1

Introduction and Main Results

This book provides a self-contained account of the functional analytic approach to the problem of construction of Markov processes with Ventcel' (Wentzell) boundary conditions in probability. More precisely, we prove existence theorems for Feller semigroups with Dirichlet boundary condition, oblique derivative boundary condition and first-order Ventcel' boundary condition for second-order, uniformly elliptic differential operators with *discontinuous* coefficients (Theorems 1.1, 1.2 and 1.3). Our approach here is distinguished by the extensive use of the ideas and techniques characteristic of the recent developments in the Calderón–Zygmund theory of singular integral operators with non-smooth kernels. It should be emphasized that singular integral operators with non-smooth kernels provide a powerful tool to deal with smoothness of solutions of elliptic boundary value problems, with minimal assumptions of regularity on the coefficients.

1.1 Formulation of the Problem

Now, let Ω be a bounded domain in Euclidean space $\mathbf{R}^n$, $n \geq 3$, with boundary $\partial\Omega$ of class $C^{1,1}$. We consider a second-order, elliptic differential operator A with real *discontinuous* coefficients of the form

$$Au := \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b^i(x) \frac{\partial u}{\partial x_i} + c(x)u. \tag{1.1}$$

In the case of continuous coefficients $a^{ij}(x)$, an L^p Schauder theory has been elaborated for second-order, uniformly elliptic differential operators (see [33]). However, the situation becomes rather difficult if we try to allow discontinuity on the $a^{ij}(x)$. In fact, it is known (see [49],

[88]) that arbitrary discontinuity of the $a^{ij}(x)$ breaks down as the L^p Schauder theory, except for the two-dimensional case ($n = 2$). In order to handle with the multidimensional case ($n \geq 3$), additional conditions on the $a^{ij}(x)$ should be required. Here we shall see that the relevant condition is that the coefficients $a^{ij}(x)$ belong to the Sarason class VMO of functions with vanishing mean oscillation. We remark that VMO consists of the John–Nirenberg class BMO of functions with bounded mean oscillation whose integral oscillation over balls shrinking to a point converge uniformly to zero (see Chapter 4 for the precise definitions and references).

Throughout this book, we assume that the coefficients $a^{ij}(x)$, $b^i(x)$ and $c(x)$ of the differential operator A satisfy the following three conditions (1), (2) and (3):

(1) $a^{ij}(x) \in \mathrm{VMO} \cap L^\infty(\Omega)$, $a^{ij}(x) = a^{ji}(x)$ for almost all $x \in \Omega$ and there exist a constant $\lambda > 0$ such that

$$\frac{1}{\lambda}|\xi|^2 \leq \sum_{i,j=1}^{n} a^{ij}(x)\xi_i\xi_j \leq \lambda\,|\xi|^2 \tag{1.2}$$

for almost all $x \in \Omega$ and all $\xi \in \mathbf{R}^n$.

(2) $b^i(x) \in L^\infty(\Omega)$ for $1 \leq i \leq n$.

(3) $c(x) \in L^\infty(\Omega)$ and $c(x) \leq 0$ for almost all $x \in \Omega$.

The differential operator A is called a *diffusion operator* which describes analytically a strong Markov process with continuous paths in the interior Ω such as Brownian motion (see Figure 1.1).

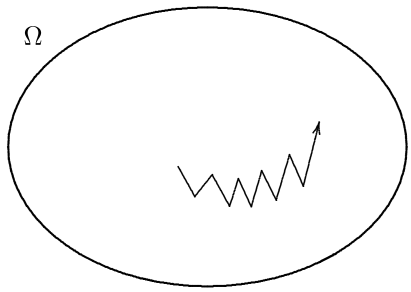

Fig. 1.1. A Markovian particle moves continuously

Moreover, we consider a first-order, boundary operator of the form

$$Lu := \mu(x')\frac{\partial u}{\partial \mathbf{n}} + \beta(x') \cdot u + \gamma(x')u - \delta(x')(Au|_{\partial\Omega}) \quad \text{on } \partial\Omega. \tag{1.3}$$

Throughout this book, we assume that the coefficients $\mu(x')$, $\beta(x')$, $\gamma(x')$ and $\delta(x')$ of the boundary operator L satisfy the following four conditions (4), (5), (6) and (7):

(4) $\mu(x')$ is a Lipschitz continuous function on $\partial\Omega$ and $\mu(x') \geq 0$ on $\partial\Omega$.
(5) $\beta(x')$ is a Lipschitz continuous vector field on $\partial\Omega$ (see Figure 1.2).
(6) $\gamma(x')$ is a Lipschitz continuous function on $\partial\Omega$ and $\gamma(x') \leq 0$ on $\partial\Omega$.
(7) $\delta(x')$ is a Lipschitz continuous function on $\partial\Omega$ and $\delta(x') \geq 0$ on $\partial\Omega$.
(8) $\mathbf{n} = (n_1, n_2, \ldots, n_n)$ is the unit inward normal to the boundary $\partial\Omega$ (see Figure 1.2).

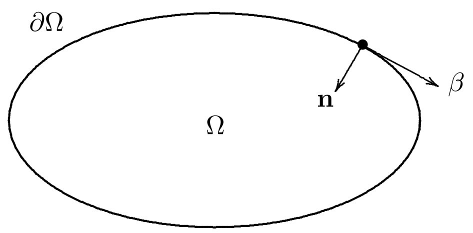

Fig. 1.2. The unit inward normal $\mathbf{n}$ to $\partial\Omega$ and the vector field β on $\partial\Omega$

The boundary condition L is called a first-order *Ventcel' boundary condition* (cf. [97]). The four terms of L

$$\gamma(x')u, \quad \mu(x')\frac{\partial u}{\partial \mathbf{n}}, \quad \beta(x') \cdot u, \quad \delta(x')(Au|_{\partial\Omega})$$

are supposed to correspond to the absorption phenomenon, the reflection phenomenon, the drift phenomenon along the boundary and the sticking (or viscosity) phenomenon, respectively (see Figures 1.3 and 1.4).

Let $C(\overline{\Omega})$ be the Banach space of real-valued, continuous functions on the closure $\overline{\Omega} = \Omega \cup \partial\Omega$, equipped with the maximum norm

$$\|f\|_{C(\overline{\Omega})} = \max_{x \in \overline{\Omega}} |f(x)|, \quad f \in C(\overline{\Omega}).$$

A strongly continuous semigroup $\{T_t\}_{t \geq 0}$ on the space $C(\overline{\Omega})$ is called a *Feller semigroup* if it is non-negative and contractive on $C(\overline{\Omega})$, that is,

$$f \in C(\overline{\Omega}), \ 0 \leq f(x) \leq 1 \quad \text{on } \overline{\Omega} \implies 0 \leq T_t f(x) \leq 1 \quad \text{on } \overline{\Omega}.$$

It is known (see [25], [79, Chapter 3]) that if T_t is a Feller semigroup on

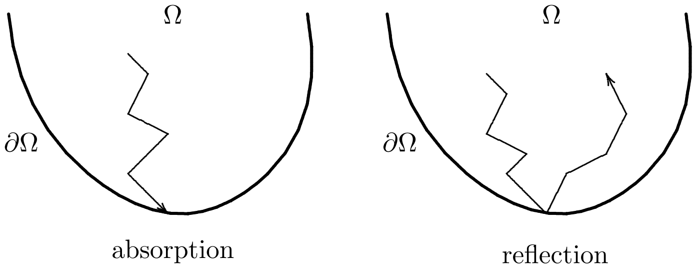

Fig. 1.3. The absorption phenomenon and the reflection phenomenon

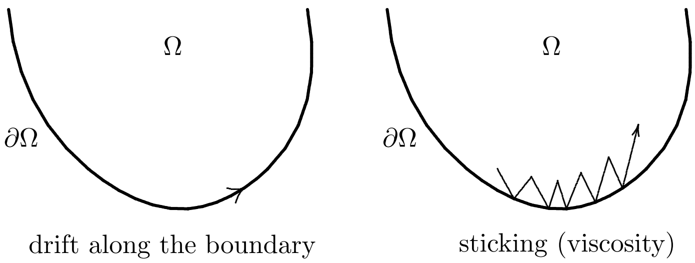

Fig. 1.4. The drift phenomenon along $\partial\Omega$ and the sticking (or viscosity) phenomenon

$C(\overline{\Omega})$, then there exists a unique Markov transition function $p_t(x,\cdot)$ on $\overline{\Omega}$ such that

$$T_t f(x) = \int_{\overline{\Omega}} p_t(x,dy) f(y) \quad \text{for all } f \in C(\overline{\Omega}).$$

Furthermore, it can be shown (see [24, Section 6.3], [79, Chapter 9]) that the function $p_t(x,\cdot)$ is the transition function of some *strong Markov process* whose paths are right-continuous and have no discontinuities other than jumps; hence the value $p_t(x,E)$ expresses the transition probability that a Markovian particle starting at position x will be found in the set E at time t.

