a,\rho)| \geq A\,\rho^n, \quad \forall\, a \in \Omega,\ 0 < \rho < \mathrm{diam}\Omega. \tag{3.66}$$

This condition excludes that Ω may have sharp outward cusps; for instance, all Lipschitz domains are of type A.

Lemma 3.34. *Assume Ω is of type A and $u \in \mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)$. Then for any $0 < r < R < \infty$, $a \in \Omega$ it follows that*

$$|u_{a,R} - u_{a,r}| \leq 2\,A^{-\frac{1}{p}}\,R^{\frac{\lambda}{p}}\,r^{-\frac{n}{p}} \cdot [u]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)}.$$

Proof.

$$\begin{aligned}
|u_{a,R} - u_{a,r}| \cdot |\Omega(a,r)|^{\frac{1}{p}} &= \|u_{a,R} - u_{a,r}\|_{L^p(\Omega(a,r))} \\
&\leq \|u - u_{a,R}\|_{L^p(\Omega(a,r))} + \|u - u_{a,r}\|_{L^p(\Omega(a,r))} \\
&\leq \|u - u_{a,R}\|_{L^p(\Omega(a,R))} + \|u - u_{a,r}\|_{L^p(\Omega(a,r))} \\
&\leq [u]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)}\,R^{\frac{\lambda}{p}} + [u]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)}\,r^{\frac{\lambda}{p}} \\
&\leq 2\,[u]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)}\,R^{\frac{\lambda}{p}}.
\end{aligned}$$

Hence the lemma follows from the assumption that $|\Omega(a,r)| \geq A\,r^n$. □

Theorem 3.35. *If Ω is of type A then $L^{p,\lambda}(\Omega;\mathbb{R}^N) \cong \mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)$ for $0 \leq \lambda < n$.*

Proof. We only need to show $\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N) \subset L^{p,\lambda}(\Omega;\mathbb{R}^N)$. Let $u \in \mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)$. Given any $a \in \Omega$, $0 < \rho < \mathrm{diam}\Omega$, we have

$$\begin{aligned}
\|u\|_{L^p(\Omega(a,\rho))} &\leq \|u - u_{a,\rho}\|_{L^p(\Omega(a,\rho))} + \|u_{a,\rho}\|_{L^p(\Omega(a,\rho))} \\
&\leq [u]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)}\,\rho^{\frac{\lambda}{p}} + C\,|u_{a,\rho}|\,\rho^{\frac{n}{p}}.
\end{aligned}$$

We choose an integer k large enough so that $\Omega(a, 2^k\rho) = \Omega$. By Lemma 3.34, we have

$$\begin{aligned}
|u_{a,\rho}| &\leq |u_{a,2^k\rho}| + \sum_{j=0}^{k-1} |u_{a,2^{j+1}\rho} - u_{a,2^j\rho}| \\
&\leq |u_\Omega| + \sum_{j=0}^{k-1} 2\,A^{-\frac{1}{p}}\,(2^{j+1}\rho)^{\frac{\lambda}{p}}\,(2^j\rho)^{-\frac{n}{p}} \cdot [u]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)} \\
&\leq |u_\Omega| + C\,\rho^{\frac{\lambda-n}{p}}\,[u]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)} \cdot \sum_{j=0}^{k} 2^{j(\lambda-n)/p} \\
&\leq |u_\Omega| + C\,[u]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)}\,\rho^{\frac{\lambda-n}{p}},
\end{aligned}$$

where u_Ω is the average of u on Ω and therefore $|u_\Omega| \leq C(\Omega)\,\|u\|_{L^p(\Omega;\mathbb{R}^N)}$. Combining these estimates, we deduce

$$\|u\|_{L^p(\Omega(a,\rho))} \leq C\,[u]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)}\,\rho^{\frac{\lambda}{p}} + C\,\|u\|_{L^p(\Omega;\mathbb{R}^N)}\,\rho^{\frac{n}{p}}$$

and, by dividing both sides by $\rho^{\frac{\lambda}{p}}$ and noting $\lambda < n$,

$$\rho^{-\frac{\lambda}{p}}\,\|u\|_{L^p(\Omega(a,\rho))} \leq C\,[u]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)} + C(\Omega)\,\|u\|_{L^p(\Omega;\mathbb{R}^N)}.$$

This proves

$$\|u\|_{L^{p,\lambda}(\Omega;\mathbb{R}^N)} \le C(\Omega)\, \|u\|_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)}.$$

□

Remark 3.19. Note that $\mathcal{L}^{p,n}(\Omega;\mathbb{R}^N) \not\cong L^{p,n}(\Omega;\mathbb{R}^N) \cong L^\infty(\Omega;\mathbb{R}^N)$. For example, let $p=\lambda=1$, $n=N=1$ and $\Omega=(0,1)$. Then $u=\ln x$ is in $\mathcal{L}^{1,1}(0,1)$ but not in $L^{1,1}(0,1)\cong L^\infty(0,1)$. In fact, $\mathcal{L}^{p,n}(\Omega;\mathbb{R}^N)\cong BMO(\Omega;\mathbb{R}^N)$, which is called the *John-Nirenberg space.*

Theorem 3.36. (Campanato '63) *Let Ω be of type A. Then for $n<\lambda\le n+p$,*

$$\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)\cong C^{0,\alpha}(\bar\Omega;\mathbb{R}^N), \quad \alpha=\frac{\lambda-n}{p},$$

whereas for $\lambda>n+p$ we have $\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)=\{constants\}$.

Proof. 1. Assume $\lambda>n$ and $v\in\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)$. For any $x\in\Omega$ and $R>0$ we define

$$\tilde v(x)=\lim_{k\to\infty} v_{x,\frac{R}{2^k}}.$$

We claim $\tilde v$ is well-defined and independent of $R>0$. We first show the limit defining $\tilde v(x)$ exists. We need to show the sequence $\{v_{x,\frac{R}{2^k}}\}$ is Cauchy. For $h>k$ we have, by Lemma 3.34,

$$\begin{aligned}|v_{x,\frac{R}{2^h}}-v_{x,\frac{R}{2^k}}| &\le \sum_{j=k}^{h-1}|v_{x,\frac{R}{2^j}}-v_{x,\frac{R}{2^{j+1}}}|\\ &\le 2\,A^{-\frac1p}\,[v]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)}\,R^{\frac{\lambda-n}{p}}\cdot\sum_{j=k}^{h-1}2^{\frac{j(n-\lambda)}{p}},\end{aligned}$$

which, since $\lambda>n$, tends to zero if $k,\,h\to\infty$. Therefore $\{v_{x,\frac{R}{2^k}}\}$ is Cauchy and the limit $\tilde v(x)$ exists. Also in the inequality above, if $k=0$ and $h\to\infty$ we also deduce

$$|v_{x,R}-\tilde v(x)|\le C\,[v]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)}\cdot R^{\frac{\lambda-n}{p}}. \tag{3.67}$$

We now prove $\tilde v(x)$ is independent of $R>0$. This follows easily since by Lemma 3.34

$$\lim_{k\to\infty}|v_{x,\frac{R}{2^k}}-v_{x,\frac{r}{2^k}}|=0.$$

2. By Lebesgue's differentiation theorem, we also have $\tilde v(x)=v(x)$ for almost every $x\in\Omega$. Therefore $\tilde v=v$ in $\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)$. We claim $\tilde v\in C^{0,\alpha}(\bar\Omega;\mathbb{R}^N)$, where $\alpha=\frac{\lambda-n}{p}$. To show this, let $x,\,y\in\Omega$ and $x\ne y$. Let $R=|x-y|$. By (3.67) it follows that

$$\begin{aligned}|\tilde v(x)-\tilde v(y)| &\le |\tilde v(x)-v_{x,2R}|+|\tilde v(y)-v_{y,2R}|+|v_{x,2R}-v_{y,2R}|\\ &\le C\,[v]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)}\cdot R^\alpha+|v_{x,2R}-v_{y,2R}|.\end{aligned}$$

We need to estimate $|v_{x,2R}-v_{y,2R}|$. To this end, let $S=\Omega(x,2R)\cap\Omega(y,2R)$. Then $\Omega(x,R)\subset S$ and hence

$$|S|\ge|\Omega(x,R)|\ge A\,R^n.$$

On the other hand, we have

$$\begin{aligned}|S|^{\frac1p}\cdot|v_{x,2R}-v_{y,2R}| &= \|v_{x,2R}-v_{y,2R}\|_{L^p(S)}\\ &\le \|v_{x,2R}-v\|_{L^p(S)}+\|v_{y,2R}-v\|_{L^p(S)}\\ &\le \|v_{x,2R}-v\|_{L^p(\Omega(x,2R))}+\|v_{y,2R}-v\|_{L^p(\Omega(y,2R))}\\ &\le 2\,[v]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)}\cdot(2R)^{\lambda/p}.\end{aligned}$$

Combining the above two estimates we have

$$|v_{x,2R} - v_{y,2R}| \le C\,[v]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)} \cdot R^{\frac{\lambda-n}{p}} = C\,[v]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)} \cdot R^{\alpha}$$

and hence

$$|\tilde v(x) - \tilde v(y)| \le C\,[v]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)} \cdot |x-y|^{\alpha}.$$

This shows

$$[\tilde v]_{C^{0,\alpha}(\bar\Omega;\mathbb{R}^N)} \le C\,[v]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)}.$$

Finally, observe that, by (3.67) with $R = \operatorname{diam}\Omega$,

$$\begin{aligned}
\|\tilde v\|_{L^\infty(\Omega;\mathbb{R}^N)} &\le |v_\Omega| + C\,[v]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)} \cdot R^{\alpha}\\
&\le C(\Omega)\,\|v\|_{L^p(\Omega;\mathbb{R}^N)} + C(\Omega)\,[v]_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)}\\
&= C(\Omega)\,\|v\|_{\mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)}.
\end{aligned}$$

3. We have thus proved that if $\lambda > n$ then every $v \in \mathcal{L}^{p,\lambda}(\Omega;\mathbb{R}^N)$ has a representation $\tilde v$ which belongs to $C^{0,\alpha}(\bar\Omega;\mathbb{R}^N)$ with $\alpha = (\lambda - n)/p$. If $\lambda > n + p$ then $\alpha > 1$ and any $u \in C^{0,\alpha}(\bar\Omega;\mathbb{R}^N)$ must be a constant (why?). The proof of Campanato's theorem is complete. □

In order to use the Campanato spaces for elliptic systems, we also need some local version of these spaces. To disperse some technicalities, we prove the following lemma.

Lemma 3.37. *Let $p = 1, 2$ and $u \in L^p_{loc}(\Omega;\mathbb{R}^N)$. Then the map $E \mapsto \int_E |u - u_E|^p$ is nondecreasing in subsets $E \subset\subset \Omega$.*

Proof. We prove the case $p = 2$ first. Let $E \subset F \subset\subset \Omega$. Then

$$\begin{aligned}
\int_E |u-u_E|^2 &= \int_E |u - u_F + u_F - u_E|^2\\
&= \int_E |u-u_F|^2 + 2\,(u_F - u_E)\cdot \int_E (u - u_F) + |E|\cdot|u_F - u_E|^2\\
&= \int_E |u-u_F|^2 - |E|\cdot|u_F-u_E|^2\\
&\le \int_F |u-u_F|^2.
\end{aligned}$$

We now prove the case $p = 1$. Note that

$$\begin{aligned}
\int_E |u-u_E| &= \int_E |u-u_F+u_F-u_E|\\
&\le \int_E |u-u_F| + \int_E |u_F - u_E|\\
&= \int_F |u-u_F| - \int_{F\setminus E} |u-u_F| + |E|\cdot|u_F-u_E|.
\end{aligned}$$

Thus we need to prove

$$|E|\cdot|u_F - u_E| \le \int_{F\setminus E} |u - u_F|. \tag{3.68}$$

Note that, by Jensen's inequality,

$$\begin{aligned}
\fint_{F\setminus E}|u-u_F| &\geq \left|\fint_{F\setminus E}(u-u_F)\right| \\
&= \frac{1}{|F\setminus E|}\left||F\setminus E|\cdot u_F - \int_{F\setminus E}u\right| \\
&= \frac{1}{|F\setminus E|}\left||F|\cdot u_F - |E|\cdot u_F - \int_F u + \int_E u\right| \\
&= \frac{1}{|F\setminus E|}\big||F|\cdot u_F - |E|\cdot u_F - |F|\cdot u_F + |E|\cdot u_E\big| \\
&= \frac{|E|}{|F\setminus E|}\,|u_F-u_E|,
\end{aligned}$$

and hence (3.68) follows. □

Theorem 3.38. *Let $p = 1,\ 2$ and $u \in L^p_{loc}(\Omega;\mathbb{R}^N)$. Assume there exists a constant $C_u > 0$ and $\alpha > 0$ such that*

$$\int_{B_\rho}|u-u_{B_\rho}|^p\,dx \leq C_u\,\rho^\lambda$$

holds for all balls $B_\rho \subset\subset \Omega$. Then for any subdomain $\Omega' \subset\subset \Omega$ we have $u \in \mathcal{L}^{p,\lambda}(\Omega';\mathbb{R}^N)$ and moreover

$$\|u\|_{\mathcal{L}^{p,\lambda}(\Omega';\mathbb{R}^N)} \leq C(\Omega')\,[C_u^{1/p} + \|u\|_{L^p(\Omega';\mathbb{R}^N)}].$$

Proof. Let $\Omega' \subset\subset \Omega$ be given. We will show $u \in \mathcal{L}^{p,\lambda}(\Omega';\mathbb{R}^N)$. Let $d = \operatorname{dist}(\Omega';\partial\Omega)$. Given any $a \in \Omega'$ and $0 < \rho < \operatorname{diam}(\Omega')$. If $\rho < \operatorname{dist}(a;\partial\Omega)$ we have by the previous lemma,

$$\int_{\Omega'(a,\rho)}|u-u_{\Omega'(a,\rho)}|^p\,dx \leq \int_{B_\rho(a)}|u-u_{B_\rho(a)}|^p\,dx \leq C_u\,\rho^\lambda.$$

If $\rho \geq \operatorname{dist}(a;\partial\Omega)$, then $\rho \geq d > 0$ and hence

$$\int_{\Omega'(a,\rho)}|u-u_{\Omega'(a,\rho)}|^p\,dx \leq 2^p\int_{\Omega'(a,\rho)}|u|^p\,dx \leq \frac{2^p\,\|u\|^p_{L^p(\Omega';\mathbb{R}^n)}}{d^\lambda}\,\rho^\lambda.$$

Therefore, for all $a \in \Omega'$, $0 < \rho < \operatorname{diam}(\Omega')$ it follows that

$$\int_{\Omega'(a,\rho)}|u-u_{\Omega'(a,\rho)}|^p \leq \left[C_u + \frac{2^p\,\|u\|^p_{L^p(\Omega';\mathbb{R}^n)}}{d^\lambda}\right]\rho^\lambda,$$

and hence by definition $u \in \mathcal{L}^{p,\lambda}(\Omega';\mathbb{R}^N)$ and moreover

$$[u]_{\mathcal{L}^{p,\lambda}(\Omega';\mathbb{R}^N)} \leq C(\Omega')\,[C_u^{1/p} + \|u\|_{L^p(\Omega';\mathbb{R}^N)}].$$

The proof is complete. □

Theorem 3.39. (Morrey) *Let $u \in W^{1,p}_{loc}(\Omega;\mathbb{R}^N)$. Suppose for some $\beta > 0$ we have*

$$\int_{B_\rho}|Du|^p\,dx \leq C\,\rho^{n-p+\beta}, \quad \forall\, B_\rho \subset\subset \Omega.$$

Then for any $\Omega' \subset\subset \Omega$ of type A, we have $u \in C^{0,\frac{\beta}{p}}(\bar{\Omega}';\mathbb{R}^N)$.

Proof. Using the Poincaré type inequality

$$\int_{B_R} |u - u_{B_R}|\, dx \le C_n\, R \int_{B_R} |Du|\, dx, \tag{3.69}$$

we have for all balls $B_\rho \subset\subset \Omega$,

$$\begin{aligned}
\int_{B_\rho} |u - u_{B_\rho}|\, dx &\le C_n\, \rho \int_{B_\rho} |Du|\, dx \\
&\le C_n\, \rho\, \|Du\|_{L^p(B_\rho;\mathbb{M}^{N\times n})} \cdot |B_\rho|^{1-\frac{1}{p}} \\
&\le C\, \rho \cdot \rho^{\frac{n-p+\beta}{p}} \cdot \rho^{n(1-\frac{1}{p})} \\
&= C\, \rho^{n+\frac{\beta}{p}}.
\end{aligned}$$

Therefore, by Theorem 3.38, $u \in \mathcal{L}^{1,n+\frac{\beta}{p}}(\Omega';\mathbb{R}^N) \cong C^{0,\frac{\beta}{p}}(\bar{\Omega}';\mathbb{R}^N)$. ☐

When $\beta = 0$ Morrey's theorem has to be replaced by the *John-Nirenberg* estimate; see G-T, P. 166, Theorem 7.21.

Theorem 3.40. (John-Nirenberg) *Let $u \in W^{1,1}(\Omega;\mathbb{R}^N)$ where Ω is convex. Suppsose there exists a constant K such that*

$$\int_{\Omega(a,R)} |Du|\, dx \le K\, R^{n-1} \quad \forall\, a \in \Omega,\ R < \text{diam}\Omega. \tag{3.70}$$

Then there exist positive constants σ_0 and C depending only on n such that

$$\int_{\Omega} \exp\left(\frac{\sigma}{K}\, |u - u_\Omega|\right)\, dx \le C\, (\text{diam}\Omega)^n, \tag{3.71}$$

where $\sigma = \sigma_0\, |\Omega|\, (\text{diam}\Omega)^{-n}$.

Remark 3.20. The set of all functions $u \in W^{1,1}(\Omega;\mathbb{R}^N)$ satisfying (3.70) is the space $BMO(\Omega;\mathbb{R}^N)$ introduced by John and Nirenberg, and for Ω cubes or balls it follows that

$$BMO(\Omega;\mathbb{R}^N) \cong \mathcal{L}^{p,n}(\Omega;\mathbb{R}^N), \quad \forall\, p \ge 1.$$

For the proof of all these results and more on BMO-spaces, we refer to Gilbarg-Trudinger's book for a proof based on the Riesz potential, and Giaquinta's book on the Calderon-Zygmund cube decomposition.

3.6.4. Estimates for systems with constant coefficients. We consider systems with constant coefficients. Let $A = A_{ij}^{\alpha\beta}$ be constants satisfying hypothesis (H2). We first have some Campanato estimates for homogeneous systems.

Theorem 3.41. *Let $u \in H^1_{loc}(\Omega;\mathbb{R}^N)$ be a weak solution of*

$$D_\alpha(A_{ij}^{\alpha\beta}\, D_\beta u^j) = 0, \quad i = 1, 2, ..., N. \tag{3.72}$$

Then there exists a constant c depending on $A_{ij}^{\alpha\beta}$ such that for any concentric balls $B_\rho \subset\subset B_R \subset\subset \Omega$,

$$\int_{B_\rho} |u|^2\, dx \le c \cdot (\frac{\rho}{R})^n \int_{B_R} |u|^2\, dx, \tag{3.73}$$

$$\int_{B_\rho} |u - u_{B_\rho}|^2\, dx \le c \cdot (\frac{\rho}{R})^{n+2} \int_{B_R} |u - u_{B_R}|^2\, dx. \tag{3.74}$$

Proof. We do scaling first. Let $B_R = B_R(a)$, where $a \in \Omega$. Define

$$v(y) = u(a + Ry),$$

where $y \in D \equiv \{y \in \mathbb{R}^n \,|\, a + Ry \in \Omega\}$, which includes $\bar{B}_1(0)$ in the y-space. Note that $v \in H^1_{loc}(D;\mathbb{R}^N)$ is a weak solution of

$$D_{y_\alpha}(A^{\alpha\beta}_{ij} D_{y_\beta} v^j) = 0.$$

Then the Caccioppoli estimates show that

$$\|v\|_{W^{k,2}(B_{1/2}(0);\mathbb{R}^N)} \le C(k)\, \|v\|_{L^2(B_1(0);\mathbb{R}^N)}, \quad \forall\, k = 1, 2, ...$$

and hence for all $0 < t \le 1/2$ it follows that

$$\begin{aligned} \int_{B_t(0)} |v|^2\, dy &\le c(n)\, t^n \sup_{y \in B_{1/2}(0)} |v(y)|^2 \\ &\le c(n)\, t^n\, \|v\|^2_{W^{k,2}(B_{1/2}(0);\mathbb{R}^N)} \\ &\le c(n,k)\, t^n\, \|v\|^2_{L^2(B_1(0);\mathbb{R}^N)}, \end{aligned}$$

where we have chosen integer $k > n/2$ and used the Sobolev embedding $W^{k,2}(B_{1/2}(0);\mathbb{R}^N) \hookrightarrow C^{0,\alpha}(B_{1/2}(0);\mathbb{R}^N)$ for some $0 < \alpha < 1$. Now if $t \ge 1/2$ we easily have

$$\int_{B_t(0)} |v|^2\, dy \le 2^n\, t^n \int_{B_1(0)} |v|^2\, dy.$$

Therefore we have proved

$$\int_{B_t(0)} |v|^2\, dy \le C(n)\, t^n \int_{B_1(0)} |v|^2\, dy, \quad \forall\, 0 < t < 1.$$

Rescaling back to $u(x)$ and letting $\rho = tR$ we have

$$\int_{B_\rho(a)} |u|^2\, dx \le C(n)\, (\frac{\rho}{R})^n \cdot \int_{B_R(a)} |u|^2\, dx, \quad \forall\, \rho < R < \operatorname{dist}(a;\partial\Omega);$$

this proves (3.73). Note that Du is also a weak solution of (3.72); therefore, by (3.73) it follows that

$$\int_{B_\rho(a)} |Du|^2\, dx \le C(n)\, (\frac{\rho}{R})^n \cdot \int_{B_R(a)} |Du|^2\, dx, \quad \forall\, \rho < R < \operatorname{dist}(a;\partial\Omega).$$

Suppose $0 < \rho < R/2$. Then we use the Poincaré inequality, the previous estimate and the Caccioppoli inequality to obtain

$$\begin{aligned} \int_{B_\rho} |u - u_{B_\rho}|^2\, dx &\le c(n)\, \rho^2 \cdot \int_{B_\rho} |Du|^2\, dx \\ &\le C(n)\, \rho^2\, (\frac{\rho}{R})^n \cdot \int_{B_{R/2}} |Du|^2\, dx \\ &\le C(n)\, (\frac{\rho}{R})^{n+2} \cdot \int_{B_R} |u - u_{B_R}|^2\, dx. \end{aligned}$$

Now if $\rho \geq R/2$ we easily have

$$\begin{aligned} \int_{B_\rho} |u-u_{B_\rho}|^2\,dx &= \int_{B_\rho} |u-u_{B_R}|^2\,dx - |B_\rho|\cdot|u_{B_\rho}-u_{B_R}|^2 \\ &\leq \int_{B_R} |u-u_{B_R}|^2\,dx \\ &\leq 2^{n+2}\,(\frac{\rho}{R})^{n+2}\cdot\int_{B_R}|u-u_{B_R}|^2\,dx. \end{aligned}$$

Therefore, for all $0<\rho<R<\operatorname{dist}(a;\partial\Omega)$,

$$\int_{B_\rho} |u-u_{B_\rho}|^2\,dx \leq C(n)\,(\frac{\rho}{R})^{n+2}\cdot\int_{B_R}|u(x)-u_{B_R}|^2\,dx.$$

The proof of both (3.73) and (3.74) is now complete. □

In (3.73) and (3.74), if we let $R\to\operatorname{dist}(a;\partial\Omega)$, we see that both estimates also hold for all balls $B_\rho\subset\subset B_R\subset\Omega$. We state this fact as follows.

Corollary 3.42. *Both estimates* (3.73) *and* (3.74) *hold for all balls* $B_\rho\subset\subset B_R\subset\Omega$.

In the following, we consider the nonhomogeneous elliptic systems with constant coefficients:

$$D_\alpha(A_{ij}^{\alpha\beta}\,D_\beta u^j) = D_\alpha f_\alpha^i,\quad i=1,2,\cdots. \tag{3.75}$$

Theorem 3.43. *Let* $A_{ij}^{\alpha\beta}$ *satisfy hypothesis* $(H2)$ *and* $u\in W^{1,2}_{loc}(\Omega;\mathbb{R}^N)$ *be a weak solution of* (3.75)*. Suppose* $f\in\mathcal{L}^{2,\lambda}_{loc}(\Omega;\mathbb{M}^{N\times n})$ *and* $0\leq\lambda<n+2$*. Then* $Du\in\mathcal{L}^{2,\lambda}_{loc}(\Omega;\mathbb{M}^{N\times n})$.

Corollary 3.44. *Under the same assumptions, if* $f\in C^{0,\mu}_{loc}(\Omega;\mathbb{M}^{N\times n})$ *and* $0<\mu<1$ *then* $Du\in C^{0,\mu}_{loc}(\Omega;\mathbb{M}^{N\times n})$.

Proof of Theorem 3.43. Let $\Omega'\subset\subset\Omega''\subset\subset\Omega$. Let $a\in\Omega'$ and $B_R(a)=B_R\subset\Omega''$. We write $u=v+w=v+(u-v)$, where $v\in H^1(B_R;\mathbb{R}^N)$ is the solution of the Dirichlet problem

$$\begin{cases} \operatorname{div}(A\,Dv)=0 & \text{in } B_R, \\ v|_{\partial B_R}=u. \end{cases}$$

The existence of solution v follows by the Lax-Milgram theorem. We now by Corollary 3.42 have for all $\rho<R$

$$\int_{B_\rho}|Dv-(Dv)_{B_\rho}|^2\,dx \leq c\cdot(\frac{\rho}{R})^{n+2}\int_{B_R}|Dv-(Dv)_{B_R}|^2\,dx. \tag{3.76}$$

From this we have

$$\begin{aligned} &\int_{B_\rho}|Du-(Du)_{B_\rho}|^2\,dx \\ =\ &\int_{B_\rho}|Dv+Dw-(Dv)_{B_\rho}-(Dw)_{B_\rho}|^2\,dx \\ \leq\ &C\cdot(\frac{\rho}{R})^{n+2}\int_{B_R}|Dv-(Dv)_{B_R}|^2\,dx+\int_{B_\rho}|Dw-(Dw)_{B_\rho}|^2\,dx \\ \leq\ &C_1\cdot(\frac{\rho}{R})^{n+2}\int_{B_R}|Du-(Du)_{B_R}|^2\,dx+C_2\int_{B_R}|Du-Dv|^2\,dx. \end{aligned}$$

Since $u - v \in W_0^{1,2}(B_R;\mathbb{R}^N)$, we use the Legendre-Hadamard condition to have

$$\begin{aligned}\nu \int_{B_R} |Du - Dv|^2\,dx &\leq \int_{B_R} A\,D(u-v)\cdot D(u-v)\,dx \\ &= \int_{B_R} (f - f_{B_R})\cdot D(u-v)\,dx \\ &\leq \frac{\nu}{2}\int_{B_R} |Du - Dv|^2\,dx + C_\nu \int_{B_R} |f - f_{B_R}|^2\,dx\end{aligned}$$

and hence

$$\int_{B_R} |Du - Dv|^2\,dx \leq C_\nu \int_{B_R} |f - f_{B_R}|^2\,dx \leq C_\nu\,[f]^2_{\mathcal{L}^{2,\lambda}(\Omega'';\mathbb{M}^{N\times n})}\cdot R^\lambda.$$

Combining what we proved above, we have

$$\begin{aligned}\int_{B_\rho} |Du - (Du)_{B_\rho}|^2\,dx &\leq C_1\cdot(\frac{\rho}{R})^{n+2}\int_{B_R} |Du - (Du)_{B_R}|^2\,dx \\ &+ C_3\,[f]^2_{\mathcal{L}^{2,\lambda}(\Omega'';\mathbb{M}^{N\times n})}\cdot R^\lambda.\end{aligned}$$

Let

$$\Phi(\rho) = \int_{B_\rho} |Du - (Du)_{B_\rho}|^2\,dx.$$

Using the Campanato lemma below, it follows that

$$\Phi(\rho) \leq C_4\left[\left(\frac{\rho}{R}\right)^\lambda \Phi(R) + [f]^2_{\mathcal{L}^{2,\lambda}(\Omega'';\mathbb{M}^{N\times n})}\cdot\rho^\lambda\right].$$

Now we have

$$\begin{aligned}\int_{\Omega'(a,\rho)} |Du - (Du)_{\Omega'(a,\rho)}|^2 &\leq \int_{\Omega'(a,\rho)} |Du - (Du)_{B_\rho(a)}|^2 \\ &\leq \int_{B_\rho} |Du - (Du)_{B_\rho}|^2 = \Phi(\rho) \\ &\leq C_5\,\rho^\lambda\left(\|Du\|^2_{L^2(\Omega'';\mathbb{M}^{N\times n})} + [f]^2_{\mathcal{L}^{2,\lambda}(\Omega'';\mathbb{M}^{N\times n})}\right).\end{aligned}$$

Therefore

$$[Du]_{\mathcal{L}^{2,\lambda}(\Omega';\mathbb{M}^{N\times n})} \leq C\left(\|Du\|_{L^2(\Omega'';\mathbb{M}^{N\times n})} + [f]_{\mathcal{L}^{2,\lambda}(\Omega'';\mathbb{M}^{N\times n})}\right).$$

The proof is complete. □

Theorem 3.45. (Campanato Lemma) *Let $\Phi(t)$ be a nonnegative nondecreasing function. Consider the inequality*

$$\Phi(\rho) \leq A\left[\left(\frac{\rho}{R}\right)^\alpha + \epsilon\right]\Phi(R) + B\,R^\beta \quad \forall\,\rho \leq R \leq R_0, \tag{3.77}$$

where A, B, α, β, ϵ are positive constants with $\alpha > \beta$. Then there exists $\epsilon_0 = \epsilon_0(A,\alpha,\beta)$ such that if (3.77) holds for some $0 \leq \epsilon \leq \epsilon_0$ then

$$\Phi(\rho) \leq C\left[\left(\frac{\rho}{R}\right)^\beta \Phi(R) + B\,\rho^\beta\right] \quad \forall\,\rho \leq R \leq R_0,$$

where C is a constant depending only on α, β, A.

Proof. For $0<\tau<1$ and $R\le R_0$, (3.77) is equivalent to

$$\Phi(\tau R)\le A\,\tau^\alpha\,(1+\epsilon\tau^{-\alpha})\,\Phi(R)+B\,R^\beta.$$

Let $\gamma\in(\beta,\,\alpha)$ be fixed and choose $\tau\in(0,\,1)$ so that $2A\tau^\alpha\le\tau^\gamma$. Let $\epsilon_0=\tau^\alpha$. Then, if (3.77) holds for some $0\le\epsilon\le\epsilon_0$, we have for every $R\le R_0$

$$\Phi(\tau R)\le\tau^\gamma\,\Phi(R)+B\,R^\beta$$

and therefore for all $k=1,2,\cdots$

$$\begin{aligned}\Phi(\tau^{k+1}R) &\le \tau^\gamma\,\Phi(\tau^k R)+B\,\tau^{k\beta}\,R^\beta\\ &\le \tau^{(k+1)\gamma}\,\Phi(R)+B\,\tau^{k\beta}\,R^\beta\sum_{j=0}^{k}\tau^{j(\gamma-\beta)}\\ &\le C\,\tau^{(k+1)\beta}(\Phi(R)+B\,R^\beta).\end{aligned}$$

Since $\Phi(t)$ is nondecreasing and $\tau^{k+2}R<\rho\le\tau^{k+1}R$ for some k, we have

$$\Phi(\rho)\le C\left(\frac{\rho}{R}\right)^\beta(\Phi(R)+B\,R^\beta)=C\,\left[\left(\frac{\rho}{R}\right)^\beta\Phi(R)+B\,\rho^\beta\right],$$

as desired. The proof is complete. □

3.6.5. Schauder estimates for systems with variable coefficients. We now study the local regularity of weak solutions of systems with variable coefficients. We first prove the regularity in the Morrey space $L^{2,\lambda}_{loc}(\Omega)$ for the gradient of the weak solutions.

Theorem 3.46. *Let $A^{\alpha\beta}_{ij}(x)$ satisfy the hypothesis $(H3)$ and $u\in W^{1,2}_{loc}(\Omega;\mathbb{R}^N)$ be a weak solution of system*

$$D_\alpha(A^{\alpha\beta}_{ij}(x)\,D_\beta u^j)=D_\alpha f^i_\alpha. \tag{3.78}$$

Suppose $f\in L^{2,\lambda}_{loc}(\Omega;\mathbb{M}^{N\times n})$ and $0\le\lambda<n$. Then $Du\in L^{2,\lambda}_{loc}(\Omega;\mathbb{M}^{N\times n})$.

Proof. Let $\Omega'\subset\subset\Omega''\subset\subset\Omega$. Let $a\in\Omega'$ and $B_R(a)=B_R\subset\Omega''$. Using the standard Korn's freezing coefficients device, u is a weak solution of system with constant coefficients

$$\operatorname{div}(A(a)\,Du)=\operatorname{div}F,\quad F=f+(A(a)-A(x))\,Du.$$

Let $v\in H^1(B_R;\mathbb{R}^N)$ be the solution of the Dirichlet problem

$$\begin{cases}\operatorname{div}(A(a)\,Dv)=0 & \text{in } B_R,\\ v|_{\partial B_R}=u.\end{cases}$$

Then, as before, using (3.73) instead of (3.74) we have

$$\begin{aligned}\int_{B_\rho}|Du|^2 &\le c\cdot(\frac{\rho}{R})^n\int_{B_R}|Du|^2+C\int_{B_R}|D(u-v)|^2\\ &\le c\cdot(\frac{\rho}{R})^n\int_{B_R}|Du|^2+C\int_{B_R}|F|^2\\ &\le c\cdot(\frac{\rho}{R})^n\int_{B_R}|Du|^2+C\int_{B_R}|f|^2+C\,\omega(R)\int_{B_R}|Du|^2\\ &\le c\,\left[(\frac{\rho}{R})^n+\omega(R)\right]\int_{B_R}|Du|^2+C\,\|f\|^2_{L^{2,\lambda}(\Omega'';\mathbb{M}^{N\times n})}\,R^\lambda,\end{aligned}$$

where $\omega(R)$ is the uniform modulus of continuity of $A(x)$:

$$\omega(R) = \sup_{|x-y|\le R} |A(x) - A(y)|.$$

We choose $R_0 > 0$ sufficiently small so that $\omega(R) < \epsilon_0$ for all $R < R_0$, where ϵ_0 is the constant appearing in the Campanato lemma above. Therefore,

$$\int_{B_\rho} |Du|^2\,dx \le C(\Omega', \Omega'') \left(\|Du\|^2_{L^2(\Omega'';\mathbb{M}^{N\times n})} + \|f\|^2_{L^{2,\lambda}(\Omega'';\mathbb{M}^{N\times n})} \right) \rho^\lambda.$$

This, by a local version similar to the Campanato space, we have

$$\|Du\|_{L^{2,\lambda}(\Omega';\mathbb{M}^{N\times n})} \le C(\Omega', \Omega'') \left(\|Du\|_{L^2(\Omega'';\mathbb{M}^{N\times n})} + \|f\|_{L^{2,\lambda}(\Omega'';\mathbb{M}^{N\times n})} \right),$$

which proves the theorem. □

We now study the regularity of the gradient of weak solutions in the Hölder spaces. This is done by proving the regularity of gradient in the Campanato space $\mathcal{L}^{2,n+2\mu}_{loc}(\Omega)$ for some $\mu \in (0, 1)$.

Theorem 3.47. *Let $A^{\alpha\beta}_{ij} \in C^{0,\mu}(\Omega)$ with some $0 < \mu < 1$ satisfy the hypothesis $(H3)$ and $u \in W^{1,2}_{loc}(\Omega;\mathbb{R}^N)$ be a weak solution of system*

$$D_\alpha(A^{\alpha\beta}_{ij}(x)\, D_\beta u^j) = D_\alpha f^i_\alpha. \tag{3.79}$$

Suppose $f \in C^{0,\mu}_{loc}(\Omega;\mathbb{M}^{N\times n})$. Then $Du \in C^{0,\mu}_{loc}(\Omega;\mathbb{M}^{N\times n})$.

Proof. Similarly as above, we have

$$\begin{aligned}
\int_{B_\rho} |Du - (Du)_{B_\rho}|^2 &\le c\cdot(\frac{\rho}{R})^{n+2} \int_{B_R} |Du - (Du)_{B_R}|^2 + C\int_{B_R} |F - f_{B_R}|^2 \\
&\le c\cdot(\frac{\rho}{R})^{n+2} \int_{B_R} |Du - (Du)_{B_R}|^2 + C\int_{B_R} |f - f_{B_R}|^2 \\
&+ C\,[A]^2_{C^{0,\mu}}\, R^{2\mu} \int_{B_R} |Du|^2 \\
&\le c\cdot(\frac{\rho}{R})^{n+2} \int_{B_R} |Du - (Du)_{B_R}|^2 + C\,[f]^2_{C^{0,\mu}(\Omega'')}\, R^{n+2\mu} \\
&+ C\,[A]^2_{C^{0,\mu}}\, R^{2\mu} \int_{B_R} |Du|^2.
\end{aligned}$$

Since by the previous theorem $Du \in L^{2,n-\epsilon}_{loc}(\Omega;\mathbb{M}^{N\times n})$ for all $\epsilon > 0$ we obtain

$$\int_{B_\rho} |Du - (Du)_{B_\rho}|^2 \le A\,(\frac{\rho}{R})^{n+2} \int_{B_R} |Du - (Du)_{B_R}|^2 + B\,R^{n+2\mu-\epsilon}.$$

Using Campanato's lemma, we have

$$\int_{B_\rho} |Du - (Du)_{B_\rho}|^2 \le C\,\rho^{n+2\mu-\epsilon}$$

and hence $Du \in \mathcal{L}^{2,n+2\mu-\epsilon}_{loc}(\Omega;\mathbb{M}^{N\times n})$ for all $\epsilon > 0$. This implies $Du \in C^{0,\beta}_{loc}(\Omega;\mathbb{M}^{N\times n})$ for $\beta = \mu - \frac{\epsilon}{2}$. In particular, Du is locally bounded. Therefore, again, using the above estimates, it follows that

$$\begin{aligned}\int_{B_\rho} |Du - (Du)_{B_\rho}|^2 &\le c\cdot(\frac{\rho}{R})^{n+2}\int_{B_R}|Du-(Du)_{B_R}|^2\\ &+ C\,[f]^2_{C^{0,\mu}(\Omega'')}\,R^{n+2\mu} + C\,[A]^2_{C^{0,\mu}}\,R^{2\mu}\int_{B_R}|Du|^2\\ &\le c\cdot(\frac{\rho}{R})^{n+2}\int_{B_R}|Du-(Du)_{B_R}|^2\\ &+ C\,[f]^2_{C^{0,\mu}(\Omega'')}\,R^{n+2\mu} + C\,[A]^2_{C^{0,\mu}}\,R^{n+2\mu}\end{aligned}$$

and using Campanato's lemma again we have $Du \in L^{2,n+2\mu}_{loc}(\Omega;\mathbb{M}^{N\times n})$ and hence $Du \in C^{0,\mu}_{loc}(\Omega;\mathbb{M}^{N\times n})$. □

Finally we remark that the following higher order regularity result can be easily deduced.

Theorem 3.48. *Let $k \ge 0$, $0<\mu<1$ and $A^{\alpha\beta}_{ij} \in C^{k,\mu}(\Omega)$ satisfy the hypothesis $(H3)$ and $u \in W^{1,2}_{loc}(\Omega;\mathbb{R}^N)$ be a weak solution of system*

$$D_\alpha(A^{\alpha\beta}_{ij}(x)\,D_\beta u^j) = D_\alpha f^i_\alpha. \tag{3.80}$$

Suppose $f \in C^{k,\mu}_{loc}(\Omega;\mathbb{M}^{N\times n})$. Then $u \in C^{k+1,\mu}_{loc}(\Omega;\mathbb{R}^N)$.

3.6.6. Systems in non-divergence form and boundary estimates. In this section, we show that the Campanato estimates can also be proved for systems that are not in divergence form and also the global estimates are valid if the boundary $\partial\Omega$ of the domain $\Omega\subset\mathbb{R}^n$ satisfies certain smoothness condition.

We first prove the interior estimates for systems in the following form:

$$A^{\alpha\beta}_{ij}(x)\,D_{\alpha\beta}u^j = f^i;\quad i=1,2,\cdots,N.$$

By a weak solution u to this system we mean a function $u \in W^{2,2}_{loc}(\Omega;\mathbb{R}^N)$ such that the system is satisfied almost everywhere in Ω.

Theorem 3.49. *Let $A^{\alpha\beta}_{ij}$, $f^i \in C^{0,\mu}_{loc}(\Omega)$ and $0<\mu<1$. If $u \in W^{2,2}_{loc}(\Omega;\mathbb{R}^N)$ is a weak solution to the system above, then $u \in C^{2,\mu}_{loc}(\Omega;\mathbb{R}^N)$.*

We now consider the regularity up to the boundary. In what follows, we assume the boundary $\partial\Omega$ of the domain Ω is of $C^{1,\mu}$; that is, for any $x_0\in\partial\Omega$, there exist an open set $U\subset\mathbb{R}^n$ containing x_0 and a $C^{1,\mu}$-diffeomorphism $y = G\colon U\to\mathbb{R}^n$ such that

$$G(x_0)=0,\quad G(U\cap\Omega)=B_1^+=\{y\in\mathbb{R}^n\,|\,|y|<1,\ y_n>0\};$$

$$G(U\cap\partial\Omega)=\Gamma_1=\{y\in\mathbb{R}^n\,|\,|y|<1,\ y_n=0\}.$$

This G is called (locally) flattening the boundary. As in the scalar case studied above, we have the following global regularity theorems for linear systems.

Theorem 3.50. *Let $\partial\Omega$ be of $C^{1,\mu}$ with $0<\mu<1$ and $A^{\alpha\beta}_{ij}$, $f^i_\alpha \in C^{0,\mu}(\bar\Omega)$ and $g^j\in C^{1,\mu}(\bar\Omega)$. Let $A(x)$ satisfy the condition $(H3)$. If $u\in H^1(\Omega;\mathbb{R}^N)$ is a weak solution to the problem*

$$\operatorname{div}(A(x)\,Du) = \operatorname{div} f,\quad u|_{\partial\Omega}=g,$$

then $u\in C^{1,\mu}(\bar\Omega;\mathbb{R}^N)$.

Theorem 3.51. *Let* $\partial\Omega$ *be of* $C^{1,\mu}$ *with* $0 < \mu < 1$ *and* $A_{ij}^{\alpha\beta} \in C^{0,\mu}(\bar{\Omega})$ *satisfy* $(H3)$. *Assume* $f^i \in C^{0,\mu}(\bar{\Omega})$, $g^j \in C^{2,\mu}(\bar{\Omega})$. *If* $u \in W^{2,2}(\Omega; \mathbb{R}^N)$ *is a weak solution to the problem*

$$A_{ij}^{\alpha\beta}(x)\, D_{\alpha\beta} u^j = f^i; \quad u^j|_{\partial\Omega} = g^j,$$

then $u \in C^{2,\mu}(\bar{\Omega}; \mathbb{R}^N)$.

Chapter 4

Linear Evolution Equations

This chapter studies various linear partial differential equations that involve time. We call these equations linear **evolution equations**. We will study two major types of evolution equations of second-order: *parabolic and hyperbolic equations*. Two methods will be used: **Galerkin method** and **Semigroup method**.

4.1. Second-order Parabolic Equations

For this chapter we assume Ω to be an open bounded subset in $\mathbb{R}^n$ and set $\Omega_T = \Omega \times (0, T]$ for some fixed time $T > 0$; the **parabolic boundary** $\partial'\Omega_T$ of Ω_T is defined by $\partial'\Omega_T = \overline{\Omega_T} \setminus \Omega_T$.

We will study the **initial-boundary value problem**

$$\begin{cases} u_t + Lu = f & \text{in } \Omega_T, \\ u = 0 & \text{on } \partial\Omega \times [0, T], \\ u = g & \text{on } \Omega \times \{t = 0\}, \end{cases} \tag{4.1}$$

where $f: \Omega_T \to \mathbb{R}$ and $g: \Omega \to \mathbb{R}$ are given and $u: \overline{\Omega_T} \to \mathbb{R}$ is the unknown function, $u = u(x, t)$.

The operator Lu denotes for each time t a second-order partial differential operator, having either the divergence form

$$Lu = -\sum_{i,j=1}^{n} D_j(a_{ij}(x,t)D_i u) + \sum_{i=1}^{n} b_i(x,t)D_i u + c(x,t)u \tag{4.2}$$

or else the nondivergence form

$$Lu = -\sum_{i,j=1}^{n} a_{ij}(x,t)D_{ij} u + \sum_{i=1}^{n} b_i(x,t)D_i u + c(x,t)u, \tag{4.3}$$

for given coefficients a_{ij}, b_i, c $(i, j = 1, 2, \cdots, n)$.

Definition 4.1. We say the operator $\frac{\partial}{\partial t} + L$ is **(uniformly) parabolic** on Ω_T if there exists a constant $\theta > 0$ such that

$$\sum_{i,j=1}^{n} a_{ij}(x,t)\xi_i\xi_j \geq \theta|\xi|^2 \quad \text{for all } (x,t) \in \Omega_T \text{ and } \xi \in \mathbb{R}^n. \tag{4.4}$$

Note that for each fixed time $t \in [0,T]$ the operator Lu is uniformly elliptic in $x \in \Omega$.

4.1.1. Weak Solutions. We consider the case that Lu has the divergence form (4.2). We assume

$$a_{ij},\ b_i,\ c \in L^\infty(\Omega_T) \quad (i,j = 1,2,\cdots,n), \quad f \in L^2(\Omega_T),\ g \in L^2(\Omega).$$

We will also assume $a_{ij} = a_{ji}$ for $i,j = 1,2,\cdots,n$.

Introduce the time-dependent bilinear form

$$B[u,v;t] = \int_\Omega \Big(\sum_{i,j=1}^{n} a_{ij}(x,t)D_iuD_jv + \sum_{i=1}^{n} b_i(x,t)D_iuv + c(x,t)uv\Big)dx \tag{4.5}$$

for $u, v \in H_0^1(\Omega)$ and almost every $t \in [0,T]$.

Definition 4.2. A **weak solution** to Problem (4.1) is a function $u \in L^2(0,T;H_0^1(\Omega))$ with weak time-derivative $u' \in L^2(0,T;H^{-1}(\Omega))$ such that

(i) $\langle u'(t), v\rangle + B[u(t),v;t] = (f(t),v)$ for each $v \in H_0^1(\Omega)$ and a.e. time $t \in [0,T]$, and

(ii) $u(0) = g$. (Note that $u \in C([0,T];L^2(\Omega))$ and thus $u(0)$ is well-defined in $L^2(\Omega)$.)

Remark 4.3. (**Motivation for definition of weak solutions.**)

Suppose $u = u(x,t)$ is a smooth solution of (4.1). Then u defines a function, still denoted by u: $[0,T] \to H_0^1(\Omega)$ by $u(t)(x) = u(x,t)$. In other words, we consider u not as a function of (x,t) but rather as a function of t into the space $H_0^1(\Omega)$. We also consider $f\colon [0,T] \to L^2(\Omega)$ in terms of $f(t)(x) = f(x,t)$.

Fix $v \in H_0^1(\Omega)$ and multiply the PDE by v and integrate by parts, and we find

$$(u_t(t), v) + B[u(t),v;t] = (f(t),v) \quad \forall\, t \in [0,T], \tag{4.6}$$

where the pairing $(\ ,\ )$ is the inner product in $L^2(\Omega)$. Note that the PDE can be written as $u_t = g^0 + \sum_{i=1}^n g^i_{x_i}$, with

$$g^0 = f - \sum_{i=1}^{n} b_iD_iu - cu, \quad g^i = \sum_{i=1}^{n} a_{ij}D_iu \quad (i = 1,2,\cdots,n).$$

Hence, with $G = (g^0, g^1, \cdots, g^n) \in L^2(\Omega;\mathbb{R}^{n+1})$,

$$\|u_t\|_{H^{-1}(\Omega)} \leq \|G\|_{L^2(\Omega)} \leq C(\|u\|_{H^1(\Omega)} + \|f\|_{L^2(\Omega)}) \quad a.e.\ t \in [0,T]. \tag{4.7}$$

This estimate suggests it is reasonable to look for weak solutions $u \in L^2(0,T;H_0^1(\Omega))$ with weak time-derivative $u'(t) \in H^{-1}(\Omega)$ for a.e. $t \in [0,T]$, which, by (4.7), also satisfies $u' \in L^2(0,T;H^{-1}(\Omega))$. In this case the first term in (4.6) should be reexpressed as $\langle u'(t), v\rangle$, as the pairing of $H^{-1}(\Omega)$ and $H_0^1(\Omega)$.

4.1.2. Galerkin Method – Existence of weak solutions. We intend to build a weak solution to (4.1) by constructing approximate solutions u_k that, for each $t \in [0,T]$, lie in a finite-dimensional space V_k of $H_0^1(\Omega)$ such that condition (i) in the definition above holds for all $v \in V_k$. Then we pass to the limit as $k \to \infty$. This is the so-called **Galerkin's method**.

Assume the functions $w_i = w_i(x)$ are smooth and

(4.8) $\{w_i\}_{i=1}^{\infty}$ forms an orthogonal basis of $H_0^1(\Omega)$ and an orthonormal basis of $L^2(\Omega)$.

(For instance, we could take $\{w_i\}$ to be the complete set of appropriately normalized eigenfunctions for $-\Delta$ in $H_0^1(\Omega)$.)

Fix now a positive integer k. Let V_k be the linear span of $\{w_1, \cdots, w_k\}$ and we look for a function $u_k \colon [0,T] \to V_k$ of the form

$$u_k(t) = \sum_{i=1}^{k} d_i(t) w_i, \tag{4.9}$$

where the coefficient functions $d_i(t)$ is selected so that

$$(u_k'(t), w_i) + B[u_k(t), w_i; t] = (f(t), w_i), \quad d_i(0) = (g, w_i), \tag{4.10}$$

for almost every $t \in [0,T]$ and $i = 1, 2, \cdots, k$.

Here and what follows, $(\,,\,)$ denotes the inner product in $L^2(\Omega)$, and $'$ denotes the time-derivative of a function (whenever it is well-defined in classical or weak sense).

Theorem 4.1. (Construction of approximate solutions) *For each $k = 1, 2, \cdots$ there exists a unique function u_k of the form* (4.9) *satisfying* (4.10).

Proof. Note that

$$(u_k'(t), w_i) = d_i'(t), \quad B[u_k(t), w_i; t] = \sum_{j=1}^{k} \alpha_i^j(t) d_j(t),$$

where $\alpha_i^j(t) = B[w_j, w_i; t]$ $(i, j = 1, 2, \cdots, k)$. Hence condition (4.10) becomes the initial value problem for the ODE system on $d(t) = (d_1(t), \cdots, d_k(t))$:

$$d_i'(t) + \sum_{j=1}^{k} \alpha_i^j(t) d_j(t) = f_i(t) \equiv (f(t), w_i), \quad d_i(0) = (g, w_i) \quad (i = 1, 2, \cdots, k).$$

Note that the coefficients α_i^j belong to $L^\infty(0,T)$ and $f_i \in L^2(0,T)$. The existence of a unique solution $d \in H^1(0,T) \subset C([0,T])$ is guaranteed by the (not so) standard existence theory for ODE (think of approximating α_i^j and f_i by smooth functions first and then pass to limits). $\square$

4.1.3. Energy Estimates.

Theorem 4.2. *Assume the uniform parabolicity condition. There exists a constant C, depending only on Ω, T, and the coefficients of L, such that, for all $k = 1, 2, \cdots$,*

$$\max_{t\in[0,T]} \|u_k(t)\|_{L^2(\Omega)} + \|u_k\|_{L^2(0,T;H_0^1(\Omega))} + \|u_k'\|_{L^2(0,T;H^{-1}(\Omega))} \le C(\|f\|_{L^2(\Omega_T)} + \|g\|_{L^2(\Omega)}). \tag{4.11}$$

Proof. 1. From the uniform parabolicity, as in the elliptic case, there exist constants $\beta > 0, \gamma \ge 0$ such that

$$\beta \|v\|_{H_0^1(\Omega)}^2 \le B[v, v; t] + \gamma \|v\|_{L^2(\Omega)}^2 \quad \forall\, v \in H_0^1(\Omega), \ \ a.e.\ t \in [0,T]. \tag{4.12}$$

Multiply (4.10) by $d_i(t)$ and sum for $i = 1, 2, \cdots, k$ to find

$$(u_k'(t), u_k(t)) + B[u_k(t), u_k(t); t] = (f(t), u_k(t)) \quad \forall\, a.e.\ t \in [0, T]. \tag{4.13}$$

Hence

$$\frac{d}{dt}\left(\|u_k(t)\|^2_{L^2(\Omega)}\right) + 2\beta\|u_k(t)\|^2_{H^1_0(\Omega)} \le C_1\|u_k(t)\|^2_{L^2(\Omega)} + C_2\|f(t)\|^2_{L^2(\Omega)} \tag{4.14}$$

for a.e. $t \in [0, T]$, and appropriate constants C_1, C_2.