The third purpose of this book is devoted to the functional analytic approach to the problem of existence of strong Markov processes in probability. More precisely, we consider the following problem:

Problem. Conversely, given analytic data (A, L), can we construct a Feller semigroup $\{T_t\}_{t\geq 0}$ whose infinitesimal generator $\mathfrak{A}$ is characterized by (A, L)?

1.2 Statement of Main Results

The next generation theorem for Feller semigroups ([83, Theorem 1.1]) asserts that there exists a Feller semigroup corresponding to such a diffusion phenomenon that a Markovian particle moves continuously in the state space, with absorption, reflection, drift and sticking phenomena at the boundary (see Figure 1.5):

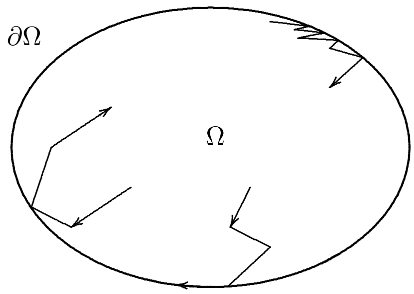

Fig. 1.5. The intuitive meaning of Theorem 1.1

Theorem 1.1. *If $n < p < \infty$, we define a linear operator $\mathfrak{A}$ from $C(\overline{\Omega})$ into itself as follows:*

(a) The domain $D(\mathfrak{A})$ is the set

$$D(\mathfrak{A}) = \left\{u \in W^{2,p}(\Omega) : Au \in C(\overline{\Omega}),\ Lu = 0 \text{ on } \partial\Omega\right\}. \quad (1.4)$$

(b) $\mathfrak{A}u = Au$ for every $u \in D(\mathfrak{A})$.

Here Au and Lu are taken in the sense of distributions.

Assume that the functions $\mu(x')$ and $\gamma(x')$ satisfy the conditions

$$\mu(x') > 0 \quad \text{on } \partial\Omega, \quad (\text{H.1})$$

and

$$\gamma(x') < 0 \quad \text{on } \partial\Omega. \quad (\text{H.2})$$

Then the operator $\mathfrak{A}$ is the infinitesimal generator of a Feller semigroup on $C(\overline{\Omega})$.

Remark 1.1. The domain $D(\mathfrak{A})$ does not depend on p, for $n < p < \infty$.

The crucial point in the proof of Theorem 1.1 is that we consider the term $\delta(x')(Au|_{\partial\Omega})$ of sticking phenomenon in the boundary condition

$$Lu = \mu(x')\frac{\partial u}{\partial \mathbf{n}} + \beta(x') \cdot u + \gamma(x')u - \delta(x')(Au|_{\partial\Omega}) \quad \text{on } \partial\Omega$$

as a term of "perturbation" of the oblique derivative boundary condition ($\delta(x') \equiv 0$)

$$L_\nu u := \mu(x')\frac{\partial u}{\partial \mathbf{n}} + \beta(x') \cdot u + \gamma(x')u \quad \text{on } \partial\Omega.$$

To do this, we prove the following generation theorem for Feller semigroups with oblique derivative boundary condition ([83, Theorem 1.2], [86, Theorem 1.1.]):

Theorem 1.2. *If $n < p < \infty$, we define a linear operator $\mathfrak{A}_\nu$ from $C(\overline{\Omega})$ into itself as follows:*

(a) The domain $D(\mathfrak{A}_\nu)$ is the set

$$D(\mathfrak{A}_\nu) = \left\{u \in W^{2,p}(\Omega) : Au \in C(\overline{\Omega}),\ L_\nu u = 0 \text{ on } \partial\Omega\right\}, \tag{1.5}$$

where

$$L_\nu u := \mu(x')\frac{\partial u}{\partial \mathbf{n}} + \beta(x') \cdot u + \gamma(x')u \quad \text{on } \partial\Omega.$$

(b) $\mathfrak{A}_\nu u = Au$ for every $u \in D(\mathfrak{A}_\nu)$.

Here Au and $L_\nu u$ are taken in the sense of distributions.

Assume that the functions $\mu(x')$ and $\gamma(x')$ satisfy the conditions

$$\mu(x') > 0 \quad \text{on } \partial\Omega, \tag{H.1}$$

and

$$\gamma(x') < 0 \quad \text{on } \partial\Omega. \tag{H.2}$$

Then the operator $\mathfrak{A}_\nu$ is the infinitesimal generator of a Feller semigroup on $C(\overline{\Omega})$.

Remark 1.2. The domain $D(\mathfrak{A}_\nu)$ does not depend on p, for $n < p < \infty$.

Rephrased, Theorem 1.2 asserts that there exists a Feller semigroup corresponding to such a diffusion phenomenon that a Markovian particle moves continuously in the state space, with absorption, reflection and drift phenomena at the boundary (see Figure 1.6).

Moreover, we construct a Feller semigroup associated with absorption phenomenon at the boundary, which we shall formulate precisely.

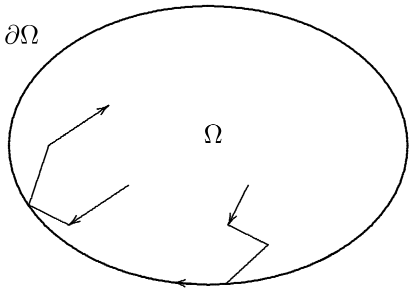

Fig. 1.6. The intuitive meaning of Theorem 1.2

We introduce a subspace of $C(\overline{\Omega})$, which is associated with Dirichlet boundary condition ($\mu(x') \equiv 0$, $\beta(x') \equiv 0$, $\gamma(x') \equiv 1$, $\delta(x') \equiv 0$), by the formula

$$C_0(\overline{\Omega}) = \left\{ u \in C(\overline{\Omega}) : u = 0 \text{ on } \partial\Omega \right\}.$$

A strongly continuous semigroup T_t on the space $C_0(\overline{\Omega})$ is called a *Feller semigroup* if it is non-negative and contractive on $C_0(\overline{\Omega})$, that is,

$$f \in C_0(\overline{\Omega}),\ 0 \leq f(x) \leq 1 \quad \text{on } \overline{\Omega} \implies 0 \leq T_t f(x) \leq 1 \quad \text{on } \overline{\Omega}.$$

It is known that if T_t is a Feller semigroup on $C_0(\overline{\Omega})$, then there exists a unique Markov transition function $p_t(x, \cdot)$ on Ω such that

$$T_t f(x) = \int_\Omega p_t(x, dy) f(y) \quad \text{for all } f \in C_0(\overline{\Omega}).$$

Furthermore, it can be shown (see [24, Section 6.3]) that the function $p_t(x, \cdot)$ is the transition function of some *strong Markov process* whose paths are right-continuous and have no discontinuities other than jumps; hence the value $p_t(x, E)$ expresses the transition probability that a Markovian particle starting at position x will be found in the set E at time t.