2. Now write

$$\eta(t) = \|u_k(t)\|^2_{L^2(\Omega)}, \quad \xi(t) = \|f(t)\|^2_{L^2(\Omega)}.$$

Then

$$\eta'(t) \le C_1\eta(t) + C_2\xi(t) \quad \forall\, a.e.\ t \in [0, T].$$

Thus, by Gronwall's inequality,

$$\eta(t) \le e^{C_1 t}\left(\eta(0) + C_2\int_0^t \xi(s)\, ds\right) \quad (0 \le t \le T).$$

Since $\eta(0) = \|u_k(0)\|^2_{L^2(\Omega)} \le \|g\|^2_{L^2(\Omega)}$, we obtain the estimate

$$\max_{t\in[0,T]} \|u_k(t)\|^2_{L^2(\Omega)} \le C(\|g\|^2_{L^2(\Omega)} + \|f\|^2_{L^2(\Omega_T)}).$$

3. Integrating (4.14) over $t \in [0, T]$, we have

$$\|u_k\|^2_{L^2(0,T;H^1_0(\Omega))} = \int_0^T \|u_k(t)\|^2_{H^1_0(\Omega)}\, dt \le C(\|g\|^2_{L^2(\Omega)} + \|f\|^2_{L^2(\Omega_T)}).$$

4. Finally we need to estimate $\|u_k'\|_{L^2(0,T;H^{-1}(\Omega))}$. So, fix any $v \in H^1_0(\Omega)$, with $\|v\|_{H^1_0(\Omega)} \le 1$. We write $v = v^1 + v^2$, where $v^1 \in V_k$, and $(v^2, w_i) = 0$ for all $i = 1, 2, \cdots, k$. (That is, v^2 is in the L^2 orthogonal complement of V_k.) Since $\{w_i\}$ are orthogonal in $H^1_0(\Omega)$, we have

$$\|v^1\|_{H^1_0(\Omega)} \le \|v\|_{H^1_0(\Omega)} \le 1.$$

Using (4.10), we have

$$(u_k'(t), v^1) + B[u_k(t), v^1; t] = (f(t), v^1).$$

Then

$$\langle u_k'(t), v\rangle = (u_k'(t), v) = (u_k'(t), v^1) = (f(t), v^1) - B[u_k(t), v^1; t]$$

and consequently

$$|\langle u_k'(t), v\rangle| \le C(\|f(t)\|_{L^2(\Omega)} + \|u_k\|_{H^1_0(\Omega)}).$$

This implies

$$\|u_k'(t)\|^2_{H^{-1}(\Omega)} \le C(\|f(t)\|^2_{L^2(\Omega)} + \|u_k(t)\|^2_{H^1_0(\Omega)}) \quad \forall\, t \in [0, T].$$

Integrate over $t \in [0, T]$ to finally obtain

$$\|u_k'\|^2_{L^2(0,T;H^{-1}(\Omega))} = \int_0^T \|u_k'(t)\|^2_{H^{-1}(\Omega)} dt \le C(\|f\|^2_{L^2(\Omega_T)} + \|u_k\|^2_{L^2(0,T;H^1_0(\Omega))}),$$

which, combining with the estimate in Step 3, derives the desired estimate. □

4.1.4. Existence and Uniqueness.

Theorem 4.3. *There exists a unique weak solution to* (4.1).

Proof. (Existence.) 1. According to the energy estimate (4.11), we see that $\{u_k\}$ is bounded in $L^2(0,T;H_0^1(\Omega))$ and $\{u_k'\}$ is bounded in $L^2(0,T;H^{-1}(\Omega))$. Consequently there exists a subsequence $\{u_{k_m}\}$ of $\{u_k\}$ with $k_m \to \infty$ and functions $u \in L^2(0,T;H_0^1(\Omega))$, $w \in L^2(0,T;H^{-1}(\Omega))$ such that

$$\begin{cases} u_{k_m} \rightharpoonup u & \text{in } L^2(0,T;H_0^1(\Omega)), \\ u'_{k_m} \rightharpoonup w & \text{in } L^2(0,T;H^{-1}(\Omega)). \end{cases} \tag{4.15}$$

Moreover, u' exists in $L^2(0,T;H^{-1}(\Omega))$ and $u' = w$. From the regularity result, $u \in C([0,T];L^2(\Omega))$.

2. Fix any integer N and let $\psi \in C^1([0,T];H_0^1(\Omega))$ have the form

$$\psi(t) = \sum_{i=1}^{N} \zeta_i(t) w_i,$$

where $\zeta_i \in C^1([0,T];\mathbb{R})$. Let $k \geq N$, multiply (4.10) by ζ_i, sum $i = 1, 2, \cdots, N$ and then integrate over $t \in [0,T]$ to find

$$\int_0^T \left(\langle u_k'(t), \psi(t) \rangle + B[u_k(t), \psi(t); t] \right) dt = \int_0^T (f(t), \psi(t)) dt. \tag{4.16}$$

Now let $k = k_m \to \infty$ and we have

$$\int_0^T \left(\langle u'(t), \psi(t) \rangle + B[u(t), \psi(t); t] \right) dt = \int_0^T (f(t), \psi(t)) dt. \tag{4.17}$$

This equality then holds for all functions $\psi \in L^2(0,T;H_0^1(\Omega))$, as functions ζ of the given form are dense in this space. We then take $\psi(t) = \zeta(t)v$ with $\zeta \in L^2(0,T)$ and $v \in H_0^1(\Omega)$ in (4.17) to obtain

$$\int_0^T \zeta(t) \left(\langle u'(t), v \rangle + B[u(t), v; t] \right) dt = \int_0^T \zeta(t)(f(t), v) dt.$$

This holding for all $\zeta \in L^2(0,T)$ yields that

$$\langle u'(t), v \rangle + B[u(t), v; t] = (f(t), v) \quad \forall\, v \in H_0^1(\Omega), \ \textit{a.e. } t \in [0,T]. \tag{4.18}$$

3. We need to show the initial data $u(0) = g$. In (4.16), (4.17), take $\psi(t) = \zeta(t)v$ with $\zeta \in C^1[0,T]$, $\zeta(T) = 0$, $\zeta(0) = -1$ and $v \in H_0^1(\Omega)$. Note that, since $\psi'(t) = \zeta'(t)v$, we have

$$\int_0^T \langle u_k'(t), \psi(t) \rangle dt = (u_k(0), v) - \int_0^T \zeta'(t)(u_k(t), v)\, dt,$$

$(u_k(0), v) \to (g, v)$ as $k \to \infty$, and

$$\int_0^T \langle u'(t), \psi(t) \rangle dt = (u(0), v) - \int_0^T \zeta'(t)(u(t), v)\, dt.$$

Hence, in (4.16), let $k = k_m \to \infty$, we eventually obtain $(u(0), v) = (g, v)$, for all $v \in H_0^1(\Omega)$; hence $u(0) = g$.

(Uniqueness.) It suffices to prove that a weak solution u with $f = g = 0$ must be zero. To show this, set $v = u(t) \in H_0^1(\Omega)$ in (4.18) with $f = 0$ to have

$$\frac{d}{dt}\left(\frac{1}{2}\|u(t)\|^2_{L^2(\Omega)}\right) + B[u(t), u(t); t] = 0 \quad (a.e.\ t \in [0, T]).$$

Since the Gårding's inequality above implies $-B[u(t), u(t); t] \le \gamma \|u(t)\|^2_{L^2(\Omega)}$, it follows that

$$\frac{d}{dt}\left(\frac{1}{2}\|u(t)\|^2_{L^2(\Omega)}\right) \le \gamma \|u(t)\|^2_{L^2(\Omega)} \quad \forall\, a.e.\ t \in [0, T].$$

So Gronwall's inequality implies $\|u(t)\|_{L^2(\Omega)} = 0$ as $u(0) = 0$. Hence $u \equiv 0$. □

4.1.5. Regularity. We now discuss the regularity of weak solutions when the initial data and coefficients are more regular. Our eventual goal is to prove that the weak solution is smooth, as long as the coefficients and initial data and the domain are all smooth. This mirrors the regularity of elliptic equations.

Before we proceed, we prove the following useful result which is Problem 9 of Chapter 7 in Evans's book; the proof follows a paper by Brezis and Evans (*Arch. Rational Mech. Analysis 71 (1979), 1–13*).

Lemma 4.4. *If Lu is uniformly elliptic with smooth coefficients, then there exist constants $\beta > 0$, $\gamma \ge 0$ such that*

$$\beta \|u\|^2_{H^2(\Omega)} \le (Lu, -\Delta u) + \gamma \|u\|^2_{L^2(\Omega)} \quad \forall\ u \in H_0^1(\Omega) \cap H^2(\Omega). \tag{4.19}$$

Proof. 1. Given a function $u \in H^2(\Omega) \cap H_0^1(\Omega)$, we claim that there exists a sequence of functions u_m in $C^3(\bar{\Omega})$ vanishing on $\partial\Omega$ such that $\|u_m - u\|_{H^2(\Omega)} \to 0$; therefore, we only need to prove (4.19) for functions $u \in C^3(\bar{\Omega})$ vanishing on $\partial\Omega$. To prove this claim, let $f = \Delta u \in L^2(\Omega)$ and choose $f_m \in C^\infty(\bar{\Omega})$ so that $\|f_m - f\|_{L^2(\Omega)} \to 0$. Let $u_m \in H_0^1(\Omega)$ be the weak solution of $\Delta u_m = f_m$ in Ω. Then the global regularity theorem shows that $u_m \in H^k(\Omega)$ if $\partial\Omega$ is of C^k and $k \ge 2$. By the general Sobolev inequalities we know that if $k > \frac{n}{2}$ then $u_m \in C^{k-[\frac{n}{2}]-1,\gamma}(\bar{\Omega})$ for some $0 < \gamma < 1$. Hence if $\partial\Omega$ is of C^k with $k = 4 + [\frac{n}{2}]$, then $u_m \in C^3(\bar{\Omega})$. Clearly $u_m = 0$ on $\partial\Omega$. Finally, since $\Delta(u_m - u) = f_m - f$, by the global H^2-estimate, we have

$$\|u_m - u\|_{H^2(\Omega)} \le C\|f_m - f\|_{L^2(\Omega)} \to 0.$$

2. In the following, assume $u \in C^3(\bar{\Omega})$ and $u = 0$ on $\partial\Omega$. Then we can write Lu in non-divergence form (using the new smooth coefficients) as

$$Lu = -a_{ij}(x)u_{x_i x_j} + b_i(x)u_{x_i} + c(x)u(x).$$

Let $x^0 \in \partial\Omega$ and assume a smooth homeomorphism $y = \Phi(x)$ from small ball $B(x^0, r)$ to a domain D in y-space satisfies that $\Phi(x^0) = 0$ and $\Phi(B(x^0, r) \cap \Omega) = D \cap \{y_n > 0\}$ and $\Phi(B(x^0, r) \cap \partial\Omega) = D \cap \{y_n = 0\}$. Let $x = \Psi(y)$ be the inverse map of $y = \Phi(x)$. (This is the standard technique of locally flattening the boundary.) Let $\zeta \in C_c^\infty(B(x^0, r)$ and $0 \le \zeta \le 1$. Then

$$\int_\Omega \zeta(x)(Lu, -\Delta u)dx = \int_{D\cap\{y_n>0\}} \psi(y)(L^1 v, L^2 v)dy,$$

where $v(y) = u(\Psi(y))$, $\psi(y) = \zeta(\Psi(y))|I(y)|$ with $I(y)$ the Jacobian determinant of $x = \Psi(y)$, and L^1, L^2 are the new differential operators of $v(y)$ from Lu and Δu, both of the form

$$L^k v = -a^k_{ij}(y)v_{y_i y_j} + b^k_i(y)v_{y_i} + c^k(y)v(y) \quad (k = 1, 2).$$

The coefficients $c^k(y)$ are bounded (in fact, $c^2(y)\equiv 0$) and the coefficients $a^k_{ij}(y), b^k_i(y)$ are smooth functions on $D\cap\{y_n\ge 0\}$ and satisfy for a constant $\theta>0$, with $k=1,2$,

$$a^k_{ij}(y)=a^k_{ji}(y),$$
$$a^k_{ij}(y)\xi_i\xi_j\ge\theta|\xi|^2\quad (y\in D\cap\{y_n\ge 0\},\ \ \xi\in\mathbb{R}^n).$$

From this, an easy algebra proof shows that for all symmetric matrices $C\in\mathbb{M}^{n\times n}$

$$\sum_{i,j,k,l=1}^{n} a^1_{ij}(y)c_{ik}a^2_{kl}(y)c_{jl}\ge\theta^2\sum_{i,j=1}^{n}c_{ij}^2\quad (y\in D\cap\{y_n\ge 0\}). \tag{4.20}$$

We need to estimate $\int_{D\cap\{y_n>0\}}\psi(y)(L^1v,L^2v)dy$.

3. We write the integrand as

$$\psi(y)(L^1v,L^2v)=\psi a^1_{ij}v_{y_iy_j}a^2_{kl}v_{y_ky_l}+\psi R,$$

where R is the term of the form

$$R=\sum A_{ijk}v_{y_iy_j}v_{y_k}+\sum B_{ij}v_{y_i}v_{y_j}+\sum C_{ij}v_{y_iy_j}v+c^1c^2v^2,$$

with bounded coefficients A,B,C. The leading term can be written as

$$\begin{aligned}\psi a^1_{ij}v_{y_iy_j}a^2_{kl}v_{y_ky_l}&=\psi a^1_{ij}v_{y_iy_k}a^2_{kl}v_{y_jy_l}\\ &\quad+(\psi a^1_{ij}a^2_{kl})_{y_j}v_{y_iy_k}v_{y_l}-(\psi a^1_{ij}a^2_{kl})_{y_k}v_{y_iy_j}v_{y_l}\\ &\quad+(\psi a^1_{ij}a^2_{kl}v_{y_iy_j}v_{y_l})_{y_k}-(\psi a^1_{ij}a^2_{kl}v_{y_iy_k}v_{y_l})_{y_j}.\end{aligned}$$

Therefore, by (4.20), we have

$$\begin{aligned}\int_{D\cap\{y_n>0\}}\psi(y)(L^1v,L^2v)dy&\ge\theta^2\sum_{i,j=1}^{n}\int_{D\cap\{y_n>0\}}\psi(y)v^2_{y_iy_j}dy\\ &\quad+\sum_{i,j,k,l=1}^{n}\int_{\partial(D\cap\{y_n>0\})}\psi a^1_{ij}a^2_{kl}(v_{y_iy_j}v_{y_l}\nu_k-v_{y_iy_k}v_{y_l}\nu_j)\,dS\\ &\quad-\varepsilon\sum_{i,j=1}^{n}\int_{D\cap\{y_n>0\}}v^2_{y_iy_j}dy-C_\varepsilon\int_{D\cap\{y_n>0\}}(|\nabla v|^2+v^2)\,dy,\end{aligned} \tag{4.21}$$

where $\nu=(\nu_1,\cdots,\nu_n)$ is the outer unit normal on the boundary.

4. Since $\psi=0$ on $\partial(D\cap\{y_n>0\})\setminus(D\cap\{y_n=0\})$, we have the boundary integral

$$\sum_{i,j,k,l=1}^{n}\int_{\partial(D\cap\{y_n>0\})}\psi a^1_{ij}a^2_{kl}(v_{y_iy_j}v_{y_l}\nu_k-v_{y_iy_k}v_{y_l}\nu_j)\,dS$$
$$=-\sum_{i,j,l=1}^{n}\int_{D\cap\{y_n=0\}}\psi a^1_{ij}a^2_{nl}v_{y_iy_j}v_{y_l}\,dy'+\sum_{i,k,l=1}^{n}\int_{D\cap\{y_n=0\}}\psi a^1_{in}a^2_{kl}v_{y_iy_k}v_{y_l}\,dy'.$$

Now since $v(y',0)=0$ we have $v_{y_i}=0,\ v_{y_iy_j}=0$ for $1\le i,j\le n-1$. So

$$\sum_{i,j,l=1}^{n}\int_{D\cap\{y_n=0\}}\psi a^1_{ij}a^2_{nl}v_{y_iy_j}v_{y_l}\,dy'=\sum_{i,j=1}^{n}\int_{D\cap\{y_n=0\}}\psi a^1_{ij}a^2_{nn}v_{y_iy_j}v_{y_n}\,dy'$$
$$=\int_{D\cap\{y_n=0\}}\psi a^1_{nn}a^2_{nn}v_{y_ny_n}v_{y_n}\,dy'+2\sum_{i=1}^{n-1}\int_{D\cap\{y_n=0\}}\psi a^1_{in}a^2_{nn}v_{y_iy_n}v_{y_n}\,dy',$$

and

$$\sum_{i,k,l=1}^{n}\int_{D\cap\{y_n=0\}}\psi a^1_{in}a^2_{kl}v_{y_iy_k}v_{y_l}\,dy' = \sum_{i,k=1}^{n}\int_{D\cap\{y_n=0\}}\psi a^1_{in}a^2_{kn}v_{y_iy_k}v_{y_n}\,dy'$$

$$= \int_{D\cap\{y_n=0\}}\psi a^1_{nn}a^2_{nn}v_{y_ny_n}v_{y_n}\,dy' + \sum_{i=1}^{n-1}\int_{D\cap\{y_n=0\}}\psi a^1_{in}a^2_{nn}v_{y_iy_n}v_{y_n}\,dy'$$

$$+\sum_{k=1}^{n-1}\int_{D\cap\{y_n=0\}}\psi a^1_{nn}a^2_{kn}v_{y_ky_n}v_{y_n}\,dy'.$$

Therefore

$$\sum_{i,j,k,l=1}^{n}\int_{\partial(D\cap\{y_n>0\})}\psi a^1_{ij}a^2_{kl}(v_{y_iy_j}v_{y_l}\nu_k - v_{y_iy_k}v_{y_l}\nu_j)\,dS$$

$$= \sum_{i=1}^{n-1}\int_{D\cap\{y_n=0\}}\psi(a^1_{nn}a^2_{in} - \psi a^1_{in}a^2_{nn})v_{y_iy_n}v_{y_n}\,dy'$$

$$= \frac{1}{2}\sum_{i=1}^{n-1}\int_{D\cap\{y_n=0\}}\psi(a^1_{nn}a^2_{in} - \psi a^1_{in}a^2_{nn})(v^2_{y_n})_{y_i}\,dy'$$

$$= -\frac{1}{2}\sum_{i=1}^{n-1}\int_{D\cap\{y_n=0\}}[\psi(a^1_{nn}a^2_{in} - \psi a^1_{in}a^2_{nn})]_{y_i}v^2_{y_n}\,dy'.$$

Hence

$$\left|\sum_{i,j,k,l=1}^{n}\int_{\partial(D\cap\{y_n>0\})}\psi a^1_{ij}a^2_{kl}(v_{y_iy_j}v_{y_l}\nu_k - v_{y_iy_k}v_{y_l}\nu_j)\,dS\right|$$

$$= \left|\frac{1}{2}\sum_{i=1}^{n-1}\int_{D\cap\{y_n=0\}}[\psi(a^1_{nn}a^2_{in} - \psi a^1_{in}a^2_{nn})]_{y_i}v^2_{y_n}\,dy'\right| \le C\int_{\partial(D\cap\{y_n>0\})}|\nabla v|^2 dS$$

$$\le \varepsilon\sum_{i,j=1}^{n}\int_{D\cap\{y_n>0\}}v^2_{y_iy_j}dy + C_\varepsilon\int_{D\cap\{y_n>0\}}|\nabla v|^2 dy,$$

by the trace inequality: for $\varepsilon > 0$,

$$\|\nabla v\|_{L^2(\partial U)} \le \varepsilon\|D^2v\|_{L^2(U)} + C_\varepsilon\|\nabla v\|_{L^2(U)} \quad \forall\, v\in H^2(U).$$

(Prove this inequality!)

5. Putting all inequalities above together in (4.21), we eventually obtain

$$\int_{D\cap\{y_n>0\}}\psi(y)(L^1v, L^2v)dy \ge \theta^2\sum_{i,j=1}^{n}\int_{D\cap\{y_n>0\}}\psi(y)v^2_{y_iy_j}dy$$

$$-\,\varepsilon\sum_{i,j=1}^{n}\int_{D\cap\{y_n>0\}}v^2_{y_iy_j}dy - C_\varepsilon\int_{D\cap\{y_n>0\}}(|\nabla v|^2 + v^2)\,dy,$$

where $\varepsilon > 0$ is arbitrary and C_ε depends on ψ.

6. Switching back to the domain Ω, we have

$$\begin{aligned}\int_\Omega \zeta(x)(Lu,-\Delta u)\,dx \geq \theta^2 \sum_{i,j=1}^n \int_\Omega \zeta(x) u_{x_ix_j}^2 dx \\ - \varepsilon \sum_{i,j=1}^n \int_\Omega u_{x_ix_j}^2 dx - C_\varepsilon \int_\Omega (|\nabla u|^2+u^2)\,dx,\end{aligned} \tag{4.22}$$

where C_ε depends on ζ. Clearly this estimate also holds when $\zeta \in C_c^\infty(\Omega)$.

7. We now cover $\partial\Omega$ by finitely many balls $B_k := B(x^k, r_k)$ with $x^k \in \partial\Omega$, $k = 1, 2, \cdots, N$, that the local flattening of the boundary works. Let $B_{N+1} := \Omega \setminus \cup_{k=1}^N \bar{B}(x^k, r_k/2)$. We find a partition of unity $\sum_{k=1}^{N+1} \zeta_k = 1$ subordinate to $\{B_1, B_2, \cdots, B_{N+1}\}$ with $0 \leq \zeta_k \leq 1$ and $\operatorname{supp}\zeta_k \subset\subset B_k$. Then using (4.22) and a choice of small $\varepsilon > 0$ we deduce that

$$\int_\Omega (Lu,-\Delta u)\,dx \geq \frac{\theta^2}{2}\|D^2u\|_{L^2(\Omega)}^2 - C\int_\Omega (|\nabla u|^2+u^2)\,dx. \tag{4.23}$$

From this, the estimate (4.19) follows since $\|D^2u\|_{L^2(\Omega)}$ is an equivalent norm for $H^2(\Omega)\cap H_0^1(\Omega)$ in $H^2(\Omega)$ and

$$\|\nabla u\|_{L^2(\Omega)} \leq \varepsilon\|D^2u\|_{L^2(\Omega)} + C_\varepsilon\|u\|_{L^2(\Omega)} \quad \forall\, u \in H^2(\Omega)\cap H_0^1(\Omega),$$

which can be seen from one of the homework problem. □

Remark 4.4. Clearly, from the proof, the estimate (4.19) holds if we replace $-\Delta u$ by another uniformly elliptic operator Mu.

In the following we assume the coefficients a_{ij}, b_i, c are as smooth as we need on Ω_T. As usual, we always assume the uniform parabolicity.

Theorem 4.5. (Improved regularity) *(i) Assume $g \in H_0^1(\Omega)$, $f \in L^2(\Omega_T)$. Suppose u is the weak solution of* (4.1)*. Then*

$$u \in L^2(0,T;H^2(\Omega))\cap L^\infty(0,T;H_0^1(\Omega)), \;\; u' \in L^2(0,T;L^2(\Omega)),$$

with the estimate

$$\begin{aligned}\|u\|_{L^\infty(0,T;H_0^1(\Omega))} + \|u\|_{L^2(0,T;H^2(\Omega))} + \|u'\|_{L^2(0,T;L^2(\Omega))} \\ \leq C(\|g\|_{H_0^1(\Omega)} + \|f\|_{L^2(\Omega_T)}),\end{aligned} \tag{4.24}$$

where C depends only on Ω, T and the coefficients of L.

(ii) If, in addition, $g \in H^2(\Omega)$, $f' \in L^2(0,T;L^2(\Omega))$, then

$$u \in L^\infty(0,T;H^2(\Omega)), \;\; u' \in L^\infty(0,T;L^2(\Omega))\cap L^2(0,T;H_0^1(\Omega)), \;\; u'' \in L^2(0,T;H^{-1}(\Omega)),$$

with the estimate

$$\begin{aligned}\|u\|_{L^\infty(0,T;H^2(\Omega))} + \|u'\|_{L^\infty(0,T;L^2(\Omega))} + \|u'\|_{L^2(0,T;H_0^1(\Omega))} \\ + \|u''\|_{L^2(0,T;H^{-1}(\Omega))} \leq C(\|g\|_{H^2(\Omega)} + \|f\|_{H^1(0,T;L^2(\Omega))}),\end{aligned} \tag{4.25}$$

where C depends only on Ω, T and the coefficients of L.

Proof. Let $\{u_k\}$ be the Galerkin approximations satisfying (4.10) constructed as above with $\{w_i\}$ being the complete collection of eigenfunctions with eigenvalues $\{\lambda_i\}$ for $-\Delta$ on $H_0^1(\Omega)$. As before, we assume $\{w_i\}$ is orthogonal on $H_0^1(\Omega)$ and orthonormal on $L^2(\Omega)$.

By the uniqueness theorem, the weak solution u is obtained as the limit of the Galerkin approximations $\{u_k\}$. We prove the theorem by deriving the same estimates for the approximate solutions u_k independent of k.

1. We first claim the following estimate for $\{u_k\}$: for each $t \in [0,T]$,

$$\|u_k(t)\|^2_{H^2(\Omega)} \le C(\|f(t)\|^2_{L^2(\Omega)} + \|u_k'(t)\|^2_{L^2(\Omega)} + \|u_k(t)\|^2_{L^2(\Omega)}), \tag{4.26}$$

where C is a constant independent of k. To prove (4.26), note

$$(Lu_k(t), w_i) = B[u_k(t), w_i; t] = (f(t) - u_k'(t), w_i) \quad (i = 1, 2, \cdots, k).$$

Multiply this equation by $\lambda_i d_i(t)$ and sum over $i = 1, 2, \cdots, k$ to deduce

$$(Lu_k(t), -\Delta u_k(t)) = B[u_k(t), -\Delta u_k(t); t] = (f(t) - u_k'(t), -\Delta u_k(t)), \tag{4.27}$$

since $-\Delta u_k(t) \in H^1_0(\Omega)$. Then (4.26) follows from (4.27) and Lemma 4.4.

2. We multiply the equation in (4.10) by $d_i'(t)$ and sum $i = 1, 2, \cdots, k$, to discover

$$(u_k'(t), u_k'(t)) + B[u_k(t), u_k'(t); t] = (f(t), u_k'(t)). \tag{4.28}$$

Write $B[u, v; t] = A[u, v; t] + C[u, v; t]$, where

$$\begin{aligned} A[u, v; t] &= \int_\Omega \left(\sum_{i,j=1}^n a_{ij}(x,t) D_i u D_j v \right) dx, \\ C[u, v; t] &= \int_\Omega \left(\sum_{i=1}^n b_i(x,t)(D_i u) v + c(x,t) uv \right) dx. \end{aligned} \tag{4.29}$$

Note that $A[u, v; t]$ is a symmetric bilinear form on $H^1_0(\Omega)$ and for any functions $u \in C^1([0,T]; H^1_0(\Omega))$,

$$A[u'(t), u(t); t] = A[u(t), u'(t); t] = \frac{1}{2}\left(\frac{d}{dt} A[u(t), u(t); t] - \tilde{A}[u(t), u(t); t] \right), \tag{4.30}$$

where

$$\tilde{A}[u, v; t] = \int_\Omega \left(\sum_{i,j=1}^n a_{ij}'(x,t) D_i u D_j v \right) dx.$$

The equation (4.28) above can be written as

$$\begin{aligned} &\|u_k'(t)\|^2_{L^2(\Omega)} + \frac{1}{2}\frac{d}{dt}\left(A[u_k(t), u_k(t); t]\right) \\ &\qquad = (f(t), u_k'(t)) - C[u_k(t), u_k'(t); t] + \frac{1}{2}\tilde{A}[u_k(t), u_k(t); t]. \end{aligned} \tag{4.31}$$

Moreover, for all $\epsilon > 0$,

$$|C[u, v; t]| \le \epsilon \|v\|^2_{L^2(\Omega)} + C_\epsilon \|u\|^2_{H^1_0(\Omega)} \quad (u,\, v \in H^1_0(\Omega)). \tag{4.32}$$

Therefore, by (4.31), we have

$$\|u_k'(t)\|^2_{L^2(\Omega)} + \frac{d}{dt}\left(A[u_k(t), u_k(t); t]\right) \le C(\|f(t)\|^2_{L^2(\Omega)} + \|u_k\|^2_{H^1_0(\Omega)}). \tag{4.33}$$

Integrate over $t \in [0,T]$ to have

$$\begin{aligned} &\int_0^T \|u_k'(t)\|^2_{L^2(\Omega)} dt + \max_{t\in[0,T]} A[u_k(t), u_k(t); t] \\ &\le A[u_k(0), u_k(0); 0] + C(\|u_k\|^2_{L^2(0,T;H^1_0(\Omega))} + \|f\|^2_{L^2(\Omega_T)}) \\ &\le C(\|g\|^2_{H^1_0(\Omega)} + \|f\|^2_{L^2(\Omega_T)}), \end{aligned}$$

where we have used (4.11) and the estimate $A[u_k(0), u_k(0); 0] \le C\|u_k(0)\|^2_{H_0^1(\Omega)} \le C\|g\|^2_{H_0^1(\Omega)}$, since

$$\|u_k(0)\|^2_{H_0^1(\Omega)} = \sum_{i=1}^k d_i^2(0)\|w_i\|^2_{H_0^1(\Omega)} = \sum_{i=1}^k (g, w_i)^2 \|w_i\|^2_{H_0^1(\Omega)} \le \sum_{i=1}^\infty (g, w_i)^2 \|w_i\|^2_{H_0^1(\Omega)} = \|g\|^2_{H_0^1(\Omega)}.$$

Therefore, using $A[v, v; t] \ge \theta\|v\|^2_{H_0^1(\Omega)}$ for all $v \in H_0^1(\Omega)$,

$$\|u_k'\|^2_{L^2(0,T;L^2(\Omega))} + \|u_k\|^2_{L^\infty(0,T;H_0^1(\Omega))} \le C(\|g\|^2_{H_0^1(\Omega)} + \|f\|^2_{L^2(\Omega_T)}).$$

From this and (4.26),

$$\|u_k\|^2_{L^2(0,T;H^2(\Omega))} \le C(\|g\|^2_{H_0^1(\Omega)} + \|f\|^2_{L^2(\Omega_T)}).$$

Note that (4.24) follows from the two estimates above.

3. Assume now the hypotheses of assertion (ii). We differentiate the equation in (4.10) with respect to t and set $\tilde{u}_k := u_k'$ to obtain

$$(4.34) \qquad (\tilde{u}_k'(t), w_i) + B[\tilde{u}_k(t), w_i; t] = (f'(t), w_i) - \tilde{B}[u_k(t), w_i; t],$$

where $\tilde{B}$ is the bilinear form defined by

$$\tilde{B}[u, v; t] = \int_\Omega \Big(\sum_{i,j=1}^n a_{ij}'(x,t) D_i u D_j v + \sum_{i=1}^n b_i'(x,t)(D_i u)v + c'(x,t)uv \Big) dx.$$

Multiplying (4.34) by $d_i'(t)$ and summing over $i = 1, 2, \cdots, k$, we discover

$$(\tilde{u}_k'(t), \tilde{u}_k(t)) + B[\tilde{u}_k(t), \tilde{u}_k(t); t] = (f'(t), \tilde{u}_k(t)) - \tilde{B}[u_k(t), \tilde{u}_k(t); t].$$

So,

$$(4.35) \qquad \frac{1}{2}\frac{d}{dt}\Big(\|\tilde{u}_k(t)\|^2_{L^2(\Omega)}\Big) + B[\tilde{u}_k(t), \tilde{u}_k(t); t] = (f'(t), \tilde{u}_k(t)) - \tilde{B}[u_k(t), \tilde{u}_k(t); t].$$

Notice that, for $u \in H_0^1(\Omega) \cap H^2(\Omega), v \in H_0^1(\Omega)$, integration by parts yields

$$\tilde{B}[u, v; t] = \int_\Omega v \Big(\sum_{i,j=1}^n (-D_j a_{ij}' D_i u - a_{ij}' D_{ij} u) + \sum_{i=1}^n b_i'(x,t) D_i u + c'(x,t)u \Big) dx,$$

and, hence, for a.e. $t \in [0, T]$ and all $u \in H^2(\Omega),\ v \in H_0^1(\Omega)$,

$$|\tilde{B}[u, v; t]| \le C(\|v\|^2_{H_0^1(\Omega)} + \|u\|^2_{H^2(\Omega)}).$$

Using this estimate and Gårding's inequality, we deduce from (4.67)

$$\frac{d}{dt}\Big(\|\tilde{u}_k(t)\|^2_{L^2(\Omega)}\Big) + \beta\,\|\tilde{u}_k(t)\|^2_{H^1(\Omega)} \le C(\|\tilde{u}_k(t)\|^2_{L^2(\Omega)} + \|u_k(t)\|^2_{H^2(\Omega)} + \|f'(t)\|^2_{L^2(\Omega)}).$$

Hence Gronwall's inequality implies

$$
\begin{aligned}
&\sup_{t\in[0,T]} \|\tilde{u}_k(t)\|^2_{L^2(\Omega)} + \beta \int_0^T \|\tilde{u}_k(t)\|^2_{H^1(\Omega)} dt \\
&\le C\left(\|\tilde{u}_k(0)\|^2_{L^2(\Omega)} + \int_0^T (\|u_k(t)\|^2_{H^2(\Omega)} + \|f'(t)\|^2_{L^2(\Omega)}) dt \right) \\
&\le C(\|\tilde{u}_k(0)\|^2_{L^2(\Omega)} + \|f\|^2_{H^1(0,T;L^2(\Omega))} + \|g\|^2_{H^1_0(\Omega)}) \\
&\le C(\|u_k(0)\|^2_{H^2(\Omega)} + \|f\|^2_{H^1(0,T;L^2(\Omega))} + \|g\|^2_{H^1_0(\Omega)}),
\end{aligned}
\tag{4.36}
$$

where, recall that $\tilde{u}_k = u'_k$ and by equation (4.10),

$$
\begin{aligned}
\|\tilde{u}_k(0)\|_{L^2(\Omega)} = \|u'_k(0)\|_{L^2(\Omega)} &\le C(\|f(0)\|_{L^2(\Omega)} + \|u_k(0)\|_{H^2(\Omega)}) \\
&\le C(\|f\|_{H^1(0,T;L^2(\Omega))} + \|u_k(0)\|_{H^2(\Omega)}).
\end{aligned}
$$

4. We must estimate $\|u_k(0)\|_{H^2(\Omega)}$. This is a little tricky. Recall that $\{w_i\}$ is the complete set of smooth eigenfunctions of $-\Delta$ on $H^1_0(\Omega)$. Since both u_k and Δu_k are in $H^1_0(\Omega) \cap H^2(\Omega)$, we have

$$
\|u_k(0)\|^2_{H^2(\Omega)} \le C\|\Delta u_k(0)\|^2_{L^2(\Omega)} = C(u_k(0), \Delta^2 u_k(0)) = C(g, \Delta^2 u_k(0)),
$$

since $(u_k(0), w_i) = (g, w_i)$, and

$$
(g, \Delta^2 u_k(0)) = (\Delta g, \Delta u_k(0)) \le \varepsilon \|u_k(0)\|^2_{H^2(\Omega)} + C_\varepsilon \|g\|^2_{H^2(\Omega)}.
$$

From these, with sufficiently small $\varepsilon > 0$, we have

$$
\|u_k(0)\|^2_{H^2(\Omega)} \le C(g, \Delta^2 u_k(0)) \le C\|g\|^2_{H^2(\Omega)}.
$$

Hence, by (4.36),

$$
\|u'_k\|_{L^\infty(0,T;L^2(\Omega))} + \|u'_k\|_{L^2(0,T;H^1(\Omega))} \le C(\|f\|^2_{H^1(0,T;L^2(\Omega))} + \|g\|^2_{H^2(\Omega)}).
$$

From this and (4.26), noting $\max_{t\in[0,T]} \|f(t)\|_{L^2(\Omega)} \le \|f\|_{H^1(0,T;L^2(\Omega))}$, we deduce

$$
\|u_k\|^2_{L^\infty(0,T;H^2(\Omega))} \le C(\|f\|^2_{H^1(0,T;L^2(\Omega))} + \|g\|^2_{H^2(\Omega)}).
$$

5. It remains to show $u'' \in L^2(0,T;H^{-1}(\Omega)))$. To do so, take $v \in H^1_0(\Omega)$ with $\|v\|_{H^1_0(\Omega)} \le 1$ and set $v = v^1 + v^2$, as above with $v^1 \in V_k$ and $(v^2, w_i) = 0$ for $i = 1, 2, \cdots, k$. Then, for a.e. $t \in [0,T]$, by (4.34),

$$
\langle u''_k(t), v\rangle = (u''_k(t), v) = (u''_k(t), v^1) = (f'(t), v^1) - B[u'_k(t), v^1; t] - \tilde{B}[u_k(t), v^1; t].
$$

Hence, since $\|v^1\|_{H^1_0(\Omega)} \le 1$,

$$
|\langle u''_k(t), v\rangle| \le C(\|f'(t)\|_{L^2(\Omega)} + \|u'_k(t)\|_{H^1_0(\Omega)} + \|u_k(t)\|_{H^1_0(\Omega)}).
$$

This proves

$$
\|u''_k(t)\|_{H^{-1}(\Omega)} \le C(\|f'(t)\|_{L^2(\Omega)} + \|u'_k(t)\|_{H^1_0(\Omega)} + \|u_k(t)\|_{H^1_0(\Omega)}).
$$

So, squaring, integrating over $t \in [0,T]$ and using the estimates obtained above, we have

$$
\begin{aligned}
\|u''_k\|^2_{L^2(0,T;H^{-1}(\Omega))} &\le C(\|f'\|^2_{L^2(0,T;L^2(\Omega))} + \|u'_k\|^2_{L^2(0,T;H^1_0(\Omega))} + \|u_k\|^2_{L^2(0,T;H^1_0(\Omega))}) \\
&\le C(\|f\|^2_{H^1(0,T;L^2(\Omega))} + \|g\|^2_{H^2(\Omega)}).
\end{aligned}
$$

By limit, this proves $u'' \in L^2(0,T;H^{-1}(\Omega))$ with the desired norm estimate. □

We now study the higher regularity. For simplicity, we assume the coefficients of L are smooth and *independent of time* t; we also assume the uniform parabolicity and the smoothness of the domain Ω.

Theorem 4.6. (Higher regularity) *Assume $m \geq 0$ in an integer and*

$$g \in H^{2m+1}(\Omega), \quad \frac{d^k f}{dt^k} \in L^2(0,T; H^{2m-2k}(\Omega)) \;\; (k = 0, 1, \cdots, m).$$

*Suppose the following m-***th-order compatibility conditions** *hold:*

$$\begin{aligned} & g_0 := g \in H_0^1(\Omega), \;\; g_1 := f(0) - Lg_0 \in H_0^1(\Omega), \\ & \cdots, \;\; g_m := \frac{d^{m-1} f}{dt^{m-1}}(0) - Lg_{m-1} \in H_0^1(\Omega). \end{aligned} \tag{4.37}$$

Then the weak solution u to (4.1) *satisfies*

$$\frac{d^k u}{dt^k} \in L^2(0,T; H^{2m-2k+2}(\Omega)) \quad (k = 0, 1, 2, \cdots, m+1),$$

with the estimate

$$\sum_{k=0}^{m+1} \| \frac{d^k u}{dt^k} \|_{L^2(0,T;H^{2m-2k+2}(\Omega))} \leq C \left(\sum_{k=0}^{m} \| \frac{d^k f}{dt^k} \|_{L^2(0,T;H^{2m-2k}(\Omega))} + \|g\|_{H^{2m+1}(\Omega)} \right). \tag{4.38}$$

Proof. The proof is an induction on m, the case $m = 0$ being the conclusion (i) of Theorem 4.5 above. Assume now the theorem is valid for some integer $m \geq 0$, and suppose then

$$g \in H^{2m+3}(\Omega), \quad \frac{d^k f}{dt^k} \in L^2(0,T; H^{2m+2-2k}(\Omega)) \;\; (k = 0, 1, \cdots, m+1)$$

and the $(m+1)$-th-order compatibility conditions hold. Let $\tilde{u} = u'$. Then the previous theorem implies that $\tilde{u} \in L^2(0,T; H_0^1(\Omega))$ and $\tilde{u}' \in L^2(0,T; H^{-1}(\Omega))$. Also $\tilde{u}$ is the weak solution to

$$\begin{cases} \tilde{u}_t + L\tilde{u} = \tilde{f} & \text{in } \Omega_T, \\ \tilde{u} = 0 & \text{on } \partial\Omega \times [0,T], \\ \tilde{u} = \tilde{g} & \text{on } \Omega \times \{t = 0\}, \end{cases}$$

where $\tilde{f} = f'$, $\tilde{g} = f(0) - Lg = g_1$. In particular, $\tilde{f}$ and $\tilde{g}$ satisfy the m-th order compatibility conditions. Then we use induction on m; details are referred to Evans's book. ☐

Remark 4.5. The condition on f implies

$$f(0) \in H^{2m-1}(\Omega), \; f'(0) \in H^{2m-3}(\Omega), \; \cdots, f^{(m-1)}(0) \in H^1(\Omega)$$

and consequently

$$g \in H^{2m+1}(\Omega), \;\; g_1 \in H^{2m-1}(\Omega), \cdots, \;\; g_m \in H^1(\Omega).$$

The compatibility conditions are precisely the requirements that, each of these functions g_k is zero on $\partial\Omega$. This is because the homogeneous boundary condition $u(x,t) = 0$ foron $x \in \partial\Omega$ (in trace), making $\frac{d^k u}{dt^k}(x,t) = 0$ on $x \in \partial\Omega$ (in trace) for each $k = 0, 1, \cdots, m$.

Theorem 4.7. (Smoothness of weak solution) *Assume $g \in C^\infty(\bar{\Omega})$, $f \in C^\infty(\overline{\Omega_T})$, and the m-th-order compatibility conditions hold for all $m = 0, 1, \cdots$. Then problem* (4.1) *has a unique solution $u \in C^\infty(\overline{\Omega_T})$.*

4.1.6. Maximum Principle. We now study some properties for classical (smooth) solutions of parabolic equations. We include such a study here in order to compare with the properties for classical solutions of hyperbolic equations we shall study later in Section 4.2.6. Although the methods we use here to treat both parabolic and hyperbolic equations are very similar, we shall see that their solutions behave quite different.

Let us denote by $C^{2,1}(\Omega_T)$ functions satisfying $u_t, u_{x_ix_j} \in C(\Omega_T)$. We shall consider general inequalities of the form

$$u_t + Lu \le 0 \quad \text{in } \Omega_T \tag{4.39}$$

where we assume L is of nondivergence form:

$$Lu = -\sum_{i,j=1}^{n} a_{ij}(x,t)\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum b_i(x,t)\frac{\partial u}{\partial x_i} + c(x,t)u$$

and satisfies the ellipticity condition

$$\sum_{i,j=1}^{n} a_{ij}(x,t)\xi_i\xi_j \ge \lambda|\xi|^2 \quad \text{for } (x,t) \in \Omega_T$$

where $\lambda > 0$ is a constant, and the coefficients a_{ij}, b_i, c are all bounded functions in Ω_T.

Theorem 4.8. (Weak Maximum Principle) *Suppose $u \in C^{2,1}(\Omega_T) \cap C(\overline{\Omega_T})$ satisfies (4.39) and $c \ge 0$. Then*

$$\max_{\Omega_T} u \le \max_{\partial'\Omega_T} u^+. \tag{4.40}$$

Theorem 4.9. (Parabolic Harnack inequality) *Assume $u \in C^{2,1}(\Omega_T)$ solves $u_t + Lu = 0$ in Ω_T and $u \ge 0$ in Ω_T. Suppose $\Omega' \subset\subset \Omega$ is connected. Then, for each $0 < t_1 < t_2 \le T$, there exists a constant C depending only on V, t_1, t_2, and the coefficients of L, such that*

$$\sup_V u(\cdot, t_1) \le C \inf_V u(\cdot, t_2). \tag{4.41}$$

Theorem 4.10. (Strong Maximum Principle) *Suppose $u \in C^{2,1}(\Omega_T) \cap C(\overline{\Omega_T})$ satisfies (4.39). Let*

$$M = \max_{\overline{\Omega_T}} u = u(x_0, t_0).$$

Assume one of the following conditions holds:

(a) $c(x,t) \equiv 0$; *(b)* $c(x,t) \ge 0$ *and* $M \ge 0$; *(c)* $c(x,t)$ *arbitrary and* $M = 0$.

Then we have the **strong maximum principle***:*

(i) If $(x_0, t_0) \in \Omega_T$, then $u(x,t) \equiv M$ for all $(x,t) \in \Omega_{t_0}$.

(ii) If $x_0 \in \partial\Omega$ and $0 < t_0 < T$, but $u(x,t) < M$ for all $x \in \Omega$, $0 < t < t_0$, then

$$\frac{\partial u}{\partial \nu}(x_0, t_0) > 0 \tag{4.42}$$

provided the exterior normal derivative exists at (x_0, t_0).

4.2. Second-order Hyperbolic Equations

We will study the **initial-boundary value problem**

$$\begin{cases} u_{tt} + Lu = f & \text{in } \Omega_T, \\ u = 0 & \text{on } \partial\Omega \times [0,T], \\ u = g, \ u_t = h & \text{on } \Omega \times \{t=0\}, \end{cases} \tag{4.43}$$

where $f\colon \Omega_T \to \mathbb{R}$ and $g, h\colon \Omega \to \mathbb{R}$ are given and $u\colon \overline{\Omega_T} \to \mathbb{R}$ is the unknown function, $u = u(x,t)$.

The operator Lu denotes for each time t a second-order partial differential operator, having either the divergence form

$$Lu = -\sum_{i,j=1}^{n} D_j(a_{ij}(x,t)D_i u) + \sum_{i=1}^{n} b_i(x,t)D_i u + c(x,t)u \tag{4.44}$$

or else the nondivergence form

$$Lu = -\sum_{i,j=1}^{n} a_{ij}(x,t)D_{ij}u + \sum_{i=1}^{n} b_i(x,t)D_i u + c(x,t)u, \tag{4.45}$$

for given coefficients a_{ij}, b_i, c $(i,j = 1,2,\cdots,n)$.

Definition 4.6. We say the operator $\frac{\partial^2}{\partial t^2} + L$ is called **(uniformly) hyperbolic** on Ω_T if there exists a constant $\theta > 0$ such that

$$\sum_{i,j=1}^{n} a_{ij}(x,t)\xi_i\xi_j \ge \theta|\xi|^2 \quad \text{for all } (x,t) \in \Omega_T \text{ and } \xi \in \mathbb{R}^n. \tag{4.46}$$

Note that for each fixed time $t \in [0,T]$ the operator Lu is uniformly elliptic on Ω.

4.2.1. Weak Solutions. We consider the case that Lu has the divergence form (4.44). Let $B[u,v;t]$ be the time-dependent bilinear form defined as above.

Definition 4.7. A **weak solution** to Problem (4.43) is a function $u \in L^2(0,T;H_0^1(\Omega))$ having weak time-derivatives $u' \in L^2(0,T;L^2(\Omega))$ and $u'' \in L^2(0,T;H^{-1}(\Omega))$ such that

(i) $\langle u''(t), v\rangle + B[u(t),v;t] = (f(t),v)$ for each $v \in H_0^1(\Omega)$ and a.e. time $t \in [0,T]$, and

(ii) $u(0) = g$, $u'(0) = h$. (Note that $u \in C([0,T];L^2(\Omega))$ and $u' \in C([0,T];H^{-1}(\Omega))$, and thus $u(0)$ and $u'(0)$ are well-defined.)

4.2.2. Galerkin Approximations. We assume

$$\begin{aligned} a_{ij},\ b_i,\ c \in C^1(\overline{\Omega_T}) \quad (i,j = 1,2,\cdots,n), \\ f \in L^2(\Omega_T),\ g \in H_0^1(\Omega),\ h \in L^2(\Omega). \end{aligned}$$

We will also assume $a_{ij} = a_{ji}$ for $i,j = 1,2,\cdots,n$.

Again, assume the functions $w_i = w_i(x)$ are smooth and

$$\{w_i\}_{i=1}^{\infty} \text{ forms an orthogonal basis of } H_0^1(\Omega) \text{ and an orthonormal basis of } L^2(\Omega). \tag{4.47}$$

(For instance, we could take $\{w_i\}$ to be the complete set of appropriately normalized eigenfunctions for $-\Delta$ in $H_0^1(\Omega)$.)

Fix now a positive integer k. Let V_k be the linear span of $\{w_1, \cdots, w_k\}$ and we look for a function $u_k\colon [0,T] \to V_k$ of the form

$$u_k(t) = \sum_{i=1}^{k} d_i(t) w_i, \tag{4.48}$$

where the coefficient functions $d_i(t)$ is selected so that

$$\begin{cases} (u_k''(t), w_i) + B[u_k(t), w_i; t] = (f(t), w_i), \\ d_i(0) = (g, w_i), \;\; d_i'(0) = (h, w_i), \end{cases} \tag{4.49}$$

for almost every $t \in [0,T]$ and $i = 1, 2, \cdots, k$. Here $(\;,\;)$ denotes the inner product in $L^2(\Omega)$.

Theorem 4.11. (Construction of approximate solutions) *For each $k = 1, 2, \cdots$ there exists a unique function u_k of the form above satisfying* (4.49).

Proof. Note that

$$(u_k''(t), w_i) = d_i''(t), \quad B[u_k(t), w_i; t] = \sum_{j=1}^{k} \alpha_i^j(t) d_j(t),$$

where $\alpha_i^j(t) = B[w_j, w_i; t]$ $(i, j = 1, 2, \cdots, k)$. Hence condition (4.10) becomes the initial value problem for the ODE system on $d(t) = (d_1(t), \cdots, d_k(t))$:

$$\begin{cases} d_i''(t) + \sum_{j=1}^{k} \alpha_i^j(t) d_j(t) = f_i(t) \equiv (f(t), w_i), \\ d_i(0) = (g, w_i), \; d_i'(0) = (h, w_i) \quad (i = 1, 2, \cdots, k). \end{cases}$$

Note that the coefficients α_i^j belong to $L^\infty(0,T)$ and $f_i \in L^2(0,T)$. The existence of a unique solution $d \in H^2(0,T) \subset C^1([0,T])$ is guaranteed by the (not so) standard existence theory for ODE (think of approximating α_i^j and f_i by smooth functions first and then pass to limits). □

4.2.3. Energy Estimates.

Theorem 4.12. *Assume the uniform hyperbolicity condition. There exists a constant C, depending only on Ω, T, and the coefficients of L, such that, for all $k = 1, 2, \cdots$,*

$$\begin{aligned} \max_{t\in[0,T]} (\|u_k(t)\|_{H_0^1(\Omega)} + \|u_k'(t)\|_{L^2(\Omega)}) + \|u_k''\|_{L^2(0,T;H^{-1}(\Omega))} \\ \le C(\|f\|_{L^2(\Omega_T)} + \|g\|_{H_0^1(\Omega)} + \|h\|_{L^2(\Omega)}). \end{aligned} \tag{4.50}$$

Proof. 1. Multiply (4.49) by $d_i'(t)$ and sum for $i = 1, 2, \cdots, k$ to find

$$(u_k''(t), u_k'(t)) + B[u_k(t), u_k'(t); t] = (f(t), u_k'(t)) \quad \forall \; a.e. \; t \in [0,T]. \tag{4.51}$$

Observe $(u_k'', u_k') = \frac{d}{dt}(\frac{1}{2}\|u_k'\|^2_{L^2(\Omega)})$; furthermore, as above,

$$B[u_k(t), u_k'(t); t] = A[u_k(t), u_k'(t); t] + C[u_k(t), u_k'(t); t] := B_1 + B_2,$$

where,

$$\begin{aligned} B_1 &= \frac{d}{dt}\left(\frac{1}{2}A[u_k(t), u_k(t); t]\right) - \frac{1}{2}\int_\Omega \sum_{i,j=1}^{n} a_{ij}' D_i u_k D_j u_k dx \\ &\ge \frac{d}{dt}\left(\frac{1}{2}A[u_k(t), u_k(t); t]\right) - C\|u_k(t)\|^2_{H_0^1(\Omega)}. \end{aligned}$$

We also note that

$$|B_2| \le C(\|u_k(t)\|^2_{H^1_0(\Omega)} + \|u_k'(t)\|^2_{L^2(\Omega)}).$$

Hence, in view of (4.51), we discover

(4.52)

$$\begin{aligned}\frac{d}{dt}\Big(\|u_k'(t)\|^2_{L^2(\Omega)} + A[u_k(t),u_k(t);t]\Big) &\le C(\|u_k'(t)\|^2_{L^2(\Omega)} + \|u_k(t)\|^2_{H^1_0(\Omega)} + \|f(t)\|^2_{L^2(\Omega)})\\ &\le C(\|u_k'(t)\|^2_{L^2(\Omega)} + A[u_k(t),u_k(t);t] + \|f(t)\|^2_{L^2(\Omega)}),\end{aligned}$$

where we used the Gårding's inequality $A[u,u;t] \ge \theta\|u\|^2_{H^1_0(\Omega)}$.

2. Now write

$$\eta(t) = \|u_k'(t)\|^2_{L^2(\Omega)} + A[u_k(t),u_k(t);t], \quad \xi(t) = \|f(t)\|^2_{L^2(\Omega)}.$$

Then

$$\eta'(t) \le C\eta(t) + C\xi(t) \quad \forall\, a.e.\ t \in [0,T].$$

Thus, by Gronwall's inequality,

$$\eta(t) \le e^{Ct}\left(\eta(0) + C\int_0^t \xi(s)\,ds\right) \quad (0 \le t \le T).$$

Since

$$\eta(0) = \|u_k'(0)\|^2_{L^2(\Omega)} + A[u_k(0),u_k(0);0] \le C(\|h\|^2_{L^2(\Omega)} + \|g\|^2_{H^1_0(\Omega)}),$$

according to the initial data in (4.49) and $\|u_k(0)\|_{H^1_0(\Omega)} \le \|g\|_{H^1_0(\Omega)}$, we thus obtain

$$\max_{t\in[0,T]}(\|u_k'(t)\|^2_{L^2(\Omega)} + A[u_k(t),u_k(t);t]) \le C(\|h\|^2_{L^2(\Omega)} + \|g\|^2_{H^1_0(\Omega)} + \|f\|^2_{L^2(\Omega_T)}).$$

This proves

(4.53) $$\max_{t\in[0,T]}(\|u_k(t)\|_{H^1_0(\Omega)} + \|u_k'(t)\|_{L^2(\Omega)}) \le C(\|f\|_{L^2(\Omega_T)} + \|g\|_{H^1_0(\Omega)} + \|h\|_{L^2(\Omega)}).$$

3. Finally we need to estimate $\|u_k''\|_{L^2(0,T;H^{-1}(\Omega))}$. So, fix any $v \in H^1_0(\Omega)$, with $\|v\|_{H^1_0(\Omega)} \le 1$. We write $v = v^1 + v^2$, where $v^1 \in V_k$, and $(v^2, w_i) = 0$ for all $i = 1,2,\cdots,k$. (That is, v^2 is in the L^2 orthogonal complement of V_k.) Since $\{w_i\}$ are orthogonal in $H^1_0(\Omega)$, we have

$$\|v^1\|_{H^1_0(\Omega)} \le \|v\|_{H^1_0(\Omega)} \le 1.$$

Using (4.49), we have

$$(u_k''(t), v^1) + B[u_k(t), v^1; t] = (f(t), v^1).$$

Then

$$\langle u_k''(t), v\rangle = (u_k''(t), v) = (u_k''(t), v^1) = (f(t), v^1) - B[u_k(t), v^1; t]$$

and consequently

$$|\langle u_k''(t), v\rangle| \le C(\|f(t)\|_{L^2(\Omega)} + \|u_k\|_{H^1_0(\Omega)}).$$

This implies

$$\|u_k''(t)\|^2_{H^{-1}(\Omega)} \le C(\|f(t)\|^2_{L^2(\Omega)} + \|u_k(t)\|^2_{H^1_0(\Omega)}) \quad \forall\, t \in [0,T].$$

Integrate over $t \in [0,T]$ to finally obtain

$$\|u_k''\|^2_{L^2(0,T;H^{-1}(\Omega))} = \int_0^T \|u_k'(t)\|^2_{H^{-1}(\Omega)} dt \le C(\|f\|^2_{L^2(\Omega_T)} + \|u_k\|^2_{L^2(0,T;H^1_0(\Omega))}),$$

which, combined with the estimate (4.53), derives the desired estimate. □

4.2.4. Existence and Uniqueness of Weak Solutions.

Theorem 4.13. (Existence) *There exists a weak solution to* (4.1).