The next generation theorem for Feller semigroups ([77, Theorem 1.2]) asserts that there exists a Feller semigroup associated with absorption phenomenon at the boundary (see Figure 1.7):

Theorem 1.3. *If $n < p < \infty$, we define a linear operator $\mathfrak{A}_D$ from $C_0(\overline{\Omega})$ into itself as follows:*

(a) The domain $D(\mathfrak{A}_D)$ is the set

$$D(\mathfrak{A}_D) = \left\{ u \in W^{2,p}(\Omega) \cap C_0(\overline{\Omega}) : Au \in C_0(\overline{\Omega}) \right\}. \tag{1.6}$$

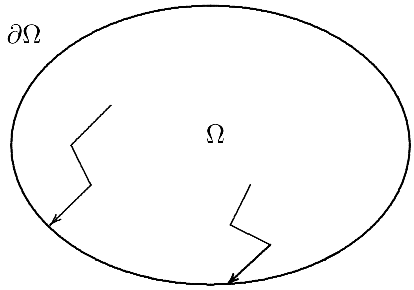

Fig. 1.7. The intuitive meanig of Theorem 1.3

(b) $\mathfrak{A}_D u = Au$ *for every* $u \in D(\mathfrak{A}_D)$.

Here Au *is taken in the sense of* distributions.

Assume that

$$c(x) \leq 0 \quad \textit{for almost all } x \in \Omega.$$

Then the operator $\mathfrak{A}_D$ *is the infinitesimal generator of a Feller semigroup on* $C_0(\overline{\Omega})$.

Remark 1.3. The domain $D(\mathfrak{A}_D)$ does not depend on p, for $n < p < \infty$.

The semigroup approach to Markov processes can be traced back to the pioneering work of Feller [26], [27] in early 1950s. Our presentation here follows the book of Dynkin [25] and also part of Lamperti's [41].

For more leisurely treatments of Markov processes and Feller semigroups, the reader is referred to Blumenthal–Getoor [8], Dynkin [24], [25], Itô–McKean [36], Lamperti [41], Revuz–Yor [59] and also Taira [79].

1.3 Summary of the Contents

This introductory chapter 1 is intended as a brief introduction to our problem and results in such a fashion that a broad spectrum of readers could understand. The contents of the book are divided into five principal parts.

The first part (Chapters 2–4) provides the elements of measure theory, functional analysis and real analysis. The material in these preparatory chapters is given for completeness, to minimize the necessity of consulting too many outside references. This makes the book fairly self-contained.

Chapter 2 is devoted to a review of standard topics from functional analysis. In Section 2.5 we formulate three pillars of functional analysis – Banach's open mapping theorem (Theorem 2.36), Banach's closed graph theorem (Theorem 2.37) and Banach's closed range theorem for closed operators (Theorem 2.40). Section 2.8 is devoted to the Riesz–Schauder theory for compact operators (Theorem 2.47). In Section 2.9 we state important properties of Fredholm operators (Theorems 2.48 through 2.53). In Section 2.10 we formulate the Riesz representation theorem for bounded linear functionals on a Hilbert space (Theorem 2.58). In the last Section 2.11 we formulate two fundamental theorems concerning spaces of continuous functions defined on a metric space – the Ascoli–Arzelà theorem (Theorem 2.67) and the Stone–Weierstrass theorem (Theorem 2.70). These topics form a necessary background for what follows.

In Chapter 3 we set forth the basic concepts of measure theory and develop the theory of integration on abstract measure spaces, paying particular attention to the Lebesgue integral on the Euclidean space $\mathbf{R}^n$. In particular, we give a complete proof of Minkowski's inequality for integrals (Theorem3.16) and Hardy's inequality (Theorem 3.18) in L^p spaces. In Section 3.9, we prove the Marcinkiewicz interpolation theorem (Theorem 3.30) that plays an important role in the proof of Theorem 9.5 in Section 9.3. In Section 3.10, as an application of Marcinkiewicz's interpolation theorem we study Riesz potentials in the classical potential theory (Theorem 3.31).

Chapter 4 is devoted to the precise definitions and statements, with some detailed proofs, of real analytic tools such as BMO and VMO functions, the Calderón–Zygmund decomposition (Theorem 4.7), the Hardy–Littlewood maximal function (Theorem 4.4), the John–Nirenberg inequality (Theorem 4.10), sharp functions (Theorem 4.14) and spherical harmonics (Theorem 4.31).

In the second part (Chapters 5–8) we study Sobolev spaces, Besov spaces and maximum principles in the framework of Sobolev spaces of L^p type that are used throughout the book.

The purpose of Chapter 5 is to study harmonic functions in the half-space in terms of Poisson integrals of functions in L^p spaces ((Theorems 5.8, 5.9 and 5.10)). In particular, we establish fundamental relationships between means of derivatives of Poisson integrals $u(x, y)$ taken with respect to the normal variable y and those taken with respect to the tangential variables x_i (Theorems 5.14 and 5.19).

In Chapter 6 we develop the theory of Besov spaces $B^{\alpha}_{p,q}(\mathbf{R}^n)$ on the

Euclidean space $\mathbf{R}^n$, paying particular attention to Poisson integrals. Besov spaces are functions spaces defined in terms of the L^p modulus of continuity. We prove a variety of equivalent norms for the Besov spaces on $\mathbf{R}^n$ via Poisson integrals (Theorems 6.3, 6.5 and 6.6).

Chapter 7 is devoted to the precise definitions and statements of function spaces with some detailed proofs. The function spaces of L^p type we treat are the generalized Sobolev spaces $W^{s,p}(\Omega)$ and $H^{s,p}(\Omega)$ and Besov spaces $B^{s,p}(\partial\Omega)$ on the boundary of a Lipschitz domain. It should be emphasized that Besov spaces enter naturally in connection with boundary value problems in the framework of Sobolev spaces of L^p type. Indeed, we need to make sense of the restriction $u|_{\partial\Omega}$ to the boundary $\partial\Omega$ as an element of a Besov space on $\partial\Omega$ when u belongs to a Sobolev space on the domain Ω. In particular, we formulate an important trace theorem (Theorem 7.5) that will be used in the study of boundary value problems in Part IV. In the last Section 7.4 we prove that

$$W^{\theta,n/\theta}(\mathbf{R}^n) \subset \text{VMO} \quad \text{for } 0 < \theta \leq 1$$

(Proposition 7.7).

In Chapter 8 we prove various maximum principles for second-order, elliptic differential operators with *discontinuous* coefficients such as the weak and strong maximum principles (Theorems 8.5 and 8.9) and the Hopf boundary point lemma (Lemma 8.8) in the framework of Sobolev spaces of L^p type that plays an essential role in the proof of uniqueness theorems for the Dirichlet problem in Part IV.

The third part (Chapters 9–11) is the heart of the subject. The Calderón–Zygmund theory of singular integral operators is a very refined mathematical tool whose full power is yet to be exploited.

Chapter 9 is devoted to a concise and accessible exposition of the most elementary part of the Calderón–Zygmund theory of singular integral operators. We present a straightforward treatment of the Calderón–Zygmund theory necessary for the study of elliptic boundary value problems, assuming only basic knowledge of real analysis and functional analysis. In particular, we present the basic theory of the Hilbert transform H that is a special case of the singular integral of a single independent variable (Theorems 9.6 and 9.14):

$$\begin{aligned} Hf(x) &:= \frac{1}{\pi}\left(\text{v.p.}\,\frac{1}{x}\right) * f \\ &= \frac{1}{\pi}\lim_{\varepsilon\downarrow 0}\int_{|x-t|>\varepsilon} \frac{f(t)}{x-t}\,dt \end{aligned}$$

$$= \frac{1}{\pi} \lim_{\varepsilon \downarrow 0} \int_{|s|>\varepsilon} f(x-s)\, \frac{ds}{s}.$$

The proof of Theorems 9.6 and 9.14 are flowcharted.