Proof. 1. According to the energy estimate (4.50), we see that $\{u_k\}$ is bounded in $L^2(0,T;H_0^1(\Omega))$, $\{u_k'\}$ is bounded in $L^2(0,T;L^2(\Omega))$, and $\{u_k''\}$ is bounded in $L^2(0,T;H^{-1}(\Omega))$. Consequently there exists a subsequence $\{u_{k_m}\}$ of $\{u_k\}$ with $k_m \to \infty$ and functions $u \in L^2(0,T;H_0^1(\Omega))$, with $u' \in L^2(0,T;L^2(\Omega))$, $u'' \in L^2(0,T;H^{-1}(\Omega))$, such that

$$\begin{cases} u_{k_m} \rightharpoonup u & \text{in } L^2(0,T;H_0^1(\Omega)), \\ u_{k_m}' \rightharpoonup u' & \text{in } L^2(0,T;L^2(\Omega)), \\ u_{k_m}'' \rightharpoonup u'' & \text{in } L^2(0,T;H^{-1}(\Omega)). \end{cases} \tag{4.54}$$

In fact, by estimate (4.50), we also have $u \in L^\infty(0,T;H_0^1(\Omega)) \cap C([0,T];L^2(\Omega))$ and $u' \in L^\infty(0,T;L^2(\Omega))$; moreover, $u' \in C([0,T];H^{-1}(\Omega))$.

2. Fix any integer N and let $\psi \in C^1([0,T];H_0^1(\Omega))$ have the form

$$\psi(t) = \sum_{i=1}^N \zeta_i(t) w_i,$$

where $\zeta_i \in C^1([0,T];\mathbb{R})$. Let $k \geq N$, multiply (4.10) by $\zeta_i(t)$, sum $i = 1, 2, \cdots, N$ and then integrate over $t \in [0,T]$ to find

$$\int_0^T \big(\langle u_k''(t), \psi(t)\rangle + B[u_k(t), \psi(t); t]\big)\, dt = \int_0^T (f(t), \psi(t)) dt. \tag{4.55}$$

Now let $k = k_m \to \infty$ and we have

$$\int_0^T \big(\langle u''(t), \psi(t)\rangle + B[u(t), \psi(t); t]\big)\, dt = \int_0^T (f(t), \psi(t)) dt. \tag{4.56}$$

This equality then holds for all functions $\psi \in L^2(0,T;H_0^1(\Omega))$, as functions ζ of the given form are dense in this space. We then take $\psi(t) = \zeta(t) v$ with $\zeta \in L^2(0,T)$ and $v \in H_0^1(\Omega)$ in (4.56) to obtain

$$\int_0^T \zeta(t) \big(\langle u''(t), v\rangle + B[u(t), v; t]\big)\, dt = \int_0^T \zeta(t)(f(t), v) dt.$$

This holding for all $\zeta \in L^2(0,T)$ yields that

$$\langle u''(t), v\rangle + B[u(t), v; t] = (f(t), v) \quad \forall\, v \in H_0^1(\Omega), \ \ a.e.\ t \in [0,T]. \tag{4.57}$$

3. We need to show the initial data $u(0) = g, u'(0) = h$. In (4.55), (4.56), we take $\psi(t) = \alpha(t) v + \beta(t) w$ with $\alpha, \beta \in C^2[0,T]$ and $v, w \in H_0^1(\Omega)$ arbitrarily given such that $\psi(T) = \psi'(T) = 0$, $\psi(0) = v$ and $\psi'(0) = w$. Note that

$$\int_0^T \langle u_k''(t), \psi(t)\rangle dt = -(u_k'(0), v) + (u_k(0), w) + \int_0^T (u_k(t), \psi''(t))\, dt,$$

$(u_k(0), w) \to (g, w)$, $(u_k'(0), v) \to (h, v)$ as $k = k_m \to \infty$, and

$$\int_0^T \langle u''(t), \psi(t)\rangle dt = -\langle u'(0), v\rangle + (u(0), w) + \int_0^T (u(t), \psi''(t))\, dt,$$

Hence, in (4.55), let $k = k_m \to \infty$, we eventually obtain

$$-\langle u'(0), v\rangle + (u(0), w) = -(h, v) + (g, w) \quad \forall\, v, w \in H_0^1(\Omega);$$

hence $u(0) = g$ and $u'(0) = h$. □

Theorem 4.14. (Uniqueness) *A weak solution of* (4.43) *is unique.*

Proof. 1. It suffices to prove that a weak solution u with $f = g = 0$ must be zero. Unlike the parabolic case, the proof here is tricky because we cannot insert $v = u'(t)$ in (4.49) since $u'(t) \notin H_0^1(\Omega)$. We instead consider, for each fixed $s \in [0, T]$, the function

$$v(t) = \begin{cases} \int_t^s u(\tau)d\tau & \text{if } t \in [0, s], \\ 0 & \text{if } t \in [s, T]. \end{cases}$$

Then for each $t \in [0, T]$, $v(t) \in H_0^1(\Omega)$, and so, by the weak solution definition,

$$\int_0^s (\langle u''(t), v(t) \rangle + B[u(t), v(t); t])\, dt = 0.$$

Since $u'(0) = v(s) = 0$, we obtain by integration by parts

$$\int_0^s (-\langle u'(t), v'(t) \rangle + B[u(t), v(t); t])\, dt = 0. \tag{4.58}$$

As above, we write $B[u, v; t] = A[u, v; t] + C[u, v; t]$. Note, for all $u, v \in H_0^1(\Omega)$ and $t \in [0, T]$,

$$C[u, v; t] = \int_\Omega \Big(\sum_{i=1}^n b_i D_i uv + cuv \Big)\, dx = \int_\Omega \Big(-\sum_{i=1}^n (D_i b_i uv + b_i u D_i v) + cuv \Big)\, dx.$$

(The trick here is to avoid the $D_i u$ terms.) Since $v'(t) = -u(t)$ on $[0, s]$, using (4.58), we can write

$$\int_0^s (\langle u'(t), u(t) \rangle - A[v'(t), v(t); t])dt = -\int_0^s C[u(t), v(t); t]dt.$$

Note that $(u'(t), u(t)) = \frac{d}{dt}(\frac{1}{2}\|u(t)\|^2_{L^2(\Omega)})$ and, since A is symmetric,

$$A[v'(t), v(t); t] = \frac{d}{dt}\left(\frac{1}{2} A[v(t), v(t); t]\right) - \frac{1}{2}\int_\Omega \sum_{i,j=1}^n a'_{ij} D_i v D_j v dx.$$

Hence

$$\begin{aligned} &\frac{1}{2}\int_0^s \frac{d}{dt}\Big(\|u(t)\|^2_{L^2(\Omega)} - A[v(t), v(t); t]\Big)\, dt \\ &= -\int_0^s C[u(t), v(t); t]dt - \frac{1}{2}\int_0^s \int_\Omega \sum_{i,j=1}^n a'_{ij} D_i v D_j v dx dt, \end{aligned} \tag{4.59}$$

and consequently,

$$\begin{aligned} \|u(s)\|^2_{L^2(\Omega)} + A[v(0), v(0); 0] &= -2\int_0^s C[u(t), v(t); t]dt - \int_0^s \int_\Omega \sum_{i,j=1}^n a'_{ij} D_i v D_j v dx dt \\ &\le C \int_0^s (\|v(t)\|^2_{H_0^1(\Omega)} + \|u(t)\|^2_{L^2(\Omega)})\, dt. \end{aligned}$$

By Gårding's inequality for $A[u, u; t]$, we obtain

$$\|u(s)\|^2_{L^2(\Omega)} + \|v(0\|^2_{H_0^1(\Omega)} \le C \int_0^s (\|v(t)\|^2_{H_0^1(\Omega)} + \|u(t)\|^2_{L^2(\Omega)})\, dt. \tag{4.60}$$

2. Now let

$$w(t) = \int_0^t u(\tau)d\tau \quad (t \in [0, T]).$$

Then $v(0) = w(s)$ and $v(t) = w(s) - w(t)$, and hence (4.60) becomes

$$\|u(s)\|^2_{L^2(\Omega)} + \|w(s)\|^2_{H^1_0(\Omega)} \le C\int_0^s (\|w(t) - w(s)\|^2_{H^1_0(\Omega)} + \|u(t)\|^2_{L^2(\Omega)})dt.$$

But

$$\|w(t) - w(s)\|^2_{H^1_0(\Omega)} \le 2\|w(t)\|^2_{H^1_0(\Omega)} + 2\|w(s)\|^2_{H^1_0(\Omega)},$$

so we have

$$\|u(s)\|^2_{L^2(\Omega)} + (1 - 2sC_1)\|w(s)\|^2_{H^1_0(\Omega)} \le C_1\int_0^s (\|w(t)\|^2_{H^1_0(\Omega)} + \|u(t)\|^2_{L^2(\Omega)})dt.$$

Choose $0 < T_1 < T$ so small that $1 - 2T_1C_1 \ge \frac{1}{2}$. Then if $0 \le s \le T_1$, we have

$$\|u(s)\|^2_{L^2(\Omega)} + \|w(s)\|^2_{H^1_0(\Omega)} \le C_2\int_0^s (\|w(t)\|^2_{H^1_0(\Omega)} + \|u(t)\|^2_{L^2(\Omega)})dt.$$

Hence Gronwall's inequality implies $u \equiv 0$ on $[0, T_1]$.

3. Apply the same argument on the intervals $[T_1, 2T_1], [2T_1, 3T_1]$, etc, to eventually deduce $u \equiv 0$ on $[0, T]$. □

4.2.5. Regularity. We now study the regularity of weak solutions when the initial data and coefficients are more regular. Our eventual goal is to prove that the weak solution is smooth, as long as the coefficients and initial data and the domain are all smooth. Although the methods and results are similar to the ones used for the regularity study of parabolic equations as above, as we shall see later, there are some quite essential differences of regularity concerning these two classes of evolution equations.

We assume the coefficients a_{ij}, b_i, c are smooth on Ω_T. As usual, we assume the uniform hyperbolicity.

Theorem 4.15. (Improved regularity) *(i) Assume $g \in H^1_0(\Omega)$, $h \in L^2(\Omega)$, $f \in L^2(\Omega_T)$. Suppose u is the weak solution of* (4.43)*. Then*

$$u \in L^\infty(0, T; H^1_0(\Omega)), \quad u' \in L^\infty(0, T; L^2(\Omega)),$$

with the estimate

$$\|u\|_{L^\infty(0,T;H^1_0(\Omega))} + \|u'\|_{L^\infty(0,T;L^2(\Omega))} \le C(\|g\|_{H^1_0(\Omega)} + \|h\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega_T)}), \tag{4.61}$$

where C depends only on Ω, T and the coefficients of L.

(ii) If, in addition, $g \in H^2(\Omega)$, $h \in H^1_0(\Omega)$, $f' \in L^2(0, T; L^2(\Omega))$, then

$$u \in L^\infty(0, T; H^2(\Omega)), \quad u' \in L^\infty(0, T; H^1_0(\Omega)),$$
$$u'' \in L^\infty(0, T; L^2(\Omega)), \quad u''' \in L^2(0, T; H^{-1}(\Omega)),$$

with the estimate

$$\begin{aligned}\|u\|_{L^\infty(0,T;H^2(\Omega))} &+ \|u'\|_{L^\infty(0,T;H^1_0(\Omega))} + \|u''\|_{L^\infty(0,T;L^2(\Omega))} \\ &+ \|u'''\|_{L^2(0,T;H^{-1}(\Omega))} \le C(\|g\|_{H^2(\Omega)} + \|h\|_{H^1_0(\Omega)} + \|f\|_{H^1(0,T;L^2(\Omega))}),\end{aligned} \tag{4.62}$$

where C depends only on Ω, T and the coefficients of L.

Proof. Let $\{u_k\}$ be the Galerkin approximations satisfying (4.49) constructed as above with $\{w_i\}$ being the complete collection of eigenfunctions with eigenvalues $\{\lambda_i\}$ for $-\Delta$ on $H^1_0(\Omega)$. As before, we assume $\{w_i\}$ is orthogonal on $H^1_0(\Omega)$ and orthonormal on $L^2(\Omega)$.

By the uniqueness theorem, the weak solution is obtained as the limit of the Galerkin approximations $\{u_k\}$. We prove the theorem by deriving the same estimates for these approximate solutions independent of k.

1. By energy estimates (4.50), we have

$$\max_{t\in[0,T]}(\|u_k(t)\|_{H_0^1(\Omega)} + \|u_k'(t)\|_{L^2(\Omega)}) \le C(\|f\|_{L^2(\Omega_T)} + \|g\|_{H_0^1(\Omega)} + \|h\|_{L^2(\Omega)}) \tag{4.63}$$

and thus we deduce (4.61).

2. Similar to the parabolic case, we claim the following estimate for $\{u_k\}$: for each $t\in[0,T]$,

$$\|u_k(t)\|^2_{H^2(\Omega)} \le C(\|f(t)\|^2_{L^2(\Omega)} + \|u_k''(t)\|^2_{L^2(\Omega)} + \|u_k(t)\|^2_{L^2(\Omega)}), \tag{4.64}$$

where C is a constant independent of k. We would easily obtain this if u_k was itself a weak solution of (4.43), since then we could use the elliptic estimate to the elliptic equation $Lu_k(t) = f(t) - u_k''(t)$ on Ω. To prove (4.64), write the equation in (4.49) as

$$B[u_k(t), w_i; t] = (f(t) - u_k''(t), w_i) \quad (i = 1, 2, \cdots, k),$$

multiply this equation by $\lambda_i d_i(t)$ and sum over $i = 1, 2, \cdots, k$ to deduce

$$(Lu_k(t), -\Delta u_k(t)) = B[u_k(t), -\Delta u_k(t); t] = (f(t) - u_k''(t), -\Delta u_k(t)), \tag{4.65}$$

since $-\Delta u_k(t) \in H_0^1(\Omega)$. Then (4.64) follows from (4.65) by using Lemma 4.4.

3. Assume now the hypotheses of assertion (ii). We differentiate the equation in (4.49) with respect to t and set $\tilde u_k := u_k'$ to obtain

$$(\tilde u_k''(t), w_i) + B[\tilde u_k(t), w_i; t] = (f'(t), w_i) - \tilde B[u_k(t), w_i; t],$$

where $\tilde B$ is the bilinear form defined by

$$\tilde B[u, v; t] = \int_\Omega \Big(\sum_{i,j=1}^n a_{ij}'(x,t) D_i u D_j v + \sum_{i=1}^n b_i'(x,t) D_i u v + c'(x,t) uv \Big) dx.$$

Multiplying by $d_i''(t)$ and summing over $i = 1, 2, \cdots, k$, we discover

$$(\tilde u_k''(t), \tilde u_k'(t)) + B[\tilde u_k(t), \tilde u_k'(t); t] = (f'(t), \tilde u_k'(t)) - \tilde B[u_k(t), \tilde u_k'(t); t].$$

So

$$\begin{aligned}(\tilde u_k''(t), \tilde u_k'(t)) + A[\tilde u_k(t), \tilde u_k'(t); t] =& (f'(t), \tilde u_k'(t)) - \tilde B[u_k(t), \tilde u_k'(t); t] \\ &- C[\tilde u_k(t), \tilde u_k'(t); t],\end{aligned}$$

where, as above, we set $B[u, v; t] = A[u, v; t] + C[u, v; t]$, with A being the symmetric part of B and C the term involving no $D_i v$ terms. We write this equation as

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\Big(\|\tilde u_k'(t)\|^2_{L^2(\Omega)} + A[\tilde u_k(t), \tilde u_k(t); t]\Big) &= \frac{1}{2}\int_\Omega \sum_{i,j=1}^n a_{ij}' D_i \tilde u_k(t) D_j \tilde u_k(t) dx \\ &+ (f'(t), \tilde u_k'(t)) - \tilde B[u_k(t), \tilde u_k'(t); t] - C[\tilde u_k(t), \tilde u_k'(t); t].\end{aligned} \tag{4.66}$$

Notice that, for $u \in H_0^1(\Omega) \cap H^2(\Omega), v \in H_0^1(\Omega)$, integration by parts yields

$$\tilde B[u, v; t] = \int_\Omega v \Big(\sum_{i,j=1}^n (-D_j a_{ij}' D_i u - a_{ij}' D_{ij} u) + \sum_{i=1}^n b_i'(x,t) D_i u + c'(x,t) u \Big) dx,$$

and, hence, for a.e. $t \in [0,T]$ and all $u \in H^2(\Omega),\ v \in H_0^1(\Omega)$,

$$|\tilde B[u, v; t]| \le C(\|u\|^2_{H^2(\Omega)} + \|v\|^2_{L^2(\Omega)}).$$

Moreover, for all $u, v \in H_0^1(\Omega)$,

$$|C[u,v;t]| \le C(\|u\|^2_{H_0^1(\Omega)} + \|v\|^2_{L^2(\Omega)}) \le C(A[u,u;t] + \|v\|^2_{L^2(\Omega)}).$$

Using these estimates in (4.66), we deduce from (4.64)

$$\tag{4.67} \begin{aligned} &\frac{d}{dt}\left(\|\tilde{u}_k'(t)\|^2_{L^2(\Omega)} + A[\tilde{u}_k(t), \tilde{u}_k(t);t]\right) \\ &\le C(\|\tilde{u}_k'(t)\|^2_{L^2(\Omega)} + A[\tilde{u}_k(t), \tilde{u}_k(t);t] + \|u_k(t)\|^2_{H^2(\Omega)} + \|f'(t)\|^2_{L^2(\Omega)}) \\ &\le C(\|\tilde{u}_k'(t)\|^2_{L^2(\Omega)} + A[\tilde{u}_k(t), \tilde{u}_k(t);t] + \|f(t)\|^2_{L^2(\Omega)} + \|u_k(t)\|^2_{L^2(\Omega)} + \|f'(t)\|^2_{L^2(\Omega)}). \end{aligned}$$

Hence Gronwall's inequality implies

$$\begin{aligned} &\|\tilde{u}_k'(t)\|^2_{L^2(\Omega)} + A[\tilde{u}_k(t), \tilde{u}_k(t);t] \\ &\le C\left(\|\tilde{u}_k'(0)\|^2_{L^2(\Omega)} + A[\tilde{u}_k(0), \tilde{u}_k(0);0]) + \int_0^T (\|f(t)\|^2_{L^2(\Omega)} + \|u_k(t)\|^2_{L^2(\Omega)} + \|f'(t)\|^2_{L^2(\Omega)})dt\right) \\ &\le C(\|\tilde{u}_k'(0)\|^2_{L^2(\Omega)} + \|\tilde{u}_k(0)\|^2_{H_0^1(\Omega)} + \|f\|^2_{H^1(0,T;L^2(\Omega))} + \|g\|^2_{H_0^1(\Omega)} + \|h\|^2_{L^2(\Omega)}). \end{aligned}$$

Recall that $\tilde{u}_k = u_k'$ and

$$\|u_k'(0)\|_{H_0^1(\Omega)} \le C\|h\|_{H_0^1(\Omega)}, \ \|u_k''(0)\|_{L^2(\Omega)} \le C(\|f(0)\|_{L^2(\Omega)} + \|u_k(0)\|_{H^2(\Omega)})$$

to simplify the previous estimate as

$$\|u_k''(t)\|^2_{L^2(\Omega)} + \|u_k'(t)\|^2_{H_0^1(\Omega)} \le C(\|u_k(0)\|^2_{H^2(\Omega)} + \|f\|^2_{H^1(0,T;L^2(\Omega))} + \|g\|^2_{H_0^1(\Omega)} + \|h\|^2_{H^1(\Omega)}).$$

Finally as above, $\|u_k(0)\|_{H^2(\Omega)} \le C\|g\|_{H^2(\Omega)}$, from which and (4.64), we deduce

$$\tag{4.68} \begin{aligned} &\sup_{t\in[0,T]} (\|u_k''(t)\|^2_{L^2(\Omega)} + \|u_k(t)\|^2_{H^2(\Omega)} + \|u_k'(t)\|^2_{H_0^1(\Omega)}) \\ &\qquad \le C(\|f\|^2_{H^1(0,T;L^2(\Omega))} + \|g\|^2_{H_0^1(\Omega)} + \|h\|^2_{H^1(\Omega)}). \end{aligned}$$

4. As in the earlier proof for the parabolic equations, we can deduce the estimate for $u''' \in L^2(0,T;H^{-1}(\Omega))$ in terms of the right-hand side of (4.68). ☐

Theorem 4.16. (Higher regularity) *Let $m \in \{0,1,2,\cdots\}$. Assume*

$$g \in H^{m+1}(\Omega), \ \ h \in H^m(\Omega), \quad \frac{d^k f}{dt^k} \in L^2(0,T;H^{m-k}(\Omega)) \ \ (k = 0,1,\cdots,m).$$

*Suppose the following m-**th order compatibility conditions** hold:*

$$\tag{4.69} \begin{cases} g_0 := g \in H_0^1(\Omega), \ h_1 := h \in H_0^1(\Omega), \cdots, \\ g_{2l} := \frac{d^{2l-2}f}{dt^{2l-2}}(0) - Lg_{2l-2} \in H_0^1(\Omega) \ \ (\ if\ m = 2l), \\ h_{2l+1} := \frac{d^{2l-1}f}{dt^{2l-1}}(0) - Lh_{2l-1} \in H_0^1(\Omega) \ \ (\ if\ m = 2l+1). \end{cases}$$

Then

$$\frac{d^k u}{dt^k} \in L^\infty(0,T;H^{m+1-k}(\Omega)) \quad (k = 0,1,2,\cdots .m+1),$$

with the estimate

$$\tag{4.70} \sum_{k=0}^{m+1} \|\frac{d^k u}{dt^k}\|_{L^\infty(0,T;H^{m-k+1}(\Omega))} \le C\left(\sum_{k=0}^{m} \|\frac{d^k f}{dt^k}\|_{L^2(0,T;H^{m-k}(\Omega))} + \|g\|_{H^{m+1}(\Omega)} + \|h\|_{H^m(\Omega)}\right).$$

Proof. Again use induction on m and differentiate the equation with respect to t. Details are referred to Evans's book. ☐

Remark 4.8. The condition on f implies

$$f(0) \in H^{m-1}(\Omega),\ f'(0) \in H^{m-2}(\Omega),\ \cdots, f^{(m-2)}(0) \in H^1(\Omega)$$

and consequently

$$g_0 \in H^{m+1}(\Omega),\ h_1 \in H^m(\Omega),\quad g_2 \in H^{m-1}(\Omega),\ h_3 \in H^{m-2}(\Omega),$$

$$\cdots, g_{2l} \in H^1(\Omega)\ \ (\text{if } m = 2l), h_{2l+1} \in H^1(\Omega)\ \ (\text{if } m = 2l+1).$$

The compatibility conditions are precisely the requirements that each of these functions is zero on $\partial\Omega$.

Theorem 4.17. **(Smoothness of weak solution)** *Assume $g,\ h \in C^\infty(\bar{\Omega}),\ f \in C^\infty(\overline{\Omega_T})$, and the m-th-order compatibility conditions hold for all $m = 0, 1, \cdots$. Then problem* (4.43) *has a unique solution $u \in C^\infty(\overline{\Omega_T})$.*

4.2.6. Propagation of Disturbances. So far our study of hyperbolic equations has much paralleled our treatment of parabolic equations, using the Galerkin method. However, we learned that the classical solution to a second-order parabolic equation has the maximum principle, which implies an **infinite propagation speed** of initial disturbances for such equations. We now study a property for second-order hyperbolic equations that is totally the opposite phenomenon, namely the **finite propagation speed** of initial disturbances.

For simplicity, we study the operator of nondivergence form

$$Lu = -\sum_{i,j=1}^n a_{ij}(x) D_{ij} u,$$

where the coefficients a_{ij} are smooth, independent of time,$a_{ij} = a_{ji}$, and satisfy the usual uniform ellipticity condition.

Assume $q(x)$ is a continuous function on $\mathbb{R}^n$ and smooth in $\mathbb{R}^n \setminus \{x_0\}$, satisfying

$$\left\{ \begin{array}{l} q(x) > 0 \ \text{ in } \mathbb{R}^n \setminus \{x_0\}, \quad q(x_0) = 0, \\ \displaystyle\sum_{i,j=1}^n a_{ij}(x) D_i q D_j q \le 1 \quad \text{in } \mathbb{R}^n \setminus \{x_0\}. \end{array} \right. \tag{4.71}$$

Given a $t_0 > 0$, define the cone-like domain with vertex (x_0, t_0)

$$K = \{(x,t) \in \mathbb{R}^n \times (0, t_0) \mid\ q(x) < t_0 - t\}.$$

For each $0 < t < t_0$, define

$$K_t = \{x \in \mathbb{R}^n \mid\ q(x) < t_0 - t\}.$$

Theorem 4.18. **(Finite propagation speed)** *Let $u = u(x,t)$ be a smooth solution of*

$$u_{tt} + Lu = 0 \quad \text{in } \mathbb{R}^n \times (0, \infty).$$

If $u = u_t \equiv 0$ on K_0, then $u \equiv 0$ within K.

Proof. 1. Define the *energy*

$$e(t) = \frac{1}{2} \int_{K_t} \left(u_t^2 + \sum_{i,j=1}^n a_{ij} D_i u D_j u \right) dx \quad (0 \le t \le t_0).$$

We compute $e'(t)$. In order to do so, note that if $f(x,t)$ is continuous in x and smooth in t then

$$\frac{d}{dt}\left(\int_{K_t} f(x,t)dx\right) = \int_{K_t} f(x,t)dx - \int_{\partial K_t} \frac{f(x,t)}{|\nabla q(x)|}\, dS,$$

according to the **co-area formula**.

2. Therefore, we compute

$$\begin{aligned} e'(t) &= \int_{K_t} \Big(u_t u_{tt} + \sum_{i,j=1}^n a_{ij} D_i u D_j u_t\Big)\, dx - \frac{1}{2}\int_{\partial K_t} \Big(u_t^2 + \sum_{i,j=1}^n a_{ij} D_i u D_j u\Big) \frac{1}{|\nabla q|}\, dS \\ &:= A - B. \end{aligned}$$

Using $a_{ij} D_i u D_j u_t = D_j(a_{ij} u_t D_i u) - u_t D_j(a_{ij} D_i u)$ and integration by parts, we have

$$\begin{aligned} A &= \int_{K_t} u_t \Big(u_{tt} - \sum_{i,j=1}^n D_j(a_{ij} D_i u)\Big) dx + \int_{\partial K_t} \sum_{i,j=1}^n a_{ij} \nu^j u_t D_i u dS \\ &= -\int_{K_t} u_t \Big(\sum_{i,j=1}^n D_j a_{ij} D_i u\Big)\, dx + \int_{\partial K_t} \sum_{i,j=1}^n a_{ij} \nu^j u_t D_i u dS, \end{aligned} \tag{4.72}$$

where $\nu = (\nu^1, \cdots, \nu^n)$ is the outer unit normal to ∂K_t. Since on ∂K_t, $q(x) = t_0 - t$, we have $\nu = \frac{\nabla q}{|\nabla q|}$ on ∂K_t; that is, $\nu^j = D_j q / |\nabla q|$ on ∂K_t. Since matrix (a_{ij}) is symmetric and positive definite, for each $x \in \mathbb{R}^n$, the form $\langle \xi, \eta \rangle = \sum_{i,j=1}^n a_{ij}(x) \xi_i \eta_j$ $(\xi,\ \eta \in \mathbb{R}^n)$ defines an inner product on $\mathbb{R}^n$, with norm $\|\xi\| = \langle \xi, \xi \rangle^{1/2}$; hence, the Cauchy-Schwarz inequality $|\langle \xi, \eta \rangle| \leq \|\xi\| \|\eta\|$ implies

$$\begin{aligned} \left|\sum_{i,j=1}^n a_{ij}(x) \nu^j D_i u\right| &\leq \left(\sum_{i,j=1}^n a_{ij}(x) D_i u D_j u\right)^{1/2} \left(\sum_{i,j=1}^n a_{ij}(x) \nu^i \nu^j\right)^{1/2} \\ &= \left(\sum_{i,j=1}^n a_{ij}(x) D_i u D_j u\right)^{1/2} \left(\sum_{i,j=1}^n a_{ij}(x) \frac{D_i q D_j q}{|\nabla q|^2}\right)^{1/2} \\ &\leq \left(\sum_{i,j=1}^n a_{ij}(x) D_i u D_j u\right)^{1/2} \frac{1}{|\nabla q|} \quad (x \in \partial K_t), \end{aligned}$$

by (4.71). Returning to (4.72), we have

$$\begin{aligned} |A| &\leq Ce(t) + \int_{\partial K_t} \left(\sum_{i,j=1}^n a_{ij}(x) D_i u D_j u\right)^{1/2} \frac{|u_t|}{|\nabla q|}\, dS \\ &\leq Ce(t) + \frac{1}{2}\int_{\partial K_t} \left(u_t^2 + \sum_{i,j=1}^n a_{ij}(x) D_i u D_j u\right) \frac{1}{|\nabla q|}\, dS \\ &= Ce(t) + B. \end{aligned}$$

3. Therefore, we deduce

$$e'(t) = A - B \leq Ce(t) + B - B = Ce(t) \quad (0 < t < t_0).$$

Since $e(0) = 0$ and $e(t) \geq 0$, we deduce from Gronwall's inequality, that $e(t) \equiv 0$ for all $0 \leq t \leq t_0$. This proves $u \equiv 0$ in K. $\square$

4.3. Hyperbolic Systems of First-order Equations

We broaden our study of hyperbolic PDE to the first-order PDE systems.

4.3.1. Notations and Definitions. Consider systems of linear first-order PDE having the form

$$\mathbf{u}_t + \sum_{j=1}^{n} \mathbf{B}_j(x,t) D_j \mathbf{u} = \mathbf{f} \quad \text{in } \mathbb{R}^n \times (0,\infty), \tag{4.73}$$

subject to the initial condition

$$\mathbf{u} = \mathbf{g} \quad \text{on } \mathbb{R}^n \times \{t=0\}. \tag{4.74}$$

The unknown is $\mathbf{u}\colon \mathbb{R}^n \times [0,\infty) \to \mathbb{R}^m$, $\mathbf{u} = (u^1, \cdots, u^m)$, and the functions $\mathbf{B}_j\colon \mathbb{R}^n \times [0,\infty) \to \mathbb{M}^{m\times m}$ $(j = 1, 2, \cdots, n)$, $\mathbf{f}\colon \mathbb{R}^n \times [0,\infty) \to \mathbb{R}^m$ and $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$ are given,

Definition 4.9. (i) The system of PDE (4.73) is called a **hyperbolic system** if the $m \times m$ matrix

$$\mathbf{B}(x,t;\xi) = \sum_{j=1} \xi_j \mathbf{B}_j(x,t)$$

is diagonalizable for each $x, \xi \in \mathbb{R}^n$, $t \geq 0$. In other words, (4.73) is a **hyperbolic system** if for each x, ξ, t the matrix $\mathbf{B}(x,t;y)$ defined above has m real eigenvalues

$$\lambda_1(x,t;\xi) \leq \lambda_2(x,t;\xi) \leq \cdots \leq \lambda_m(x,t;\xi)$$

and corresponding eigenvectors $\{\mathbf{r}_k(x,t;\xi)\}_{k=1}^m$ that form a basis of $\mathbb{R}^m$.

(ii) We say (4.73) is a **symmetric hyperbolic system** if $\mathbf{B}_j(x,t)$ is symmetric for each $x \in \mathbb{R}^n$, $t \geq 0$.

(iii) The system is called **strictly hyperbolic** if if for each x, ξ, t the matrix $\mathbf{B}(x,t;\xi)$ defined above has m distinct real eigenvalues

$$\lambda_1(x,t;\xi) < \lambda_2(x,t;\xi) < \cdots < \lambda_m(x,t;\xi).$$

4.3.2. Vanishing Viscosity Method. We study the initial value problem (4.73),(4.74), with

$$\begin{aligned} &\mathbf{B}_j \in C^2(\mathbb{R}^n \times [0,T]; \mathbb{M}^{m\times m}) \text{ is symmetric}, \\ &\sup_{\mathbb{R}^n\times[0,T]} (|\mathbf{B}_j| + |D_{x,t}\mathbf{B}_j| + |D^2_{x,t}\mathbf{B}_j|) < \infty, \\ &\mathbf{g} \in H^1(\mathbb{R}^n;\mathbb{R}^m), \quad \mathbf{f} \in H^1(\mathbb{R}^n \times (0,T);\mathbb{R}^m). \end{aligned} \tag{4.75}$$

In this section, we do not need the hyperbolicity of the system. We define the bilinear form

$$B[\mathbf{u},\mathbf{v};t] := \Big(\sum_{j=1}^{n} \mathbf{B}_j(\cdot,t) D_j \mathbf{u}, \mathbf{v}\Big) = \int_{\mathbb{R}^n} \sum_{j=1}^{n} (\mathbf{B}_j(x,t) D_j \mathbf{u}(x)) \cdot \mathbf{v}(x)\, dx \tag{4.76}$$

for all $\mathbf{u}, \mathbf{v} \in H^1(\mathbb{R}^n;\mathbb{R}^m)$, $t \in [0,T]$.

Definition 4.10. We say a function

$$\mathbf{u} \in L^2(0,T; H^1(\mathbb{R}^n;\mathbb{R}^m)), \ \text{ with } \ \mathbf{u}' \in L^2(0,T; L^2(\mathbb{R}^n;\mathbb{R}^m)),$$

is a **weak solution** of the initial value problem (4.73), (4.74) provided

(i) $(\mathbf{u}', \mathbf{v}) + B[\mathbf{u},\mathbf{v};t] = (\mathbf{f},\mathbf{v})$ for each $\mathbf{v} \in H^1(\mathbb{R}^n;\mathbb{R}^m)$ and a.e. $t \in [0,T]$, and

(ii) $\mathbf{u}(0) = \mathbf{g}$. Again, by regularity, $\mathbf{u} \in C([0,T]; L^2(\mathbb{R}^n;\mathbb{R}^m))$, so $\mathbf{u}(0)$ is well-defined.

We shall use the **vanishing viscosity method** to prove the existence of weak solution. To this end, we approximate the initial value problem by the parabolic problem

$$\begin{cases} \mathbf{u}_t - \varepsilon\Delta\mathbf{u} + \sum_{j=1}^n \mathbf{B}_j D_j\mathbf{u} = \mathbf{f} & \text{in } \mathbb{R}^n \times (0,T], \\ \mathbf{u} = \mathbf{g}^\epsilon & \text{on } \mathbb{R}^n \times \{t=0\}, \end{cases} \tag{4.77}$$

for $0 < \varepsilon \le 1$, $\mathbf{g}_\varepsilon := \eta_\varepsilon \star \mathbf{g}$. The second-order term $-\varepsilon\Delta\mathbf{u}$ is called the **viscosity term**, which tends to regularize the original first-order system.

Theorem 4.19. (Existence of approximate solutions) *For each* $0 < \varepsilon \le 1$, *there exists a unique solution* $\mathbf{u} = \mathbf{u}_\varepsilon$ *of* (4.77), *with*

$$\mathbf{u}_\varepsilon \in L^2(0,T;H^3(\mathbb{R}^n;\mathbb{R}^m)), \quad \mathbf{u}_\varepsilon' \in L^2(0,T;H^1(\mathbb{R}^n;\mathbb{R}^m)).$$

Proof. 1. Set $X = L^\infty(0,T;H^1(\mathbb{R}^n;\mathbb{R}^m))$. For each $\mathbf{v} \in X$, consider the linear system

$$\begin{cases} \mathbf{u}_t - \varepsilon\Delta\mathbf{u} = \mathbf{f} - \sum_{j=1}^n \mathbf{B}_j D_j\mathbf{v} & \text{in } \mathbb{R}^n \times (0,T], \\ \mathbf{u} = \mathbf{g}^\epsilon & \text{on } \mathbb{R}^n \times \{t=0\}. \end{cases} \tag{4.78}$$

The right-hand side is bounded in L^2, there exists a unique solution $\mathbf{u} \in L^2(0,T;H^2(\mathbb{R}^n;\mathbb{R}^m))$, with $\mathbf{u}' \in L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^m))$. This solution $\mathbf{u}$ can be expressed by the Duhamel formula using the heat kernel. From this we can also show that $\mathbf{u} \in X = L^\infty(0,T;H^1(\mathbb{R}^n;\mathbb{R}^m))$. Hence we define a map $S\colon X \to X$ by setting $\mathbf{u} = S(\mathbf{v})$. Let $\mathbf{v}_1 \in X$ and $\mathbf{u}_1 = S(\mathbf{v}_1)$. Set $\mathbf{w} = \mathbf{u} - \mathbf{u}_1$ and $\mathbf{z} = \mathbf{v} - \mathbf{v}_1$. Then

$$\begin{cases} \mathbf{w}_t - \varepsilon\Delta\mathbf{w} = -\sum_{j=1}^n \mathbf{B}_j D_j\mathbf{z} & \text{in } \mathbb{R}^n \times (0,T], \\ \mathbf{w} = 0 & \text{on } \mathbb{R}^n \times \{t=0\}. \end{cases}$$

From the representation of $\mathbf{w}$ in terms of the heat kernel and $\sum_{j=1}^n \mathbf{B}_j D_j\mathbf{z}$, we have

$$\begin{aligned} \|\mathbf{w}\|_{L^\infty(0,T;H^1(\mathbb{R}^n;\mathbb{R}^m))} &\le C(\varepsilon)\|\sum_{j=1}^n \mathbf{B}_j D_j\mathbf{z}\|_{L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^m))} \\ &\le C(\varepsilon)\|\mathbf{z}\|_{L^2(0,T;H^1(\mathbb{R}^n;\mathbb{R}^m))} \\ &\le C(\varepsilon)T^{1/2}\|\mathbf{z}\|_{L^\infty(0,T;H^1(\mathbb{R}^n;\mathbb{R}^m))}. \end{aligned}$$

Thus $\|\mathbf{w}\|_X \le C(\varepsilon)T^{1/2}\|\mathbf{z}\|_X$; that is,

$$\|S(\mathbf{v}) - S(\mathbf{v}_1)\|_X \le C(\varepsilon)T^{1/2}\|\mathbf{v} - \mathbf{v}_1\|_X. \tag{4.79}$$

3. If $C(\varepsilon)T^{1/2} < 1$ then S is a strict contraction on X; hence it has a unique fixed point $\mathbf{u} = \mathbf{u}_\varepsilon$: $S(\mathbf{u}_\varepsilon) = \mathbf{u}_\varepsilon$. Then $\mathbf{u} = \mathbf{u}_\varepsilon$ solves (4.77) for such a $T > 0$. If $C(\varepsilon)T^{1/2} \ge 1$, then we choose $0 < T_1 < T$ so that $C(\varepsilon)T_1^{1/2} < 1$ and repeat the above argument on the time intervals $[0,T_1]$, $[T_1, 2T_1]$, etc, to obtain a weak solution $\mathbf{u} = \mathbf{u}_\varepsilon$ for all $T > 0$. Finally the high regularity of such a solution $\mathbf{u}$ follows from parabolic regularity theory. □

Theorem 4.20. (Energy estimates) *There exists a constant* C, *depending only on* n *and the coefficients,such that*

$$\begin{aligned} &\max_{t\in[0,T]} (\|\mathbf{u}_\varepsilon(t)\|_{H^1(\mathbb{R}^n;\mathbb{R}^m)} + \|\mathbf{u}_\varepsilon'(t)\|_{L^2(\mathbb{R}^n;\mathbb{R}^m)}) \\ &\qquad \le C(\|\mathbf{g}\|_{H^1(\mathbb{R}^n;\mathbb{R}^m)} + \|\mathbf{f}\|_{L^2(0,T;H^1(\mathbb{R}^n;\mathbb{R}^m)} + \|\mathbf{f}'\|_{L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^m)}) \end{aligned} \tag{4.80}$$

for all $0 < \varepsilon \le 1$.

Proof. 1. We compute

$$\frac{d}{dt}\left(\frac{1}{2}\|\mathbf{u}_\varepsilon(t)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)}\right) = (\mathbf{u}_\varepsilon(t), \mathbf{u}'_\varepsilon(t)) = \left(\mathbf{u}_\varepsilon, \mathbf{f}(t) + \varepsilon\Delta\mathbf{u}_\varepsilon(t) - \sum_{j=1}^n \mathbf{B}_j D_j \mathbf{u}_\varepsilon(t)\right).$$

Note that

$$(\mathbf{u}_\varepsilon(t), \varepsilon\Delta\mathbf{u}_\varepsilon(t)) = -\varepsilon\|\nabla\mathbf{u}_\varepsilon(t)\|^2_{L^2(\mathbb{R}^n)} \le 0 \quad (0 < t \le T).$$

2. Suppose $\mathbf{v} \in C_0^\infty(\mathbb{R}^n;\mathbb{R}^m)$. Then, by the symmetry of $\mathbf{B}_j$ (this is the only place the symmetry assumption is used: if $\mathbf{B}$ is symmetric then $\mathbf{Ba}\cdot\mathbf{b} = \mathbf{Bb}\cdot\mathbf{a}$),

$$\begin{aligned} B[\mathbf{v},\mathbf{v};t] &= \Big(\mathbf{v}, \sum_{j=1}^n \mathbf{B}_j D_j \mathbf{v}\Big) = \int_{\mathbb{R}^n} \sum_{j=1}^n (\mathbf{B}_j D_j \mathbf{v})\cdot\mathbf{v}\,dx \\ &= \frac{1}{2}\sum_{j=1}^n \int_{\mathbb{R}^n} D_j[(\mathbf{B}_j\mathbf{v})\cdot\mathbf{v}]dx - \frac{1}{2}\sum_{j=1}^n \int_{\mathbb{R}^n} [(D_j\mathbf{B}_j)\mathbf{v})\cdot\mathbf{v}]dx \\ &= -\frac{1}{2}\sum_{j=1}^n \int_{\mathbb{R}^n} [(D_j\mathbf{B}_j)\mathbf{v})\cdot\mathbf{v}]dx. \end{aligned}$$

Hence, for all $\mathbf{v} \in C_0^\infty(\mathbb{R}^n;\mathbb{R}^m)$,

$$|B[\mathbf{v},\mathbf{v};t]| = |\Big(\mathbf{v}, \sum_{j=1}^n \mathbf{B}_j D_j \mathbf{v}\Big)| \le C\|\mathbf{v}\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)}. \tag{4.81}$$

By approximation, (4.81) holds for all $\mathbf{v} \in H^1(\mathbb{R}^n;\mathbb{R}^m)$. Hence

$$|\Big(\mathbf{u}_\varepsilon(t), \sum_{j=1}^n \mathbf{B}_j D_j \mathbf{u}_\varepsilon(t)\Big)| \le C\|\mathbf{u}_\varepsilon(t)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)}.$$

We therefore deduce

$$\frac{d}{dt}\left(\frac{1}{2}\|\mathbf{u}_\varepsilon(t)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)}\right) \le C\Big(\|\mathbf{u}_\varepsilon(t)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)} + \|\mathbf{f}(t)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)}\Big).$$

So Gronwall's inequality and $\|\mathbf{g}_\varepsilon\|_{L^2} \le \|\mathbf{g}\|_{L^2}$ will yield

$$\max_{t\in[0,T]} \|\mathbf{u}_\varepsilon(t)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)} \le C\Big(\|\mathbf{g}\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)} + \|\mathbf{f}\|^2_{L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^m))}\Big). \tag{4.82}$$

3. Differentiating the equation with respect x_k, estimating $D_k\mathbf{u}_\varepsilon$ and summing $k = 1, 2, \cdots, n$, we deduce

$$\max_{t\in[0,T]} \|\nabla\mathbf{u}_\varepsilon(t)\|^2_{L^2(\mathbb{R}^n)} \le C\Big(\|\nabla\mathbf{g}\|^2_{L^2(\mathbb{R}^n)} + \|\mathbf{f}\|^2_{L^2(0,T;H^1(\mathbb{R}^n;\mathbb{R}^m))}\Big), \tag{4.83}$$

where we have used $\|\nabla\mathbf{g}_\varepsilon\|_{L^2} \le C\|\nabla\mathbf{g}\|_{L^2}$.

4. Next differentiating the equation with respect to t and setting $\mathbf{v} = \mathbf{u}'_\varepsilon$, we have

$$\begin{cases} \mathbf{v}_t - \varepsilon\Delta\mathbf{v} + \sum_{j=1}^n \mathbf{B}_j D_j \mathbf{v} = \mathbf{f}' - \sum_{j=1}^n \mathbf{B}'_j D_j D_j \mathbf{u}_\varepsilon & \text{in } \mathbb{R}^n \times (0,T], \\ \mathbf{v} = \mathbf{f} + \varepsilon\Delta\mathbf{g}_\varepsilon - \sum_{j=1}^n \mathbf{B}_j D_j \mathbf{g}_\varepsilon & \text{on } \mathbb{R}^n \times \{t=0\}. \end{cases} \tag{4.84}$$

Reasoning as above, we deduce

$$\begin{aligned} \max_{t\in[0,T]} \|\mathbf{u}'_\varepsilon(t)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)} &\le C\Big(\|\nabla\mathbf{g}\|^2_{L^2(\mathbb{R}^n)} + \varepsilon^2\|\Delta\mathbf{g}_\varepsilon\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)} \\ &\quad + \|\mathbf{f}(0)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)} + \|\mathbf{f}\|^2_{L^2(0,T;H^1(\mathbb{R}^n;\mathbb{R}^m))} + \|\mathbf{f}'\|^2_{L^2(0,T;L^2(\mathbb{R}^n;\mathbb{R}^m))}\Big). \end{aligned} \tag{4.85}$$

Now, using $\mathbf{g}_\varepsilon = \eta_\varepsilon \star \mathbf{g}$, we have

$$\|\Delta \mathbf{g}_\varepsilon\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)} \le \frac{C}{\varepsilon^2}\|\nabla \mathbf{g}\|^2_{L^2(\mathbb{R}^n)}.$$

Furthermore,

$$\|\mathbf{f}(0)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)} \le C\|\mathbf{f}\|^2_{H^1(0,T;L^2(\mathbb{R}^n;\mathbb{R}^m))}.$$

Combining all these estimates completes the proof. □

Theorem 4.21. (Existence of weak solution by vanishing viscosity) *There exists a weak solution to problem* (4.77) *as certain limit of* $\{\mathbf{u}_\varepsilon\}$ *along a sequence* $\varepsilon \to 0^+$.

Proof. This follows from the energy estimates above in exactly the same fashion as before. □

Theorem 4.22. (Uniqueness of weak solution) *A weak solution to problem* (4.77) *is unique.*

Proof. It suffices to show that the only weak solution to (4.77) with $\mathbf{f} \equiv \mathbf{g} \equiv 0$ is $\mathbf{u} \equiv 0$. To verify this, note that $(\mathbf{u}'(t), \mathbf{u}(t)) + B[\mathbf{u}(t), \mathbf{u}(t); t] = 0$ for a.e. $t \in [0,T]$ and, by (4.81),

$$|B[\mathbf{u}(t), \mathbf{u}(t); t]| \le C\|\mathbf{u}(t)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)},$$

so we have

$$\frac{d}{dt}\left(\|\mathbf{u}(t)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)}\right) \le C\|\mathbf{u}(t)\|^2_{L^2(\mathbb{R}^n;\mathbb{R}^m)},$$

whence Gronwall's inequality forces $\mathbf{u} \equiv 0$. □

4.3.3. Systems with Constant Coefficients. In this section, we study

$$\begin{cases} \mathbf{u}_t + \sum_{j=1}^n \mathbf{B}_j D_j \mathbf{u} = 0 & \text{in } \mathbb{R}^n \times (0,\infty), \\ \mathbf{u} = \mathbf{g} & \text{on } \mathbb{R}^n \times \{t=0\}. \end{cases} \tag{4.86}$$

Here we assume that the coefficients $\mathbf{B}_j$ are *constant* $m \times m$ matrices and that, for each $\xi \in \mathbb{R}^n$, the matrix

$$\mathbf{B}(\xi) = \sum_{j=1}^n \xi_j \mathbf{B}_j$$

has *all* real eigenvalues

$$\lambda_1(\xi) \le \lambda_2(\xi) \le \cdots \le \lambda_m(\xi).$$

There is no hypothesis concerning the eigenvectors and there is no symmetry assumption on $\mathbf{B}_j$. The weak notion of *hyperbolicity* lies in the assumption that matrix $\mathbf{B}(\xi)$ has all real eigenvalues for all $\xi \in \mathbb{R}^n$. We will apply the Fourier transform to solve the corresponding problem (4.86).

Theorem 4.23. (Existence of solution by Fourier transform) *Assume* $\mathbf{g} \in H^s(\mathbb{R}^n;\mathbb{R}^m)$ *with* $s > \frac{n}{2} + m$. *There exists a unique weak solution* $\mathbf{u} \in C^1(\mathbb{R}^n \times [0,\infty);\mathbb{R}^m)$ *to the initial value problem* (4.86).

Proof. If $\mathbf{u} \in L^2(\mathbb{R}^n;\mathbb{R}^m)$, $\mathbf{u} = (u^1, \cdots, u^m)$, we use

$$\hat{\mathbf{u}}(\xi) = (\hat{u}^1(\xi), \cdots, \hat{u}^m(\xi)),$$

where $\hat{v}(\xi)$ stands for the Fourier transform a function $v \in L^2(\mathbb{R}^n)$. If $\mathbf{u} = \mathbf{u}(x,t) \in L^2_x(\mathbb{R}^n\ \mathbb{R}^m)$ for each t, then we use $\hat{\mathbf{u}}(\xi, t)$ to denote the Fourier transform of $\mathbf{u}(\cdot, t)$; we do not transform with respect to t.

1. Taking the spatial Fourier transform, System (4.86) becomes

$$\hat{\mathbf{u}}_t(\xi,t) + i\mathbf{B}(\xi)\hat{u}(\xi,t) = 0, \quad \hat{\mathbf{u}}(\xi,0) = \hat{\mathbf{g}}(\xi). \tag{4.87}$$

For fixed $\xi \in \mathbb{R}^n$, we can solve (4.87) to find

$$\hat{\mathbf{u}}(\xi,t) = e^{-it\mathbf{B}(\xi)}\hat{\mathbf{g}}(\xi) \quad (\xi \in \mathbb{R}^n, \;\; t \geq 0). \tag{4.88}$$

Consequently, via the Fourier inverse transform,

$$\mathbf{u}(x,t) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ix\cdot\xi} e^{-it\mathbf{B}(\xi)}\hat{\mathbf{g}}(\xi)\, d\xi \quad (x \in \mathbb{R}^n, \;\; t \geq 0). \tag{4.89}$$

2. We need to verify that the formula (4.89) indeed defines a function $\mathbf{u} \in C^1(\mathbb{R}^n \times [0,\infty);\mathbb{R}^m)$ that is a solution to (4.86). Since $\mathbf{g} \in H^s(\mathbb{R}^n;\mathbb{R}^m)$, we have

$$h(\xi) := (1+|\xi|^s)\hat{\mathbf{g}}(\xi) \in L^2(\mathbb{R}^n;\mathbb{R}^m).$$

To show the integral in (4.89) converges, we must estimate $\|e^{-it\mathbf{B}(\xi)}\|$.

3. Let $\lambda_1(\xi), \cdots, \lambda_m(\xi)$ be the eigenvalues of $\mathbf{B}(\xi)$. Let Γ be the circle $\partial B(0,r)$ in the complex plane, traversed counterclockwise, with radius $r > 0$ so large that all $\lambda_j(\xi)$ lie inside Γ and $\mathrm{dist}(\lambda_j(\xi);\Gamma) \geq 2$. By Cauchy's theorem in complex analysis, we have

$$e^{-it\mathbf{B}(\xi)} = \frac{1}{2\pi i}\int_\Gamma e^{-itz}(zI - \mathbf{B}(\xi))^{-1}\, dz.$$

4. Define a new path Δ in the complex plane by

$$\Delta = \partial\left(\bigcup_{j=1}^m B(\lambda_j(\xi),1)\right),$$

traversed counterclockwise. Deforming Γ to Δ, we have

$$e^{-it\mathbf{B}(\xi)} = \frac{1}{2\pi i}\int_\Delta e^{-itz}(zI - \mathbf{B}(\xi))^{-1}\, dz.$$

Note that

$$|e^{-itz}| \leq e^t \quad (z \in \Delta),$$

and

$$|\det(zI - \mathbf{B}(\xi))| = |(z-\lambda_1(\xi))(z-\lambda_2(\xi))\cdots(z-\lambda_m(\xi))| \geq 1 \quad (z \in \Delta).$$

So we can estimate the inverse matrix by

$$\|(zI - \mathbf{B}(\xi))^{-1}\| \leq \|\mathrm{cof}(zI - \mathbf{B}(\xi))\| \leq C(1+|\xi|^{m-1} + \|\mathbf{B}(\xi)\|^{m-1}) \leq C(1+|\xi|^{m-1}),$$

where we used the fact $|\lambda_k(\xi)| \leq C|\xi|$ since $\mathbf{B}(\xi)$ is linear in ξ. Combining these estimates, we have

$$\|e^{-it\mathbf{B}(\xi)}\| \leq e^t(1+|\xi|^{m-1}) \quad (\xi \in \mathbb{R}^n).$$

5. Therefore,

$$\begin{aligned}\int_{\mathbb{R}^n} |e^{ix\cdot\xi}e^{-it\mathbf{B}(\xi)}\hat{\mathbf{g}}(\xi)|\, d\xi &\leq C\int_{\mathbb{R}^n} \frac{\|e^{-it\mathbf{B}(\xi)}\| h(\xi)}{1+|\xi|^s}\, d\xi \\ &\leq Ce^t\int_{\mathbb{R}^n} h(\xi)(1+|\xi|^{-s})(1+|\xi|^{m-1})\, d\xi \\ &\leq Ce^t\|h\|_{L^2(\mathbb{R}^n)}\|(1+|\xi|^{m-1-s})\|_{L^2(\mathbb{R}^n)} < \infty,\end{aligned}$$

since $s > \frac{n}{2} + m > \frac{n}{2} + m - 1$. Hence the integral in (4.89) converges, and it follows easily that the function $\mathbf{u}(x,t)$ defined is continuous on $\mathbb{R}^n \times [0,\infty)$.