Moreover, we study bounded kernels (Theorem 9.2), continuous kernels (Theorem 9.5), odd kernels (Theorem 9.15) and Riesz kernels

$$R_j(x) = -\frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}} \frac{x_j}{|x|^{n+1}} \quad \text{for } 1 \leq j \leq n$$

(Theorem 9.16) in great detail. The proof of the continuous kernel case is based on a version of the Calderón–Zygmund decomposition (Lemma 9.3). The study of *odd* kernels $K(x)$ and Riesz kernels $R_j(x)$ is reduced to that of the Hilbert transform H. More precisely, we have the following formula for the odd kernel $K(x)$ (see formula (9.83)):

$$\begin{aligned} K * f(x) &:= \lim_{\varepsilon \downarrow 0} \int_{|x-y|>\varepsilon} K(x-y)\, f(y)\, dy \\ &= \frac{1}{2} \int_{\Sigma_{n-1}} K(\sigma) \left(\lim_{\varepsilon \downarrow 0} \int_{|t|>\varepsilon} f(x-t\sigma) \frac{dt}{t} \right) d\sigma. \end{aligned}$$

On the other hand, singular integral operators with *even* kernel $K(x)$ can be expressed as a finite sum of products of singular integral operators with odd kernel (Theorem 9.24). we have the following decomposition formula:

$$K * f = -\sum_{j=1}^{n} R_j * (R_j * K) * f, \tag{1.7}$$

where the operators R_j and $R_j * K$ have respectively the odd kernels. This decomposition formula (1.7) may be rephrased as follows:

$$\{\text{Even kernels}\} = \{\text{Riesz kernels}\} * \{\text{Odd kernels}\}\,.$$

In this way, we can prove the existence of the singular integral in the general case (Theorem 9.25). The proof of Theorem 9.25 is flowcharted.

The results discussed in Chapter 9 are adapted from the original paper of Calderón–Zygmund [15] and also Tanabe [89] and [90].

The first main result in Chapter 10 (Theorem 10.1) asserts the existence of singular integral operators and the second main result (Theorem 10.2) concerns commutators of BMO functions and singular integral operators. It should be emphasized again that singular integral operators with non-smooth kernels provide a powerful tool to deal with smoothness of solutions of partial differential equations, with minimal assumptions of regularity on the coefficients.

In Chapter 11 we consider singular integrals with kernels depending on a parameter, and prove theorems about singular integrals and commutators of L^∞ functions and singular integral operators (Theorems 11.1 and 11.2), generalizing Theorems 10.1 and 10.2. The main idea of proof is to reduce the variable kernel case to the constant kernel case. This is done by expanding the kernel into a series of spherical harmonics, each term defining a constant kernel operator treated in Chapter 10. Theorems about singular integrals and commutators are usually formulated in the whole space $\mathbf{R}^n$. However, our application to the theory of elliptic equations with discontinuous coefficients will require a local version of Theorems 11.1 and 11.2 (Theorems 11.3 and 11.4 and Corollary 11.5).

Our subject proper starts with the fourth part (Chapters 12–15).

Chapter 12 is devoted to the study of the *homogeneous* Dirichlet problem

$$\begin{cases} \mathcal{L}u = \sum_{i,j=1}^n a^{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j} = f & \text{in } \Omega, \\ \gamma_0 u = 0 & \text{on } \partial\Omega \end{cases}$$

in the framework of Sobolev spaces of L^p type. We state interior and global *a priori* estimates for the Dirichlet problem (Theorems 12.1 and 12.2) that plays an essential role in the proof of the unique solvability theorem of the Dirichlet problem in Chapter 15. The proofs of Theorems 12.1 and 12.2 are flowcharted. Our approach can be traced back to the pioneering work of Schauder [64] and [65] on the Dirichlet problem for second-order, elliptic differential operators.

Chapter 13 is devoted to the proof of Theorem 12.1 (Theorem 13.3) that is based on some local interior *a priori* estimates for the solutions of the Dirichlet problem (Lemma 13.2). The main idea of proof may be considered as an integral perturbation about the constant coefficient case, which goes back to Eugenio Elia Levi [42] (Theorem 13.1). The VMO assumption on the coefficients is of the greatest relevance in the study of an error term expressed by singular commutators (Corollary 11.5). The desired interior *a priori* estimate (12.2) follows in a standard way from Lemma 13.2 by a covering argument if we make use of Sobolev's imbedding theorem (Theorem 7.3), the contraction mapping principle (Theorem 2.1) and the interpolation inequality (Theorem 13.4).

In Chapter 14, we prove the global *a priori* estimate for the homogeneous Dirichlet problem stated in Theorem 12.2. The desired global estimate is consequences of the following two facts (I) and (II):

(I) The explicit representation formula (14.2) for the solutions of the homogeneous Dirichlet problem, which is obtained from the half space Green function, involves the same integral operators as in the interior case.

(II) An L^p boundedness of the singular integral operators and boundary commutators appearing in formula (14.2) (Theorems 14.2 and 14.5).

The results of Chapter 14 are flowcharted.

Chapter 15 is devoted to the study of the *homogeneous* Dirichlet problem for a second-order, uniformly elliptic differential operator with VMO coefficients in the framework of Sobolev spaces of L^p type. We prove an existence and uniqueness theorem for the homogeneous Dirichlet problem (Theorem 15.1). Our proof is based on some interior and boundary *a priori* estimates for the solutions of problem (15.2) (Theorem 12.1 and Theorem 12.2). Both the interior and boundary *a priori* estimates are consequences of explicit representation formulas (12.1) and (13.1) for the solutions of the homogeneous Dirichlet problem (Theorem 13.1 and Theorem 14.1) and also of the L^p-boundedness of Calderón–Zygmund singular integral operators and boundary commutators appearing in those representation formulas (Theorem 14.2 and Theorem 14.5). It should be emphasized that we make use of the method of continuity (Theorem 2.14) in order to prove the existence theorem for the homogeneous Dirichlet problem.

The fourth part (Chapters 16–20) is devoted to the study of the *regular* oblique derivative problem for a second-order, uniformly elliptic differential operator with discontinuous coefficients

$$\begin{cases} \mathcal{L}u(x) = \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} = f(x) & \text{for almost all } x \in \Omega, \\ \mathcal{B}u(x') = \frac{\partial u}{\partial \ell} + \sigma(x')u = \varphi(x') & \text{in the sense of traces on } \partial\Omega \end{cases}$$

in the framework of Sobolev spaces of L^p type. More precisely, we consider a second-order, uniformly elliptic differential operator with VMO coefficients and an oblique derivative boundary operator that is *nowhere* tangential to the boundary.

In Chapter 16 we state global regularizing property of the oblique derivative problem in the framework of Sobolev spaces of L^p type (Theorem 16.1). Furthermore, we state an exsitence and uniqueness theorem for the oblique derivative problem in the framework of Sobolev spaces of L^p type (Theorem 16.2).

In Chapter 17, for a given boundary function, we construct a special auxiliary function that satisfies an oblique derivative boundary condition (Lemma 17.1). This result will allow us to represent, locally near the boundary, the solution of the non-homogeneous oblique derivative problem in Chapter 19 (see formula (19.10)). In this way, we are reduced to the study of the *homogeneous* oblique derivative problem.

In Chapter 18 we prove boundary representation formulas for solutions of the homogeneous oblique derivative problem, by using the half-space Green function. The first step is to derive a boundary representation formula for the solution of the homogeneous oblique derivative problem with constant coefficients operators and homogeneous boundary conditions (Lemma 18.1). The second step is to derive integral representation formulas for the second derivatives of solutions of the homogeneous oblique derivative problem for variable coefficients differential operators and constant coefficients boundary operators (Theorem 18.3).

The purpose of Chapter 19 is to prove boundary Sobolev regularity of the solutions of the non-homogeneous oblique derivative problem towards the proof of Theorem 16.1 (Lemma 19.1). A combination of this regularity result with the interior regularity (Theorem 12.1) will prove the main result in the next Chapter 20.