6. To show $\mathbf{u}$ is $C^1(\mathbb{R}^n \times [0,\infty);\mathbb{R}^m)$, observe for $0 < |h| \le 1$,

$$\frac{\mathbf{u}(x,t+h)-\mathbf{u}(x,t)}{h} = \frac{1}{(2\pi)^{n/2}h}\int_{\mathbb{R}^n} e^{x\cdot\xi}(e^{-i(t+h)\mathbf{B}(\xi)} - e^{-it\mathbf{B}(\xi)})\hat{\mathbf{g}}(\xi)\,d\xi.$$

Since

$$e^{-i(t+h)\mathbf{B}(\xi)} - e^{-it\mathbf{B}(\xi)} = -i\int_t^{t+h} \mathbf{B}(\xi)e^{-is\mathbf{B}(\xi)}\,ds,$$

we can easily estimate

$$\left|\frac{1}{h}\left(e^{-i(t+h)\mathbf{B}(\xi)} - e^{-it\mathbf{B}(\xi)}\right)\right| \le Ce^{t+1}(1+|\xi|^m).$$

From this we have

$$\left|\frac{\mathbf{u}(x,t+h)-\mathbf{u}(x,t)}{h}\right| \le Ce^{t+1}\int_{\mathbb{R}^n} h(\xi)(1+|\xi|^m)(1+|\xi|^s)^{-1}\,d\xi < \infty,$$

since $s > \frac{n}{2}+m$. Therefore $\mathbf{u}_t$ exists and is continuous on $\mathbb{R}^n \times [0,\infty)$. A similar argument shows that $D_j\mathbf{u}$ exists and is continuous for all $j = 1, 2, \cdots, n$ (the proof is similar since ξ is like $\mathbf{B}(\xi)$). Furthermore, we can differentiate inside the integral in (4.89) to confirm that $\mathbf{u}$ solves the system (4.86). □

4.4. Semigroup Theory

Semigroup theory is an abstract study of first-order ordinary differential equations with values in Banach spaces, driven by linear, but possibly unbounded operators. The method provides an elegant alternative to some of the existence theory for evolution equations set forth above.

The whole idea of semigroups springs from properties of solutions of the elementary initial value problem in $u \in \mathbb{R}^m$:

$$\frac{du}{dt} = Au \ \ (t > 0), \quad u(0) = u_0 \tag{4.90}$$

where A is a constant $m \times m$ matrix. The solution of course is

$$u(t) = T(t)u_0 = e^{tA}u_0. \tag{4.91}$$

This operator $T(t)$ has some obvious properties:

$$\text{(a)} \quad \lim_{t\to 0} T(t) = 1, \qquad \text{(b)} \quad T(t_1)T(t_2) = T(t_1+t_2).$$

Effectively, $T(t)$ maps the initial data u_0 into the current value of the solution $u(t)$.

Before we discuss the general theory for (4.90) with A being a linear operator defined in a subspace of a Banach space, we review some definitions and elementary properties.

4.4.1. Definitions and Elementary Properties. Let X, Y be normed spaces. A linear operator $T : \mathcal{D}(T) \subset X \to Y$ is said to be **closed** if whenever $\{x_n\} \subset \mathcal{D}(T)$ is a sequence satisfying

$$x_n \to x, \ Tx_n \to y$$

then $x \in \mathcal{D}(T)$ and $Tx = y$.

Example 4.24. Let $L : \mathcal{D}(L) \subset L^2(0,1) \to L^2(0,1)$ be the differential operator $L = d/dx$, where $\mathcal{D}(L) = H_0^1(0,1)$. To show that L is closed, let $u_n \to u$, $u_n' \to f$ in $L^2(0,1)$, where $u_n \in \mathcal{D}(L)$. Passing to the limit in

$$\int_0^1 u_n v' dx = -\int_0^1 u_n' v dx \quad \text{for all} \quad v \in C_0^\infty(0,1)$$

we see, by the definition of weak derivative, that $u' = f$ and $u_n \to u$ in $H^1(0,1)$. Since $H_0^1(0,1)$ is closed, we have $u \in H_0^1(0,1)$.

Example 4.25. Let Ω be a bounded domain in $\mathbb{R}^n$ and let $X, Y = L^2(\Omega)$. Let

$$\mathcal{D}(L) = H^2(\Omega) \cap H_0^1(\Omega)$$

$$Lu = \Delta u, \quad u \in \mathcal{D}(L).$$

Note that we are considering L as an operator on $L^2(\Omega)$. Clearly L is densely defined. It is unbounded, for if we consider $\{\varphi_n\}$, the sequence of eigenfunctions of $-\Delta$, then $\|\varphi_n\|_2 = 1$ while $\|L\varphi_n\|_2 = \lambda_n \to \infty$ as $n \to \infty$.

To see that $L : \mathcal{D}(L) \subset L^2(\Omega) \to L^2(\Omega)$ is a closed operator, let $u_n \in \mathcal{D}(L)$ with $u_n \to u$, $Lu_n \to f$. Applying the estimate $\|u\|_{2,2} \leq c\|Lu\|_2$ to $u_n - u_m$, it follows that $\{u_n\}$ is a Cauchy sequence in $H^2(\Omega)$ and thus $\|u_n - v\|_{2,2} \to 0$ for some $v \in H^2(\Omega)$. Clearly $u = v$ and $u \in \mathcal{D}(L)$. Since $L : H^2(\Omega) \to L^2(\Omega)$ is continuous, $Lu_n \to Lu$ which yields $Lu = f$. Hence L is closed.

Theorem 4.26. (Closed Graph Theorem) *If X, Y are Banach spaces and if $T : \mathcal{D}(T) \subset X \to Y$ is a closed linear operator with closed domain, then T is bounded.*

Corollary 4.27. *If X, Y are Banach spaces and if $T : \mathcal{D}(T) \subset X \to Y$ is a closed linear operator which is one-to-one, then T^{-1} is bounded iff $\mathcal{R}(T)$ is closed. (In particular, if T is 1-1 and onto, then T^{-1} is bounded.)*

Definition 4.11. Let X be a Banach space and let $T\colon \mathcal{D}(T) \subset X \to X$ be a closed operator on X.

(i) We say a real number λ belongs to the **resolvent set** $\rho(T)$ of T, provided the operator

$$\lambda I - T\colon \mathcal{D}(T) \to X$$

is one-to-one and onto.

(ii) If $\lambda \in \rho(T)$, the **resolvent operator** $R_\lambda\colon X \to X$ is defined by

$$R_\lambda u := (\lambda I - T)^{-1} u \quad (u \in X);$$

that is, $R_\lambda = (\lambda I - T)^{-1}$.

Remark 4.12. (i) Note that, since T is closed, by (4.27), R_λ is a bounded linear operator on X if $\lambda \in \rho(T)$. Furthermore,

$$TR_\lambda u = R_\lambda T u \quad (u \in \mathcal{D}(T)).$$

(ii) We have the **resolvent identity**:

$$R_\lambda - R_\mu = (\mu - \lambda) R_\lambda R_\mu, \quad R_\lambda R_\mu = R_\mu R_\lambda \quad (\lambda\ \mu \in \rho(T)).$$

4.4.2. C_0 Semigroups of Operators.

Definition 4.13. Let X be a Banach space. A family $\{T(t)\} \subset \mathcal{B}(X)$ $(0 \le t < \infty)$ is called a **strongly continuous semigroup of operators** if

(i) $T(t)T(s) = T(t+s), \quad t, s \ge 0 \quad$ (semigroup property)

(ii) $T(0) = I$

(iii) For all $u \in X$, $T(t)u$ is strongly continuous in $t \in [0, \infty)$, i.e.,

$$\|T(t+h)u - T(t)u\| \to 0 \text{ as } h \to 0.$$

For simplicity we say that $T(t)$ is a C_0 semigroup.

Moreover, for fixed $u \in X$, we write $T(\cdot)u \in C([0,\infty); X)$. If in addition the map $t \to T(t)$ is continuous in the uniform operator topology, i.e., $\|T(t+h) - T(t)\| \to 0$ for $t, t+h \ge 0$, then the family $\{T(t)\}$ is called a **uniformly continuous semigroup**. If the C_0 semigroup $\{T(t)\}$ satisfies the property $\|T(t)\| \le 1$ for $t \ge 0$, then it is called a **contraction semigroup**.

Note that $T(t)$ and $T(s)$ commute as a consequence of (i), and that $T(t)u$ is also strongly continuous in u for each fixed $t \ge 0$. ($\|T(t)u - T(t)v\| \le \|T(t)\|\|u - v\|$.)

Example 4.28. Let $A \in \mathcal{B}(X)$ where X is a Banach space. Then the series $\sum_{n=0}^{\infty}(A^n/n!)t^n$ converges in the uniform operator topology for any real number t. In fact, set

$$S_n = \sum_{k=0}^{n}(A^k/k!)t^k$$

and observe that for $m < n$

$$\|S_n - S_m\| \le \sum_{k=m+1}^{n}(\|A\|^k/k!)|t|^k \to 0 \text{ as } m, n \to \infty.$$

Thus $\{S_n\}$ converges to a bounded linear operator, in the uniform operator topology, which we denote by e^{tA}. From the above estimate we see that

$$\|e^{tA}\| \le e^{|t|\|A\|}.$$

Let $u(t) = e^{tA}u_0$ where $u_0 \in X$. Since $e^{(t+s)A} = e^{tA}e^{sA}$, it follows that

$$\frac{u(t+h) - u(t)}{h} - Au = \left(\frac{e^{hA} - I}{h} - A\right)u.$$

However,

$$\begin{aligned}\left\|\frac{e^{hA} - I}{h} - A\right\| &\le \sum_{k=2}^{\infty}(\|A\|^k/k!)|h|^{k-1} \\ &= \frac{e^{|h|\|A\|} - 1}{|h|} - \|A\| \to 0 \quad \text{as } h \to 0.\end{aligned}$$

Hence $u'(t)$ exists for all t and equals Au, and so we have shown that $u(t) = e^{tA}u_0$ satisfies the Cauchy problem

$$\frac{du}{dt} = Au \ \ (t > 0), \quad u(0) = u_0.$$

Finally, we show that $\{e^{tA}\},\ t \geq 0$, is a uniformly continuous semigroup. In fact

$$\begin{aligned} \|e^{(t+h)A} - e^{tA}\| &= \|e^{tA}(e^{hA} - I)\| \\ &\leq e^{|t|\|A\|}(e^{|h|\|A\|} - 1) \to 0 \quad \text{as } h \to 0. \end{aligned}$$

Lemma 4.29. *If $\{T(t)\}$ is a C_0 semigroup, then there exist constants $\omega \geq 0$ and $M \geq 1$ such that*

$$\|T(t)\| \leq Me^{\omega t}, \quad t \geq 0$$

i.e., $\|T(t)\|$ grows slower than an exponential.

Proof. In fact, since the function $g(t) = \|T(t)u\|$ is continuous on $[0,1]$ for each fixed $u \in X$, we have $\sup_{t\in[0,1]} \|T(t)u\| < \infty$. Hence, by the Uniform Boundedness Principle, there is a constant $M > 0$ such that $\|T(t)\| \leq M$ for $t \in [0,1]$. Let $\omega = \log M$. Then $\omega \geq 0$, since $M \geq 1$ by virtue of $T(0) = I$. Now let t be given and let n be the least integer greater than or equal to t. By virtue of the semigroup property $(T(t/n))^n = T(t)$ and thus

$$\|T(t)\| = \|(T(t/n))^n\| \leq M^n \leq M^{t+1} = Me^{\omega t}.$$

□

4.4.3. Infinitesimal Generator. Let $\{T(t)\}$ be a C_0 semigroup on the Banach space X. For $h > 0$ we define the linear operator A_h by the formula

$$A_h u = \frac{T(h)u - u}{h}, \quad u \in X.$$

Definition 4.14. (i) Let $\mathcal{D}(A)$ be the set of all $u \in X$ for which $\lim_{h\to 0^+} A_h u$ exists. Define the operator A on $\mathcal{D}(A)$ by the relation

$$Au = \lim_{h\to 0^+} A_h u = \frac{d}{dh}T(h)u|_{h=0^+} \quad (u \in \mathcal{D}(A)).$$

The operator A is called the **infinitesimal generator** of the semigroup $\{T(t)\}$.

(ii) Given an operator A on $\mathcal{D}(A)$, we say that it generates a C_0 semigroup $\{T(t)\}$ if A coincides with the infinitesimal generator of $\{T(t)\}$.

Example 4.30. Clearly $A \in \mathcal{B}(X)$ is the infinitesimal generator of $\{e^{tA}\},\ t \geq 0$.

Theorem 4.31. *Let $\{T(t)\}$ be a C_0 semigroup on the Banach space X and let $A : \mathcal{D}(A) \to X$ be its infinitesimal generator. Then the following hold:*

(a) *$\mathcal{D}(A)$ is a subspace of X and A is a linear operator.*

(b) *If $u \in \mathcal{D}(A)$, then $T(t)u \in \mathcal{D}(A)$, $0 \leq t < \infty$, is strongly differentiable in t and*

$$(d/dt)T(t)u = AT(t)u = T(t)Au, \ \ t \geq 0 \tag{4.92}$$

(c) *If $u \in \mathcal{D}(A)$, then*

$$T(t)u - T(s)u = \int_s^t T(h)Au\, dh, \ \ t, s \geq 0$$

(d) *If $f(t)$ is a real-valued continuous function for $t \geq 0$, then*

$$\lim_{h\to 0} h^{-1} \int_t^{t+h} f(s)T(s)u\, ds = f(t)T(t)u, \ \ u \in X, \ t \geq 0$$

(e) *$\int_0^t T(s)u\, ds \in \mathcal{D}(A)$ and $T(t)u = u + A\int_0^t T(s)u\, ds$, $u \in X$*

(f) *$\overline{\mathcal{D}(A)} = X$ and A is a closed operator.*

Proof. (a) This follows directly from the definition since A_h is linear.

(b) Since $T(t)$ and $T(h)$ commute and $\|T(t)\| < \infty$, we have

$$A_h T(t)u = T(t)A_h u \to T(t)Au, \text{ as } h \to 0^+.$$

Hence

$$T(t)u \in \mathcal{D}(A), \quad AT(t)u = T(t)Au = D^+T(t)u.$$

Next we consider $D^-T(t)u$ if $t > 0$. Note that

$$\begin{aligned} \frac{T(t)u - T(t-h)u}{h} - T(t)Au &= T(t-h)\left(\frac{T(h)u - u}{h} - Au\right) \\ &+ (T(t-h) - T(t))Au. \end{aligned}$$

But

$$\|T(t-h)\left(\frac{T(h)u - u}{h} - Au\right)\| \le Me^{\omega t}\|\frac{T(h)u - u}{h} - Au\| \to 0$$

as $h \to 0^+$ and

$$\|(T(t-h) - T(t))Au\| \to 0 \text{ as } h \to 0^+$$

which implies the desired result.

(c) The abstract function $T(t)u$ is differentiable by (b) and its derivative $T(t)Au$ is continuous in t. The conclusion follows by integrating (4.92).

(d)

$$\|h^{-1}\int_t^{t+h} f(s)T(s)u\,ds - f(t)T(t)u\| = \|h^{-1}\int_t^{t+h}(f(s)T(s)u - f(t)T(t)u)\,ds\|$$

$$\le h^{-1}\int_t^{t+h}|f(s)|\|(T(s)u - T(t)u\|\,ds + h^{-1}\int_t^{t+h}\|T(t)u\||f(s) - f(t)|\,ds.$$

Since the functions $T(t)u$ and $f(t)$ are continuous in t, by choosing h small enough so that $\|(T(s)u - T(t)u\| < \varepsilon$ and $|f(s) - f(t)| < \varepsilon$ for $|t - s| < h$, the result easily follows.

(e) Let $h > 0$ and consider

$$\begin{aligned} \left(\frac{T(h) - I}{h}\right)\int_0^t T(s)u\,ds &= \frac{1}{h}\int_0^t (T(s+h)u - T(s)u)ds \\ &= \frac{1}{h}\left(\int_t^{t+h} T(s)u\,ds - \int_0^h T(s)u\,ds\right) \\ &\to T(t)u - u \end{aligned}$$

by (d). Thus $\int_0^t T(s)u\,ds \in \mathcal{D}(A)$ and (e) follows by the definition of A.

(f) Let $u \in X$. Then $\int_0^h T(s)u\,ds \in \mathcal{D}(A)$ and by (d), $\lim_{h\to 0} h^{-1}\int_0^h T(s)u\,ds = T(0)u = u$. Thus $\overline{\mathcal{D}(A)} = X$. Let $u_n \in \mathcal{D}(A)$ with $u_n \to u$, $Au_n \to v$ in X. Now

$$A_h u = \lim_{n\to\infty}\frac{T(h)u_n - u_n}{h} = \lim_{n\to\infty} h^{-1}\int_0^h T(s)Au_n\,ds$$

where the last term follows from (c). But $Au_n \to v$ and so

$$A_h u = h^{-1}\int_0^h T(s)v\,ds \to v \text{ as } h \to 0$$

by virtue of (d). Thus $u \in \mathcal{D}(A)$ and $Au = v$. □

Remark 4.15. Since the map $t \to T(t)Au$ is continuous, it follows from (b) that $T(t)u$ is continuously differentiable from $[0, \infty)$ with values in X. Also $T(t)u \in \mathcal{D}(A)$ as proved, so it has values in $\mathcal{D}(A)$ as well. Further, the continuous differentiability into X also proves the continuity into $\mathcal{D}(A)$ (with the graph norm).

4.4.4. Application to Abstract Cauchy Problems. As an application we shall prove that the abstract Cauchy problem

$$\frac{du}{dt} = Au \ (t \geq 0), \quad u(0) = u_0, \ \ u_0 \in \mathcal{D}(A) \tag{4.93}$$

has a unique solution.

Theorem 4.32. *Let $A : \mathcal{D}(A) \to X$ be the infinitesimal generator of a C_0 semigroup $\{T(t)\}$. Then the Cauchy problem* (4.93) *has the unique solution*

$$u(t) = T(t)u_0, \ \ t \geq 0.$$

Proof. The existence is a consequence of Theorem 4.31 (b). To prove uniqueness, let $v(t)$ be any solution of the Cauchy problem and set $F(s) = T(t-s)v(s)$. Then

$$F(s+h) - F(s) = T(t-s-h)\left(v(s+h) - v(s)\right) - \left(T(t-s) - T(t-s-h)\right)v(s).$$

Since $v(s) \in \mathcal{D}(A)$, Theorem 4.31 implies that $F(s)$ is strongly differentiable in s and

$$\begin{aligned}(d/ds)F(s) &= -AT(t-s)v(s) + T(t-s)v'(s) \\ &= -AT(t-s)v(s) + T(t-s)Av(s) = 0, \ \ 0 \leq s \leq t.\end{aligned}$$

Hence by the mean value theorem, $F(s)$ =constant for $0 \leq s \leq t$. In particular, $F(t) = F(0)$ or $v(t) = T(t)u_0 = u(t)$. □

Remark 4.16. (i) If $T(t)u_0$ is differentiable for every $u_0 \in X$ and $t \geq 0$, then in particular, $(d/dt)(T(t)u_0)|_{t=0} = Au_0$. Hence, $\mathcal{D}(A) = X$ and A (being a closed operator) must be bounded. Thus if A is unbounded, then $T(t)u_0$ is not differentiable for all $u_0 \in X$. We can however consider $u(t) = T(t)u_0$ as a **generalized solution** of the Cauchy problem.

(ii) From the uniqueness in Theorem 4.32 we have that a linear operator $A : \mathcal{D}(A) \to X$ which is densely defined can be the infinitesimal generator of at most one C_0 semigroup $\{T(t)\}$. Moreover, if $\{T(t)\}$ is a C_0 semigroup whose infinitesimal generator A is bounded, then $T(t) = e^{At}$ since then A is the infinitesimal generator of $T(t)$ and e^{tA}.

4.4.5. Characterization of Generators. We know the generator A of a C_0 semigroup is a closed, densely defined linear operator. Now given a closed, densely defined linear operator A on a Banach space X, we would like to know whether A generates a C_0 semigroup on X; if it does, then the abstract Cauchy problem (4.93) has a unique solution.

Suppose A generates a C_0 contraction semigroup $\{T(t)\}$, i.e., $\|T(t)\| \leq 1$ for all $t \geq 0$. For $u \in X$, $\lambda > 0$, the integral

$$\int_s^t e^{-\lambda\tau} T(\tau)u d\tau$$

is well defined, and as $\|T(\tau)u\| \leq \|u\|$ for all τ, we deduce that this integral tends to zero as $t, s \to \infty$. Hence the integral

$$R(\lambda; A)u = \int_0^\infty e^{-\lambda\tau} T(\tau)u d\tau$$

exists as an improper Riemann integral. Since

$$\|R(\lambda;A)u\| \le \|u\| \int_0^\infty e^{-\lambda\tau} d\tau = (1/\lambda)\|u\|$$

it follows that $R(\lambda;A)$ is a bounded linear operator and $\|R(\lambda;A)\| \le 1/\lambda$ for all $\lambda > 0$.

Lemma 4.33. *If A is the infinitesimal generator of a contraction semigroup $\{T(t)\}$, then $(\lambda I - A)$ is invertible for every $\lambda > 0$ and*

$$(\lambda I - A)^{-1} = R(\lambda;A).$$

In particular, for every $\lambda > 0$, $\|(\lambda I - A)^{-1}\| \le \lambda^{-1}$. (Thus the resolvent operator $R_\lambda = (\lambda I - A)^{-1}$ is the Laplace transform of the semigroup.)

Proof. Let $h > 0$. Then for $u \in X$

$$\begin{aligned} \left(\frac{T(h)-I}{h}\right) R(\lambda;A)u &= (1/h)\int_0^\infty e^{-\lambda\tau}(T(\tau+h)u - T(\tau)u)d\tau \\ &= (1/h)\int_h^\infty e^{-\lambda(\tau-h)}T(\tau)ud\tau - (1/h)\int_0^\infty e^{-\lambda\tau}T(\tau)ud\tau \\ &= \left(\frac{e^{\lambda\tau}-1}{h}\right)\int_0^\infty e^{-\lambda h}T(\tau)ud\tau - \frac{e^{-\lambda h}}{h}\int_0^h e^{-\lambda\tau}T(\tau)ud\tau \\ &\to \lambda R(\lambda;A)u - u \end{aligned}$$

Thus $R(\lambda;A)u \in \mathcal{D}(A)$ and $AR(\lambda;A)u = \lambda R(\lambda;A)u - u$, i.e., $(\lambda I - A)R(\lambda;A)u = u$.

Now let $u \in \mathcal{D}(A)$. Then

$$\begin{aligned} R(\lambda;A)Au &= \int_0^\infty e^{-\lambda\tau}T(\tau)Aud\tau = \int_0^\infty e^{-\lambda\tau}\frac{d}{d\tau}(T(\tau)u)d\tau \\ &= \lambda\int_0^\infty e^{-\lambda\tau}T(\tau)ud\tau - u \\ &= \lambda R(\lambda;A) - u \end{aligned}$$

i.e., $R(\lambda;A)(\lambda I - A)u = u$. This proves $R(\lambda;A) = (\lambda I - A)^{-1} = R_\lambda$. □

Theorem 4.34. (Hille-Yosida Theorem) *A linear operator $A : \mathcal{D}(A) \to X$ is the generator of a C_0 contraction semigroup if and only if A is closed, densely defined and*

$$(0,\infty) \subset \rho(A), \quad \|(\lambda I - A)^{-1}\| \le \lambda^{-1} \ \ \forall\, \lambda > 0. \tag{4.94}$$

Proof. 1. If A is the generator, (4.94) follows from the previous lemma.

2. Now assume A is closed, densely defined, and satisfies (4.94). We must build a C_0 contraction semigroup whose infinitesimal generator is A. For this, fix $\lambda > 0$ and defined

$$A_\lambda := -\lambda I + \lambda^2 R_\lambda = \lambda A R_\lambda.$$

The operator A_λ is a kind of regularization of A.

3. We first claim

$$\lim_{\lambda\to\infty} A_\lambda u = Au \quad (u \in \mathcal{D}(A)). \tag{4.95}$$

Indeed, for each $u \in \mathcal{D}(A)$, since $\lambda R_\lambda u - u = AR_\lambda u = R_\lambda Au$, we have

$$\|\lambda R_\lambda u - u\| \le \|R_\lambda\|\|Au\| \le \frac{1}{\lambda}\|Au\| \to 0 \quad \text{as } \lambda \to \infty.$$

Hence $\lambda R_\lambda u \to u$ as $\lambda \to \infty$ if $u \in \mathcal{D}(A)$. But since $\|\lambda R_\lambda\| \leq 1$ and $\mathcal{D}(A)$ is dense, we have

$$\lim_{\lambda\to\infty} \lambda R_\lambda u \ = u \quad (u \in X).$$

So, if $u \in \mathcal{D}(A)$,

$$\lim_{\lambda\to\infty} A_\lambda u = \lim_{\lambda\to\infty} \lambda A R_\lambda u = \lim_{\lambda\to\infty} \lambda R_\lambda A u = Au.$$

4. Define

$$T_\lambda(t) := e^{tA_\lambda} = e^{-\lambda t} e^{\lambda^2 t R_\lambda} = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda^2 t)^k}{k!} R_\lambda^k.$$

Observe, since $\|R_\lambda\| \leq 1/\lambda$,

$$\|T_\lambda(t)\| \leq e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda^2 t)^k}{k!} = 1 \quad (t \geq 0).$$

Consequently, $\{T_\lambda(t)\}_{t\geq 0}$ is a contraction semigroup, with generator A_λ.