Chapter 20 is devoted to the study of the non-homogeneous oblique derivative problem. We prove an existence and uniqueness theorem (Theorem 20.1). By Lemma 17.1, we are reduced to the study of the *homogeneous* oblique derivative problem. Our proof is based on some interior and boundary *a priori* estimates for the solutions of the homogeneous oblique derivative problem (Theorem 12.1 and Lemma 19.1). Both the interior and boundary *a priori* estimates are consequences of explicit representation formulas (19.10) and (19.11) for the solutions of homogeneous oblique derivative problem and also of the L^p-boundedness of Calderón–Zygmund singular integral operators and boundary commutators appearing in those representation formulas (Theorem 14.2 and Theorem 14.5). We make use of the method of continuity (Theorem 2.14) in order to prove the existence theorem for the homogeneous oblique derivative problem.

The fifth and final part (Chapters 21–25) is devoted to the functional analytic approach to the problem of construction of Markov processes with first-order Ventcel' boundary condition for second-order, uniformly elliptic differential operators with *discontinuous* coefficients. Our approach is distinguished by the extensive use of the ideas and techniques in the Calderón–Zygmund theory of singular integral operators with non-smooth kernels developed in Parts II and III.

Chapter 21 provides a brief description of the basic definitions and results about a class of semigroups (Feller semigroups) associated with Markov processes in probability, which forms a functional analytic background for the proof of Theorem 1.1. In particular, we formulate a version of the Hille–Yosida theorem adapted to the present context (Theorem 21.9). Moreover, we give two useful criteria in order that a linear operator be the infinitesimal generator of a Feller semigroup (Theorem 21.11 and Corollary 21.12).

In Chapter 22 we consider the Dirichlet problem for the diffusion operator with VMO coefficients in the framework of Sobolev spaces of L^p type, and prove an existence and uniqueness theorem for the Dirichlet problem (Theorem 22.2). The uniqueness result in Theorem 22.2 follows from a variant of the Bakel'man–Aleksandrov maximum principle in the framework of Sobolev spaces due to Bony [9] (Theorem 8.1). Moreover, we construct a Feller semigroup associated with absorption phenomenon at the boundary (Theorem 1.3).

In Chapter 23 we study the oblique derivative problem in the framework of Sobolev spaces of L^p type, and prove un existence and uniqueness theorem for the oblique derivative problem with VMO coefficients (Theorem 23.2). The uniqueness result in Theorem 23.2 follows from a variant of the Bakel'man–Aleksandrov maximum principle in the framework of Sobolev spaces due to Lieberman [43] (Theorem 23.5). Moreover, we construct a Feller semigroup associated with absorption, reflection and drift phenomena at the boundary (Theorem 1.2).

The purpose of Chapter 24 is to prove a general existence theorem for Feller semigroups with Ventcel' boundary condition in terms of boundary value problems (Theorem 24.9). Intuitively, Theorem 24.9 tells us that we can "piece together" a Markov process on the boundary with a diffusion in the interior to construct a Markov process on the closure of the domain.

Chapter 25 is devoted to the proof of Theorem 1.1. The crucial point in the proof is that we consider the term of sticking in the boundary condition as a term of "perturbation" of the oblique derivative boundary condition. More precisely, we make use of a generation theorem for Feller semigroups with oblique derivative boundary condition to verify all the conditions in a version of the Hille–Yosida theorem (Theorem 21.9).

In the last Chapter 26 we give two overviews for general results on generation theorems for Feller semigroups based on the theory of pseudo-differential operators [73], [74], [83], [79] and [80] and based on the theory of singular integral operators [75], [76], [77] and [83], respectively.

Bibliographical references are discussed primarily in Notes and Comments at the end of each chapter. These notes are intended to supplement the text and place it in better perspective.

1.4 A Bird's Eye View of the Contents

The following diagram gives a bird's eye view of Markov processes, Feller semigroups and boundary value problems and how these relate to each other:

Probability	**Functional analysis**	**Boundary value problems**
Strong Markov process (X_t)	Feller semigroup $\{T_t\}$	Infinitesimal generator $\mathfrak{A}$
Markov transition function $p_t(x,\cdot)$	$T_t f(x) = \int p_t(x,dy) f(y)$	$T_t = \exp[t\mathfrak{A}]$
Chapman–Kolmogorov equation	Semigroup property $T_{t+s} = T_t \cdot T_s$	Diffusion operator A
Various diffusion phenomena	Function spaces $C(\overline{\Omega})$, $C_0(\overline{\Omega})$	Ventcel' condition L

Table 1.1. *A bird's eye view of Markov processes, Feller semigroups and boundary value problems*

The paper [85] is devoted to the functional analytic approach to the problem of construction of Feller semigroups in the *characteristic case* via the Fichera function. Probabilistically, our result may be stated as follows: We construct a Feller semigroup corresponding to such a diffusion phenomenon that a Markovian particle moves continuously in the interior of the state space, without reaching the boundary (see [85, Theorem 1.2]). We make use of the Hille–Yosida–Ray theorem (Theorem 21.11) that is a Feller semigroup version of the classical Hille–Yosida theorem in terms of the positive maximum principle. Our proof is based on a method of *elliptic regularizations* essentially due to Oleĭnik and Radkevič [55].

Part I

A Short Course on Functional Analysis and Real Analysis

2
Elements of Functional Analysis

This chapter is devoted to a review of standard topics from functional analysis such as quasinormed and normed linear spaces and closed and continuous (bounded) linear operators between Banach spaces. Most of the material will be quite familiar to the reader and may be omitted. This chapter, included for the sake of completeness, should serve to settle questions of notation and such.

Section 2.5 is devoted to the three pillars of functional analysis – Banach's open mapping theorem (Theorem 2.36), Banach's closed graph theorem (Theorem 2.37) and Banach's closed range theorem for closed operators (Theorem 2.40). Section 2.8 is devoted to the Riesz–Schauder theory for compact operators (Theorem 2.47). In Section 2.9 we state important properties of Fredholm operators (Theorems 2.48 through 2.53). In Section 2.10 we formulate the Riesz representation theorem for bounded linear functionals on a Hilbert space (Theorem 2.58). In the last Section 2.11 we formulate two fundamental theorems concerning spaces of continuous functions defined on a metric space – the Ascoli–Arzelà theorem (Theorem 2.67) and the Stone–Weierstrass theorem (Theorem 2.70).

2.1 Metric Spaces and the Contraction Mapping Principle

A set X is called a *metric space* if there is defined a real-valued function ρ on the Cartesian product $X \times X$ such that

(D1) $0 \leq \rho(x,y) < +\infty$;
(D2) $\rho(x,y) = 0$ if and only if $x = y$;
(D3) $\rho(x,y) = \rho(y,x)$ (symmetry);
(D4) $\rho(x,y) \leq \rho(x,z) + \rho(y,z)$ (the triangle inequality).

The function ρ is called a *metric* or *distance function* on X.

If $x \in X$ and $\varepsilon > 0$, then $B(x;\varepsilon)$ will denote the open ball of radius ε about x, that is,

$$B(x;\varepsilon) = \{y \in X : \rho(x,y) < \varepsilon\}.$$

The countable family

$$\left\{B\left(x;\frac{1}{n}\right) : n \in \mathbf{N}\right\}$$

of open balls forms a fundamental neighborhood system of x; hence a metric space is a *topological space* which satisfies the first axiom of countability.

A topological space X is said to be *metrizable* if we can introduce a metric ρ on X in such a way that the induced topology on X by ρ is just the original topology on X.

Two metrics ρ_1 and ρ_2 on the same set X are said to be *equivalent* if, for each $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\begin{cases} \rho_1(x,y) < \delta \Longrightarrow \rho_2(x,y) < \varepsilon, \\ \rho_2(x,y) < \delta \Longrightarrow \rho_1(x,y) < \varepsilon. \end{cases}$$

Equivalent metrics induce the same topology on X.

If x is a point of X and A is a subset of X, then we define the *distance* $\operatorname{dist}(x,A)$ from x to A by the formula

$$\operatorname{dist}(x,A) = \inf_{a \in A} \rho(x,a).$$

Let (X,ρ) be a metric space. A sequence $\{x_n\}$ in X is called a *Cauchy sequence* if it satisfies Cauchy's convergence condition

$$\lim_{n,m\to\infty} \rho(x_n,x_m) = 0.$$

A metric space X is said to be *complete* if every Cauchy sequence in X converges to a point in X.

Let (X,ρ) be a metric space. A map T from a subset X_0 of X into X is called a *contraction* on X_0 if there exists a number $0 < \theta < 1$ such that

$$\rho(T(x),T(y)) \leq \theta\, \rho(x,y) \quad \text{for all } x,\, y \in X_0. \tag{2.1}$$

The next theorem is the basis of many important existence theorems in analysis (cf. [30, Chapter 3, Theorem 3.8.2]):

Theorem 2.1 (the contraction mapping principle)**.** *Let T be a map of a* complete *metric space (X, ρ) into itself. If T is a* contraction*, then there exists a unique point $z \in X$ such that $T(z) = z$.*

A point z for which $T(z) = z$ is called a *fixed point* of T. Hence Theorem 2.1 is also called a *fixed point theorem*.

2.2 Linear Operators and Functionals

Let X, Y be linear spaces over the same scalar field $\mathbf{K}$. A mapping T defined on a linear subspace $\mathcal{D}$ of X and taking values in Y is said to be *linear* if it preserves the operations of addition and scalar multiplication:

(L1) $T(x_1 + x_2) = Tx_1 + Tx_2$ for all $x_1, x_2 \in \mathcal{D}$.

(L2) $T(\alpha x) = \alpha Tx$ for all $x \in \mathcal{D}$ and $\alpha \in \mathbf{K}$.

We often write Tx, rather than $T(x)$, if T is linear. We let

$$\begin{aligned} D(T) &= \mathcal{D}, \\ R(T) &= \{Tx : x \in D(T)\}, \\ N(T) &= \{x \in D(T) : Tx = 0\}, \end{aligned}$$

and call them the *domain*, the *range* and the *null space* of T, respectively. The mapping T is called a *linear operator* from $D(T) \subset X$ into Y. We also say that T is a linear operator from X into Y with domain $D(T)$. In In the particular case when $Y = \mathbf{K}$, the mapping T is called a *linear functional* on $D(T)$. In other words, a linear functional is a $\mathbf{K}$-valued function on $D(T)$ that satisfies conditions (L1) and (L2).

If a linear operator T is a one-to-one map of $D(T)$ onto $R(T)$, then it is easy to see that the inverse mapping T^{-1} is a linear operator on $R(T)$ onto $D(T)$. The mapping T^{-1} is called the *inverse operator* or simply the *inverse* of T. A linear operator T admits the inverse T^{-1} if and only if $Tx = 0$ implies that $x = 0$.

Let T_1 and T_2 be two linear operators from a linear space X into a linear space Y with domains $D(T_1)$ and $D(T_2)$, respectively. Then we say that $T_1 = T_2$ if and only if $D(T_1) = D(T_2)$ and $T_1 x = T_2 x$ for all $x \in D(T_1) = D(T_2)$. If $D(T_1) \subset D(T_2)$ and $T_1 x = T_2 x$ for all $x \in D(T_1)$, then we say that T_2 is an *extension* of T_1 and also that T_1 is a *restriction* of T_2, and we write $T_1 \subset T_2$.

2.3 Quasinormed Linear Spaces

Let X be a linear space over the real or complex number field $\mathbf{K}$. A real-valued function p defined on X is called a *seminorm* on X if it satisfies the following three conditions (S1), (S2) and (S3):

(S1) $0 \leq p(x) < \infty$ for all $x \in X$.
(S2) $p(\alpha x) = |\alpha| p(x)$ for all $\alpha \in \mathbf{K}$ and $x \in X$.
(S3) $p(x+y) \leq p(x) + p(y)$ for all $x, y \in X$.

Let $\{p_i\}$ be a countable family of seminorms on X such that

$$p_1(x) \leq p_2(x) \leq \cdots \leq p_i(x) \leq \cdots \quad \text{for each } x \in X, \tag{2.2}$$

and define

$$V_{ij} = \left\{ x \in X : p_i(x) < \frac{1}{j} \right\}, \quad i, j = 1, 2, \ldots.$$

Then it is easy to verify that a countable family of the sets

$$x + V_{ij} = \{x + y : y \in V_{ij}\}$$

satisfies the axioms of a fundamental neighborhood system of x; hence X is a topological space which satisfies the first axiom of countability.

Furthermore, we have the following:

Theorem 2.2. *Let $\{p_i\}$ be a countable family of seminorms on a linear space X which satisfies condition* (2.2). *Assume that*

For every non-zero $x \in X$, there exists a seminorm p_i such that $p_i(x) > 0$. (2.3)

Then the space X is metrizable *by the metric*

$$\rho(x, y) = \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{p_i(x)}{1 + p_i(x)} \quad \text{for all } x, y \in X.$$

If we let

$$|x| = \rho(x, 0) = \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{p_i(x)}{1 + p_i(x)} \quad \text{for } x \in X, \tag{2.4}$$

then the quantity $|x|$ enjoys the following four properties (Q1), (Q2), (Q3) and (Q4):

(Q1) $|x| \geq 0$; $|x| = 0$ *if and only if* $x = 0$.
(Q2) $|x + y| \leq |x| + |y|$ *(the triangle inequality).*
(Q3) $\alpha_n \to 0$ *in* $\mathbf{K} \Longrightarrow |\alpha_n x| \to 0$ *for every* $x \in X$.

(Q4) $|x_n| \to 0 \Longrightarrow |\alpha x_n| \to 0$ *for every* $\alpha \in \mathbf{K}$.

This quantity $|x|$ is called a *quasinorm* of x, and the space X is called a *quasinormed linear space.*

Theorem 2.2 may be restated as follows:

Theorem 2.3. *A linear space* X*, topologized by a countable family* $\{p_i\}$ *of seminorms satisfying conditions* (2.2) *and* (2.3)*, is a quasinormed linear space with respect to the quasinorm* $|x|$ *defined by formula* (2.4)*.*

Let X be a quasinormed linear space. The convergence

$$\lim_{n\to\infty} |x_n - x| = 0$$

in X is denoted by $s - \lim_{n\to\infty} x_n = x$ or simply by $x_n \to x$, and we say that the sequence $\{x_n\}$ *converges strongly* to x. A sequence $\{x_n\}$ is called a *Cauchy sequence* if it satisfies Cauchy's condition

$$\lim_{m,n\to\infty} |x_m - x_n| = 0.$$

A quasinormed linear space X is called a *Fréchet space* if it is complete, that is, if every Cauchy sequence in X converges strongly to a point in X. If a quasinormed linear space X is topologized by a countable family $\{p_i\}$ of seminorms which satisfies conditions (2.2) and (2.3), then the above definitions may be reformulated in terms of seminorms as follows:

(i) A sequence $\{x_n\}$ in X converges strongly to a point x in X if and only if, for every seminorm p_i and every $\varepsilon > 0$, there exists a positive integer $N = N(i, \varepsilon)$ such that

$$n \geq N \Longrightarrow p_i(x_n - x) < \varepsilon.$$

(ii) A sequence $\{x_n\}$ in X is a Cauchy sequence if and only if, for every seminorm p_i and every $\varepsilon > 0$, there exists a positive integer $N = N(i, \varepsilon)$ such that

$$m,\, n \geq N \Longrightarrow p_i(x_m - x_n) < \varepsilon.$$

Let X be a quasinormed linear space. A linear subspace of X is called a *closed subspace* if it is a closed subset of X. For example, the closure $\overline{M}$ of a linear subspace M is a closed subspace. Indeed, since the elements of $\overline{M}$ are limits of sequences in M, we have the assertions

$$\begin{cases} x = \lim_{n\to\infty} x_n, & x_n \in M, \\ y = \lim_{n\to\infty} y_n, & y_n \in M \end{cases}$$

$$\Longrightarrow$$

$$\begin{cases} x + y = \lim_{n\to\infty}(x_n + y_n), \\ \alpha x = \lim_{n\to\infty} \alpha x_n \quad \text{for all } \alpha \in \mathbf{K}. \end{cases}$$

This proves that $x + y \in \overline{M}$ and $\alpha x \in \overline{M}$ for all $\alpha \in \mathbf{K}$.