5. By the **resolvent identity** above, we see $A_\lambda A_\mu = A_\mu A_\lambda$ for all $\lambda, \mu > 0$. So

$$A_\mu T_\lambda(t) = T_\lambda(t) A_\mu \quad (t > 0).$$

We thus compute

$$T_\lambda(t)u - T_\mu(t)u = \int_0^t \frac{d}{d\tau}[T_\mu(t-\tau)T_\lambda(\tau)u]d\tau = \int_0^t T_\mu(t-\tau)T_\lambda(\tau)(A_\lambda u - A_\mu u)d\tau.$$

Consequently, if $u \in \mathcal{D}(A)$, then

$$\|T_\lambda(t)u - T_\mu(t)u\| \leq t\|A_\lambda u - A_\mu u\| \to 0 \quad \text{as } \lambda, \mu \to \infty.$$

This proves that $\{T_\lambda(t)u\}_{\lambda>0}$ is a Cauchy sequence in X. Hence define

$$T(t)u := \lim_{\lambda\to\infty} T_\lambda(t)u \quad (u \in \mathcal{D}(A),\ t \geq 0).$$

From $\|T_\lambda(t)\| \leq 1$, it is straightforward to show that $\{T(t)\}_{t\geq 0}$ is a $C-0$ contraction semigroup.

6. Finally we show the generator of $\{T(t)\}_{t\geq 0}$ is A. Write B to denote this generator. Now

$$T_\lambda(t)u - u = \int_0^t T_\lambda(\tau)A_\lambda u d\tau$$

and

$$T_\lambda(\tau)A_\lambda u - T(\tau)Au\| \leq \|T_\lambda(\tau)\|\|A_\lambda u - Au\| + \|T_\lambda(\tau)Au - T(\tau)Au\| \to 0$$

as $\lambda \to \infty$ for all $u \in \mathcal{D}(A)$. Thus we have

$$T(t)u - u = \int_0^t T(\tau)A_\lambda u d\tau \quad (u \in \mathcal{D}(A)).$$

This, recalling the definition of $\mathcal{D}(B)$, proves $\mathcal{D}(A) \subseteq \mathcal{D}(B)$ and, for all $u \in \mathcal{D}(A)$,

$$Bu = \lim_{h\to 0^+} \frac{T(h)u - u}{h} = Au.$$

To show $\mathcal{D}(B) \subseteq \mathcal{D}(A)$, note $(0, \infty) \subset \rho(A) \cap \rho(B)$ and also $(\lambda I - B)(\mathcal{D}(A)) = (\lambda I - A)(\mathcal{D}(A)) = X$. So, if $y \in \mathcal{D}(B)$, then there exists a $x \in \mathcal{D}(A)$, such that $(\lambda I - B)y = \lambda y - By = (\lambda I - A)x = \lambda x - Ax = \lambda x - Bx$ and hence $y = x \in \mathcal{D}(A)$. This proves $\mathcal{D}(B) \subseteq \mathcal{D}(A)$. So $A = B$. $\square$

Remark 4.17. A C_0 semigroup $\{T(t)\}$ is called a C_0 **ω-contraction semigroup**, provided for some $\omega \in \mathbb{R}$,

$$\|T(t)\| \leq e^{\omega t} \quad (t \geq 0). \tag{4.96}$$

In this case, $T_1(t) = e^{-\omega t}T(t)$ will be a C_0 contraction semigroup. If A is the generator of T, then $A - \omega I$ is the generator of T_1 and if A is the generator of T_1, then $A + \omega I$ is the generator of T. Thus we can deduce the following result:

Theorem 4.35. *A linear operator $A : \mathcal{D}(A) \to X$ is the generator of a C_0 ω-contraction semigroup $\{T(t)\}$ if and only if A is closed, densely defined and satisfies*

$$(\omega, \infty) \subset \rho(A), \quad \|(\lambda I - A)^{-1}\| \leq \frac{1}{\lambda - \omega} \quad \forall\, \lambda > \omega.$$

To characterize the infinitesimal generators of general C_0 semigroups, we usually renorm the Banach space so that $\{T(t)\}$ becomes a C_0 contraction semigroup in the new (equivalent) norm. We just state the following general result without proof.

Theorem 4.36. (Hille-Yosida-Phillips) *A linear operator $A : \mathcal{D}(A) \to X$ is the generator of a C_0 semigroup $\{T(t)\}$ if and only if A is closed, densely defined and there exist constants $M \geq 1$, $\omega \in \mathbb{R}$ such that $\lambda \in \rho(A)$ for each $\lambda > \omega$ and*

$$\|(\lambda I - A)^{-n}\| \leq \frac{M}{(\lambda - \omega)^n}, \quad \forall\, \lambda > \omega, \quad n = 1, 2, \ldots.$$

In this case $\|T(t)\| \leq Me^{\omega t}$.

4.4.6. Another Characterization of C_0 Contraction Semigroups. We now give a different characterization of generators of semigroups of contractions in a Hilbert space H.

Definition 4.18. (i) A linear operator $A : \mathcal{D}(A) \to H$ is said to be **dissipative** if

$$\mathrm{Re}(Au, u) \leq 0 \quad \text{for all} \quad u \in \mathcal{D}(A).$$

(Note that if H is a real Hilbert space, then A is dissipative iff $-A$ is monotone.)

(ii) We say that A is **maximal dissipative** if in addition $\mathcal{R}(I - A) = H$. It is maximal in the sense that there exists no linear operator B with the same properties and such that B is an extension of A. Indeed, suppose such an extension exists. Let $u \in \mathcal{D}(B)$ and $v \in \mathcal{D}(A)$ be such that $(I - A)v = (I - B)u$. Since $Av = Bv$, we have $(I - B)(v - u) = 0$. Multiplying this by $(v - u)$ we get

$$0 \leq \|v - u\|^2 = \mathrm{Re}(B(v - u), v - u) \leq 0$$

whence $u = v$ or $u \in \mathcal{D}(A)$. Thus $B = A$.

Lemma 4.37. *If A is dissipative, then*

$$\|(\lambda I - A)u\| \geq \lambda\|u\| \quad \textit{for all} \quad \lambda > 0, \quad u \in \mathcal{D}(A). \tag{4.97}$$

Proof.

$$\|(\lambda I - A)u\|\|u\| \geq \mathrm{Re}((\lambda I - A)u, u) = \lambda\|u\|^2 - \mathrm{Re}(Au, u) \geq \lambda\|u\|^2.$$

□

Theorem 4.38. (Lumer-Phillips) *Let $A : \mathcal{D}(A) \to H$ be a densely defined linear operator.*

(a) *If A is a generator of a C_0 contraction semigroup $\{T(t)\}$, then A is dissipative and $\mathcal{R}(\lambda I - A) = H$ for all $\lambda > 0$.*

(b) *If A is dissipative and there exists a $\lambda_0 > 0$ such that $\mathcal{R}(\lambda_0 I - A) = H$, then A is a generator of a C_0 contraction semigroup. In particular, A is maximal dissipative.*

Proof. (a) By the Hille-Yosida Theorem, $(0,\infty) \subset \rho(A)$ and therefore $\lambda I - A$ is onto H for all $\lambda > 0$. Furthermore,

$$|(T(t)u, u)| \leq \|T(t)\| \|u\|^2 \leq \|u\|^2.$$

Hence

$$\mathrm{Re}\left(\frac{T(t)u - u}{t}, u\right) = \frac{1}{t}(\mathrm{Re}(T(t)u, u) - \|u\|^2) \leq 0.$$

Let $u \in \mathcal{D}(A)$ and let $t \to 0^+$ obtaining $\mathrm{Re}(Au, u) \leq 0$.

(b) Since $\mathcal{R}(\lambda_0 I - A) = H$, it follows from (4.97) that $\lambda_0 \in \rho(A)$ and A is closed. If $\mathcal{R}(\lambda I - A) = H$ for all $\lambda > 0$, then $(0,\infty) \subset \rho(A)$ and $\|\mathcal{R}(\lambda; A)\| \leq \lambda^{-1}$ by (4.97). The desired result then follows from the Hille-Yosida Theorem.

In order to prove that $\mathcal{R}(\lambda I - A) = H$ for all $\lambda > 0$, consider the open set

$$\Gamma = \{\lambda : \lambda > 0 \text{ and } \mathcal{R}(\lambda I - A) = H\}.$$

Note that $\lambda \in \Gamma$ implies $\lambda \in \rho(A)$, and since $\rho(A)$ is open, there is a neighborhood of λ whose intersection with the real line is in Γ. By hypothesis, $\lambda_0 \in \Gamma$. Hence if Γ is closed in $(0,\infty)$, then $\Gamma = (0,\infty)$. Let $\lambda_n \to \lambda > 0$, $\lambda_n \in \Gamma$. For every $v \in H$, there exist u_n's such that

$$\lambda_n u_n - A u_n = v \quad \text{for all} \quad n. \tag{4.98}$$

From (4.97) it follows that $\|u_n\| \leq \lambda_n^{-1}\|v\| \leq c$. Now by (4.97) again,

$$\begin{aligned} \lambda_n \|u_n - u_m\| &\leq \|\lambda_n(u_n - u_m) - A(u_n - u_m)\| \\ &= \|v - \lambda_n u_m + A u_m\| = \|v - \lambda_n u_m + \lambda_m u_m - v\| \\ &= |\lambda_n - \lambda_m| \|u_m\| \leq c|\lambda_n - \lambda_m|. \end{aligned}$$

Therefore $\{u_n\}$ is a Cauchy sequence with limit, say, u. Thus by (4.98), $Au_n \to \lambda u - v$ and since A is closed, $u \in \mathcal{D}(A)$ and $\lambda u - Au = v$. Thus $\lambda \in \Gamma$. ☐

Corollary 4.39. *Let $A : \mathcal{D}(A) \to H$ be a densely defined closed linear operator. If both A and A^* are dissipative, then A is a generator of a C_0 contraction semigroup.*

Proof. In view of (b) of Theorem 4.38, it suffices to show that $\mathcal{R}(I - A) = H$. Since A is closed so is $I - A$. Moreover, (4.97) and Corollary 4.27 imply $\mathcal{R}(I - A)$ is a closed subspace of H and therefore $\mathcal{R}(I - A) = (\mathcal{N}(I - A^*))^\perp$. Consequently, it suffices to show that $I - A^*$ is one-to-one. But this follows from the fact that A^* is dissipative. ☐

While we could solve the initial value problem in $[0,\infty)$ when $u_0 \in \mathcal{D}(A)$ and not for general $u_0 \in H$, one can show that if A is self-adjoint then we can solve the problem for any initial data $u_0 \in H$. The price we pay for this is the lack of differentiability at $t = 0$.

Theorem 4.40. *Let A be a self-adjoint maximal dissipative operator and let $u_0 \in H$. Then there exists a unique u such that*

$$u \in C([0,\infty); H) \cap C^1((0,\infty); H) \cap C((0,\infty); \mathcal{D}(A))$$

and u satisfies the initial value problem (4.93). (Continuity in $\mathcal{D}(A)$ is with respect to the graph norm.)

Proof. We only prove uniqueness. Let u_1, u_2 be two solutions of (4.93). Set $v(t) = \|u_1(t) - u_2(t)\|^2$ which is continuous and such that $v(0) = 0$. Taking the inner product of $u' = Au$ with $u_1(t) - u_2(t)$ for $u = u_1$ and $u = u_2$ and using the dissipativity of A, we get

$$\frac{d}{dt}v(t) \leq 0$$

whence it follows that $v(t) \equiv 0$ and so $u_1(t) = u_2(t)$ for $t \geq 0$. □

4.4.7. Applications. We demonstrate in this section that certain second-order parabolic and hyperbolic PDE can be realized within the semigroup framework. The solvability of the Cauchy problem for $u' = Au$ is reduced to the study of the solvability and a-priori estimate of the solutions of the usually simpler problem $\lambda u - Au = v,\ u \in \mathcal{D}(A)$ where $v \in X$ and $\lambda > \omega$.

A. Second-order Parabolic Equations. We consider the following initial-boundary value problem (IBVP):

$$\begin{cases} u_t + Lu = 0 & \text{in } \Omega_T, \\ u = 0 & \text{on } \partial\Omega \times [0, T], \\ u = g & \text{on } \Omega \times \{t = 0\}, \end{cases} \tag{4.99}$$

which a special case studied above. We assume L has the divergence structure, satisfies the uniform ellipticity condition, and has coefficients that are smooth and *independent of time* t. We assume domain Ω is bounded and has smooth boundary $\partial\Omega$.

We propose problem (4.99) as the flow determined by a semigroup on $X = L^2(\Omega)$. For this we set

$$\mathcal{D}(A) = H_0^1(\Omega) \cap H^2(\Omega)$$

and define

$$Au := -Lu \quad (u \in \mathcal{D}(A)).$$

Recall the Gårding's inequality

$$\beta\|u\|^2_{H_0^1(\Omega)} \leq B[u, u] + \gamma\|u^2_{L^2(\Omega)}. \tag{4.100}$$

Theorem 4.41. (Second-order parabolic PDE as semigroup) *The operator A generates a C_0 γ-contraction semigroup $\{T(t)\}_{t\geq 0}$ on $L^2(\Omega)$. Therefore $u(t) = T(t)g$ defines the solution to* (4.99) *for any given $g \in H_0^1(\Omega) \cap H^2(\Omega)$.*

Proof. We must verify the hypotheses of the Hille-Yosida-Phillips Theorem. Therefore the evolution problem becomes a study of the spectral properties of the elliptic operator A defined above.

1. $\mathcal{D}(A)$ is clearly dense in $L^2(\Omega)$.
2. We show A is closed. Indeed, let $\{u_k\}$ be a sequence in $\mathcal{D}(A)$ with

$$u_k \to u, \ \ Au_k \to f \quad \text{in } L^2(\Omega).$$

According the global regularity,

$$\|u_k - u_l\|_{H^2(\Omega)} \leq C(\|Au_k - Au_l\|_{L^2(\Omega)} + \|u_k - u_l\|_{L^2(\Omega)})$$

for all k, l. This shows that $\{u_k\}$ is a Cauchy sequence in $H^2(\Omega)$ and so

$$u_k \to u \in H^2(\Omega) \quad \text{in } H^2(\Omega).$$

Moreover, $u \in H_0^1(\Omega)$. (Trace operator is continuous from $H^1(\Omega)$ to $L^2(\partial\Omega)$.) Therefore $u \in \mathcal{D}(A)$. Furthermore the strong convergence $u_k \to u$ in $H^2(\Omega)$ implies $Au_k \to Au$ in $L^2(\Omega)$, and thus $f = Au$. By definition, A is closed.

3. We next prove the condition $(\gamma, \infty) \subset \rho(A)$; that is, $\lambda I - A$ is one-to-one and onto for all $\lambda > \gamma$. By Lax-Milgram's Theorem, for all $\lambda \geq \gamma$, the BVP

$$\begin{cases} Lu + \lambda u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{4.101}$$

has a unique weak solution $u \in H_0^1(\Omega)$ for each $f \in L^2(\Omega)$. The global regularity theory shows that $u \in \mathcal{D}(A)$, and

$$\lambda u - Au = f.$$

Thus $(\lambda I - A)\colon \mathcal{D}(A) \to X$ is one-to-one and onto, for all $\lambda \geq \gamma$. This proves $[\gamma, \infty) \subset \rho(A)$.

4. Let $R_\lambda = R(\lambda, A) = (\lambda I - A)^{-1}$. We will show

$$\|R_\lambda\| \leq \frac{1}{\lambda - \gamma} \quad (\lambda > \gamma)$$

as required for generating a C_0 γ-contraction semigroup. To show this, consider the weak form of problem (4.101):

$$B[u, v] + \lambda(u, v) = (f, v) \quad \forall\, v \in H_0^1(\Omega),$$

where, as usual, $(,)$ stands for the L^2-inner product. Set $v = u$ and recall the Gårding's inequality to compute

$$(\lambda - \gamma)\|u\|^2_{L^2(\Omega)} \leq \|f\|_{L^2(\Omega)}\|u\|_{L^2(\Omega)}.$$

This implies, as $u = R_\lambda f$,

$$\|R_\lambda f\|_{L^2(\Omega)} = \|u\|_{L^2(\Omega)} \leq \frac{1}{\lambda - \gamma}\|f\|_{L^2(\Omega)}.$$

This bound is valid for all $f \in L^2(\Omega)$, which proves the desired claim. □

Example 4.42. We study the heat equation:

$$\begin{cases} u_t(x,t) - \Delta u(x,t) = 0 & \text{in } \Omega \times [0, \infty), \\ u(x,t) = 0 & \text{on } \partial\Omega \times [0, \infty), \\ u(x, 0) = u_0(x), & x \in \Omega. \end{cases} \tag{4.102}$$

Theorem 4.43. *Let $u_0 \in L^2(\Omega)$. Then there exists a unique solution u of (4.102) such that*

$$u \in C([0,\infty); L^2(\Omega)) \cap C^1((0,\infty); L^2(\Omega)) \cap C((0,\infty); H^2(\Omega) \cap H_0^1(\Omega)).$$

Proof. Let $H = L^2(\Omega)$ and define $A : \mathcal{D}(A) \subset H \to H$ by

$$\mathcal{D}(A) = H^2(\Omega) \cap H_0^1(\Omega); \quad Au = \Delta u \text{ for } u \in \mathcal{D}(A).$$

Using the concept of an abstract function, the IBVP can be posed as the abstract Cauchy problem

$$\frac{du}{dt} = Au \ (t \geq 0), \ \ u(0) = u_0 \in H.$$

Since for $u \in H^2(\Omega) \cap H_0^1(\Omega)$

$$(Au, u) = \int_\Omega u \Delta u \, dx = -\int_\Omega |\nabla u|^2 \, dx \leq 0,$$

we have that A is dissipative. It is maximal dissipative; that is, $\mathcal{R}(I - A) = H$, for, by the Lax-Milgram, there exists a unique $u \in H_0^1(\Omega)$ such that

$$(I - A)u = u - \Delta u = f \quad \text{for all} \quad f \in L^2(\Omega).$$

Also by elliptic regularity, $u \in H^2(\Omega)$ and so $u \in \mathcal{D}(A)$; hence $\mathcal{R}(I - A) = H$. Finally, we saw earlier that A is self-adjoint. Hence we can apply Theorem 4.40 to deduce the desired conclusions. □

Remark 4.19. Note that however badly behaved $u_0 \in L^2(\Omega)$ may be, $u(x,t)$ is very smooth for all $t > 0$. This is known as the **strong regularizing effect** of the heat operator. In particular, this shows that the heat equation is **irreversible** in time, i.e., we cannot always solve the problem

$$\begin{cases} u_t(x,t) - \Delta u(x,t) = 0 & \text{in } \Omega \times [0,T), \\ u(x,t) = 0 & \text{on } \partial\Omega \times (0,T), \\ u(x,T) = u_T(x), & x \in \Omega. \end{cases}$$

B. Second-order Hyperbolic Equations. We consider the following initial-boundary value problem (IBVP):

$$\begin{cases} u_{tt} + Lu = 0 & \text{in } \Omega_T, \\ u = 0 & \text{on } \partial\Omega \times [0,T], \\ u = g, \; u_t = h & \text{on } \Omega \times \{t = 0\}, \end{cases} \tag{4.103}$$

which a special case studied above. We assume domain Ω is bounded and has smooth boundary $\partial\Omega$ and L has the symmetric form:

$$Lu = -\sum_{i,j=1}^{n} D_j(a_{ij}(x)D_i u) + c(x)u,$$

where $a_{ij}(x) = a_{ji}(x)$ and $c(x) \geq 0$ are all smooth functions and $(a_{ij}(x)) \geq \theta I$. Hence

$$B[u,u] \geq \theta \|u\|^2_{H_0^1(\Omega)} \quad (u \in H_0^1(\Omega)). \tag{4.104}$$

We recast problem (4.103) as a first-order system by setting $v := u_t$. Then (4.103) reads

$$\begin{cases} u_t = v, \;\; v_t + Lu = 0 & \text{in } \Omega_T, \\ u = 0 & \text{on } \partial\Omega \times [0,T], \\ u = g, \; v = h & \text{on } \Omega \times \{t = 0\}, \end{cases} \tag{4.105}$$

Now take Hilbert space $X = H_0^1(\Omega) \times L^2(\Omega)$, with the inner product

$$\langle\langle (u,v), (f,g) \rangle\rangle = B[u,f] + (v,g)_{L^2(\Omega)}$$

and the norm

$$\|(u,v)\| := (B[u,u] + \|v\|^2_{L^2(\Omega)})^{1/2}.$$

Define

$$\mathcal{D}(A) := [H^2(\Omega) \cap H_0^1(\Omega)] \times H_0^1(\Omega)$$

and define

$$A(u,v) := (v, -Lu) \quad \forall\, (u,v) \in \mathcal{D}(A).$$

Theorem 4.44. (Second-order hyperbolic PDE as semigroup) *The operator A generates a C_0 contraction semigroup $\{T(t)\}_{t\geq 0}$ on $H_0^1(\Omega) \times L^2(\Omega)$. Therefore $(u(t), v(t)) = T(t)(g,h)$ defines the solution to* (4.105) *for any given $(g,h) \in [H_0^1(\Omega) \cap H^2(\Omega)] \times H_0^1(\Omega)$.*

Proof. We must verify the hypotheses of the Hille-Yosida-Phillips Theorem. Therefore the evolution problem again becomes a study of the spectral properties of the linear operator A defined above.

1. $\mathcal{D}(A)$ is clearly dense in $L^2(\Omega)$.

2. We show A is closed. Indeed, let $\{(u_k, v_k)\}$ be a sequence in $\mathcal{D}(A)$ with

$$(u_k, v_k) \to (u, v), \quad A(u_k, v_k) \to (f, g) \quad \text{in } X = H_0^1(\Omega) \times L^2(\Omega).$$

Since $A(u_k, v_k) = (v_k, -Lu_k)$, we conclude $f = v$ and $Lu_k \to -g$ in $L^2(\Omega)$. As before, by elliptic estimates and regularity, $u_k \to u$ in $H^2(\Omega)$ and $g = -Lu$. Thus $(u, v) \in \mathcal{D}(A)$, $A(u, v) = (v, -Lu) = (f, g)$. By definition, A is closed.

3. We next prove $(0, \infty) \subset \rho(A)$; that is, $\lambda I - A$ is one-to-one and onto for all $\lambda > 0$. Now given $\lambda > 0$ and $(f, g) \in X = H_0^1(\Omega) \times L^2(\Omega)$, consider the operator equation

$$(\lambda I - A)(u, v) = \lambda(u, v) - A(u, v) = (f, g).$$

This is equivalent to the two scalar equations:

$$\begin{cases} \lambda u - v = f & (u \in H^2(\Omega) \cap H_0^1(\Omega)), \\ \lambda v + Lu = g & (v \in H_0^1(\Omega)). \end{cases} \tag{4.106}$$

But this implies

$$\lambda^2 u + Lu = \lambda f + g \quad (u \in H^2(\Omega) \cap H_0^1(\Omega)).$$

Since $\lambda^2 > 0$, by existence and regularity of elliptic PDE, this problem has a unique solution $u \in H^2(\Omega) \cap H_0^1(\Omega)$. Once u is found, define $v = \lambda u - f \in H_0^1(\Omega)$. We have shown that (4.106) has a unique solution (u, v). This proves that $(\lambda I - A)\colon \mathcal{D}(A) \to X$ is one-to-one and onto, for each $\lambda > 0$; hence $(0, \infty) \subset \rho(A)$.

4. Let $R_\lambda = R(\lambda, A) = (\lambda I - A)^{-1}$. We will show

$$\|R_\lambda\| \le \frac{1}{\lambda} \quad (\lambda > 0),$$

as required for generating a C_0 contraction semigroup. Note that $R_\lambda(f, g) = (u, v)$ if and only if (4.106) holds. From the second equation in (4.106), we deduce

$$\lambda \|v\|_{L^2(\Omega)}^2 + B[u, v] = (g, v)_{L^2(\Omega)}.$$

Putting $v = \lambda u - f$, we obtain

$$\lambda(\|v\|_{L^2(\Omega)}^2 + B[u, u]) = (v, g)_{L^2(\Omega)} + B[u, f] = \langle\langle (u, v), (f, g) \rangle\rangle \le \|(u, v)\| \|(f, g)\|.$$

This implies

$$\|(u, v)\| \le \frac{1}{\lambda} \|(f, g)\|;$$

so

$$\|R_\lambda(f, g)\| = \|(u, v)\| \le \frac{1}{\lambda} \|(f, g)\|.$$

This bound is valid for all $(f, g) \in X$, which proves the desired claim $\|R_\lambda\| \le \frac{1}{\lambda}$ for all $\lambda > 0$. $\square$

4.4.8. Nonhomogeneous Problems. We consider the nonhomogeneous Cauchy problem

$$\frac{du}{dt} - Au = f(t) \ (t > 0), \quad u(0) = u_0, \ \ u_0 \in \mathcal{D}(A). \tag{4.107}$$

A function $u(t) \in \mathcal{D}(A)$ is called a **classical solution** of (4.107) if it continuous for $t \geq 0$, continuously differentiable for $t > 0$ and satisfies (4.107).

Theorem 4.45. *Let $A : \mathcal{D}(A) \to X$ be the infinitesimal generator of a C_0 semigroup $\{T(t)\}$. Let $f : [0, \infty) \to X$ be continuously differentiable. Then the Cauchy problem (4.107) has a unique solution*

$$u(t) = T(t)u_0 + \int_0^t T(t-s)f(s)\, ds, \ \ t \geq 0. \tag{4.108}$$

Proof. Obviously $u(0) = u_0$. Define the function

$$v(t) = \int_0^t T(t-s)f(s)\, ds = \int_0^t T(s)f(t-s)\, ds.$$

Since $T(t)$ is bounded for each $t \geq 0$ and $f(s)$ is continuous, the above integrals exist. Now

$$\begin{aligned} [v(t+h) - v(t)]/h &= h^{-1} \int_0^{t+h} T(s)f(t+h-s)ds - h^{-1} \int_0^t T(s)f(t-s)ds \\ &= \int_0^t T(s)[f(t+h-s) - f(t-s)]/h\, ds \\ &+ h^{-1} \int_t^{t+h} T(s)f(t+h-s)\, ds. \end{aligned}$$

Hence $v'(t)$ exists and

$$v'(t) = \int_0^t T(s)f'(t-s)\, ds + T(t)f(0).$$

On the other hand, for $h > 0$ we have

$$\begin{aligned} [v(t+h) - v(t)]/h &= h^{-1} \int_0^{t+h} T(t+h-s)f(s)ds - h^{-1} \int_0^t T(t-s)f(s)ds \\ &= [T(h) - I]/h \int_0^t T(t-s)f(s)\, ds \\ &+ h^{-1} \int_t^{t+h} T(t+h-s)f(s)\, ds. \end{aligned}$$

Since the limit on the left exists and also

$$\lim_{h \to 0} h^{-1} \int_t^{t+h} T(t+h-s)f(s)ds = f(t)$$

it follows that

$$\lim_{h \to 0} A_h \int_0^t T(t-s)f(s)\, ds = A \int_0^t T(t-s)f(s)\, ds$$

which yields

$$v'(t) = A \int_0^t T(t-s)f(s)\, ds + f(t).$$

As a result we get

$$\begin{aligned}\frac{du}{dt} &= AT(t)u_0 + A\int_0^t T(t-s)f(s)\,ds + f(t)\\ &= Au(t) + f(t)\end{aligned}$$

and the proof is complete. □

Remark 4.20. The expression (4.108) is called the **variation of constants** or **Duhamel formula**. If the function f is integrable, then (4.108) still makes sense. Then u defined by that formula is called a **generalized solution** or **mild solution** of (4.107). It can be shown that a generalized solution always exists, but need not be a classical solution.

Index