2.3.1 Compact Sets

A collection $\{U_\lambda\}_{\lambda\in\Lambda}$ of open sets of a topological space X is called an *open covering* of X if $X = \cup_{\lambda\in\Lambda} U_\lambda$. A topological space X is said to be *compact* if every open covering $\{U_\lambda\}$ of X contains some finite subcollection of $\{U_\lambda\}$ which still covers X. If s subset of X is compact considered as a topological subspace of X, then it is called a *compact subset* of X.

A subset of a topological space X is said to be *relatively compact* (or *precompact*) if its closure is a compact subset of X. A topological space X is said to be *locally compact* if every point of X has a relatively compact neighborhood.

A subset of a topological space X is called a *σ-compact subset* if it is a countable union of compact sets.

Compactness is such a useful property that, given a non-compact space $(X, \mathcal{O})$, it is worthwhile constructing a compact space $(X', \mathcal{O}')$ with X being its dense subset. Such a space is called a *compactification* of $(X, \mathcal{O})$. The simplest way in which this can be achieved is by adjoining one extra point ∞ to the space X; a topology $\mathcal{O}'$ can be defined on $X' = X \cup \{\infty\}$ in such a way that $(X', \mathcal{O}')$ is compact and that $\mathcal{O}$ is the relative topology induced on X by $\mathcal{O}'$. The topological space $(X', \mathcal{O}')$ is called the *one-point compactification* of $(X, \mathcal{O})$, and the point ∞ is called the *point at infinity.*

Let X be a quasinormed linear space. A subset Y of X is called a *sequentially compact* if every sequence $\{y_n\}$ in Y contains a subsequence $\{y_{n'}\}$ which converges to a point y of Y:

$$\lim_{n'\to\infty} |y_{n'} - y| = 0.$$

Then we have the following criterion for compactness:

Theorem 2.4. *A subset of a quasinormed linear space X is compact if and only if it is sequentially compact.*

2.3.2 Bounded Sets

Let $(X, |\cdot|)$ be a quasinormed linear space. A set B in X is said to be *bounded* if it satisfies the condition

$$\sup_{x \in B} |x| < \infty.$$

We remark that every compact set is bounded.

A subset K of X is said to be *totally bounded* if, for any given $\varepsilon > 0$ there is a finite number of balls

$$B(x_i, \varepsilon) = \{x \in X : |x - x_i| < \varepsilon\}, \quad 1 \leq i \leq n,$$

of radius ε about $x_i \in X$ that cover K:

$$K \subset \cup_{i=1}^{n} B(x_i, \varepsilon).$$

Example 2.1. Let $(X, |\cdot|)$ be a quasinormed linear space. Assume that a subset A satisfies the following three conditions:

(a) For every $h > 0$, there exists a totally bounded subset A_h of X.
(b) For each point $x \in A$, there exists a point $y \in A_h$ such that $|x - y| \leq h$.
(c) For each point $z \in A_h$, there exists a point $w \in A$ such that $|z - w| \leq h$.

Then the subset A is totally bounded.

Proof. For any given $\varepsilon > 0$, we assume that

$$A \subset \cup_{x \in A} B(x, \varepsilon).$$

Choose a number h_0 such that

$$0 < h_0 < \frac{\varepsilon}{3}.$$

Since we have the assertion

$$A_{h_0} \subset \cup_{y \in A_{h_0}} B\left(y, \varepsilon/3\right),$$

it follows from condition (a) that there is a finite number of points $\{y_1, \ldots, y_N\}$ of A_{h_0} such that

$$A_{h_0} \subset \cup_{i=1}^{N} B\left(y_i, \varepsilon/3\right). \tag{2.5}$$

Moreover, by condition (c) we can find a finite number of points

$$\{x_1, \ldots, x_N\}$$

of A such that

$$|x_i - y_i| \leq h_0, \quad 1 \leq i \leq N.$$

Then we have the assertion

$$A \subset \cup_{i=1}^{N} B(x_i, \varepsilon), \tag{2.6}$$

which proves that A is totally bounded.

Indeed, if x is an arbirary point of A, by condition (b) we can find a point $y_0 \in A_{h_0}$ such that

$$|x - y_0| \leq h_0.$$

Moreover, by assertion (2.5), there exists a point $y_i \in A_{h_0}$ such that

$$|y_0 - y_i| \leq \frac{\varepsilon}{3}.$$

Hence we have the inequality

$$\begin{aligned}|x - x_i| &\leq |x - y_0| + |y_0 - y_i| + |y_i - x_i| \leq h_0 + \frac{\varepsilon}{3} + h_0 \\ &< \varepsilon.\end{aligned}$$

This proves the desired assertion (2.6).

The proof of Example 2.1 is complete. □

Finally, we have the following criterion for compactness:

Theorem 2.5. *Let X be a complete quasinormed linear space. A closed subset of X is compact if and only if it is totally bounded.*

Throughout the rest of this section, let X and Y be quasinormed linear spaces over the same scalar field $\mathbf{K}$, topologized respectively by countable families $\{p_i\}$ and $\{q_i\}$ of seminorms which satisfy conditions (2.2) and (2.3).

2.3.3 Continuity of Linear Operators

Let T be a linear operator from X into Y with domain $D(T)$. By virtue of the linearity of T, it follows that T is continuous everywhere on $D(T)$ if and only if it is continuous at one point of $D(T)$. Furthermore, we have the following:

Theorem 2.6. *A linear operator T from X into Y with domain $D(T)$ is continuous everywhere on $D(T)$ if and only if, for every seminorm q_j on Y, there exist a seminorm p_i on X and a constant $C > 0$ such that*

$$q_j(Tx) \leq C p_i(x) \quad \textit{for all } x \in D(T).$$

2.3.4 Topologies of Linear Operators

We let

$L(X, Y) =$ the collection of continuous linear operators on X into Y.

We define in the set $L(X, Y)$ addition and scalar multiplication of operators in the usual way:

$$(T + S)x = Tx + Sx, \quad x \in X,$$
$$(\alpha T)x = \alpha(Tx), \quad \alpha \in \mathbf{K},\ x \in X.$$

Then $L(X, Y)$ is a linear space.

We introduce three different topologies on the space $L(X, Y)$.

(1) *Simple convergence topology*: This is the topology of convergence at each point of X; a sequence $\{T_n\}$ in $L(X, Y)$ converges to an element T of $L(X, Y)$ in the simple convergence topology if and only if $T_n x \to Tx$ in Y for each $x \in X$.
(2) *Compact convergence topology*: This is the topology of uniform convergence on compact sets in X; $T_n \to T$ in the compact convergence topology if and only if $T_n x \to Tx$ in Y uniformly for x ranging over compact sets in X.
(3) *Bounded convergence topology*: This is the topology of uniform convergence on bounded sets in X; $T_n \to T$ in the bounded convergence topology if and only if $T_n x \to Tx$ in Y uniformly for x ranging over bounded sets in X.

The simple convergence topology is weaker than the compact convergence topology, and the compact convergence topology is weaker than the bounded convergence topology.

2.3.5 Product Spaces

Let X and Y be quasinormed linear spaces over the same scalar field $\mathbf{K}$. Then the Cartesian product $X \times Y$ becomes a linear space over $\mathbf{K}$ if we define the algebraic operations coordinatewise

$$\{x_1, y_1\} + \{x_2, y_2\} = \{x_1 + x_2, y_1 + y_2\},$$
$$\alpha \{x, y\} = \{\alpha x, \alpha y\} \quad \text{for } \alpha \in \mathbf{K}.$$

It is easy to verify that the quantity

$$|\{x, y\}| = \left(|x|_X^2 + |y|_Y^2\right)^{1/2} \tag{2.7}$$

satisfies axioms (Q1) through (Q4) of a quasinorm; hence the product space $X \times Y$ is a quasinormed linear space with respect to the quasinorm defined by formula (2.7). Furthermore, if X and Y are Fréchet spaces, then so is $X \times Y$. In other words, the completeness is inherited by the product space.

2.4 Normed Linear Spaces

A quasinormed linear space is called a *normed linear space* if it is topologized by just one seminorm that satisfies condition (2.3). We give the precise definition of a normed linear space.

Let X be a linear space over the real or complex number field $\mathbf{K}$. A real-valued function $\|\cdot\|$ defined on X is called a *norm* on X if it satisfies the following three conditions (N1), (N2) and (N3):

(N1) $\|x\| \geq 0$; $\|x\| = 0$ if and only if $x = 0$.
(N2) $\|\alpha x\| = |\alpha|\|x\|$, $\alpha \in \mathbf{K}$, $x \in X$.
(N3) $\|x + y\| \leq \|x\| + \|y\|$, x, $y \in X$ (the triangle inequality).

A linear space X equipped with a norm $\|\cdot\|$ is called a normed linear space. The topology on X is defined by the metric

$$\rho(x, y) = \|x - y\|.$$

The convergence

$$\lim_{n\to\infty} \|x_n - x\| = 0$$

in X is denoted by $s - \lim_{n\to\infty} x_n = x$ or simply $x_n \to x$, and we say that the sequence $\{x_n\}$ converges strongly to x. A sequence $\{x_n\}$ in X is called a *Cauchy sequence* if it satisfies the condition

$$\lim_{n,m\to\infty} \|x_n - x_m\| = 0.$$

A normed linear space X is called a *Banach space* if it is complete, that is, if every Cauchy sequence in X converges strongly to a point in X.

Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ defined on the same linear space X are said to be *equivalent* if there exist constants $c > 0$ and $C > 0$ such that

$$c\,\|x\|_1 \leq \|x\|_2 \leq C\,\|x\|_1 \quad \text{for all } x \in X.$$

Equivalent norms induce the same topology.

If X and Y are normed linear spaces over the same scalar field $\mathbf{K}$, then the product space $X \times Y$ is a normed linear space by the norm

$$\|\{x, y\}\| = \left(\|x\|_X^2 + \|y\|_Y^2\right)^{1/2}.$$

If X and Y are Banach spaces, then so is $X \times Y$.

Let X be a normed linear space. If Y is a closed linear subspace of X, then the factor space X/Y is a normed linear space by the norm

$$\|\widetilde{x}\| = \inf_{z \in \widetilde{x}} \|z\|. \tag{2.8}$$

If X is a Banach space, then so is X/Y. The space X/Y, normed by formula (2.8), is called a *normed factor space*.

2.4.1 Linear Operators on Normed Spaces

Throughout the rest of this section, the letters X, Y, Z denote normed linear spaces over the same scalar field $\mathbf{K}$.

The next theorem is a normed linear space version of Theorem 2.6:

Theorem 2.7. *Let T be a linear operator from X into Y with domain $D(T)$. Then T is continuous everywhere on $D(T)$ if and only if there exists a constant $C > 0$ such that*

$$\|Tx\| \leq C\|x\| \quad \text{for all } x \in D(T). \tag{2.9}$$

Remark 2.1. In inequality (2.9), the quantity $\|x\|$ is the norm of x in X and the quantity $\|Tx\|$ is the norm of Tx in Y. Frequently several norms appear together, but it is clear from the context which is which.

One of the consequences of Theorem 2.7 is the following extension theorem for a continuous linear operator:

Theorem 2.8. *If T is a continuous linear operator from X into Y with domain $D(T)$ and if Y is a Banach space, then T has a unique continuous extension $\widetilde{T}$ whose domain is the closure $\overline{D(T)}$ of $D(T)$.*

As another consequence of Theorem 2.7, we give a necessary and sufficient condition for the existence of the continuous inverse of a linear operator:

Theorem 2.9. *Let T be a linear operator from X into Y with domain $D(T)$. Then T admits a continuous inverse T^{-1} if and only if there exists a constant $c > 0$ such that*

$$\|Tx\| \geq c\|x\| \quad \text{for all } x \in D(T).$$

A linear operator T from X into Y with domain $D(T)$ is called an *isometry* if it is norm-preserving, that is, if we have the formula

$$\|Tx\| = \|x\| \quad \text{for all } x \in D(T).$$

It is clear that if T is an isometry, then it is injective and both T and T^{-1} are continuous.

If T is a continuous, one-to-one linear mapping of X onto Y and if its inverse T^{-1} is also a continuous mapping, then it is called an *isomorphism* of X onto Y. Two normed linear spaces are said to be *isomorphic* if there is an isomorphism between them.

By combining Theorems 2.7 and 2.9, we obtain the following:

Theorem 2.10. *Let T be a linear operator on X onto Y. Then T is an isomorphism if and only if there exist constants $c > 0$ and $C > 0$ such that*

$$c\|x\| \leq \|Tx\| \leq C\|x\| \quad \textit{for all } x \in X.$$

If T is a continuous linear operator from X into Y with domain $D(T)$, we let

$$\|T\| = \inf\{C : \|Tx\| \leq C\|x\|,\ x \in D(T)\}.$$

Then, in view of the linearity of T we have the formula

$$\|T\| = \sup_{\substack{x \in D(T) \\ x \neq 0}} \frac{\|Tx\|}{\|x\|} = \sup_{\substack{x \in D(T) \\ \|x\|=1}} \|Tx\| = \sup_{\substack{x \in D(T) \\ \|x\| \leq 1}} \|Tx\|. \tag{2.10}$$

This proves that $\|T\|$ is the smallest non-negative number such that

$$\|Tx\| \leq \|T\| \cdot \|x\| \quad \text{for all } x \in D(T). \tag{2.11}$$

Theorem 2.7 asserts that a linear operator T on X into Y is continuous if and only if it maps bounded sets in X into bounded sets in Y. Thus a continuous linear operator on X into Y is usually called a *bounded linear operator* on X into Y. We let

$$\begin{aligned} &L(X, Y) \\ &= \text{the space of bounded (continuous) linear operators on } X \text{ into } Y. \end{aligned}$$

In the case of normed linear spaces, the simple convergence topology on $L(X, Y)$ is usually called the *strong topology* of operators, and the bounded convergence topology on $L(X, Y)$ is called the *uniform topology* of operators. In view of formulas (2.10) and (2.11), it follows that the quantity $\|T\|$ satisfies axioms (N1), (N2) and (N3) of a norm; hence the space $L(X, Y)$ is a normed linear space by the norm $\|T\|$ given by formula (2.10). The topology on $L(X, Y)$ induced by the operator norm $\|T\|$ is just the uniform topology of operators.

We give a sufficient condition for the space $L(X, Y)$ to be complete:

Theorem 2.11. *If Y is a Banach space, then so is $L(X, Y)$.*

If T is a linear operator from X into Y with domain $D(T)$ and S is a linear operator from Y into Z with domain $D(S)$, then we define the product ST as follows:

(a) $D(ST) = \{x \in D(T) : Tx \in D(S)\}$,
(b) $(ST)(x) = S(Tx)$ for every $x \in D(ST)$.

As for the product of linear operators, we have the following:

Proposition 2.12. *If $T \in L(X, Y)$ and $S \in L(Y, Z)$, then it follows that $ST \in L(X, Z)$. Moreover, we have the inequality*

$$\|ST\| \leq \|S\| \cdot \|T\|.$$

We make use of the following theorem in constructing the bounded inverse of a bounded linear operator:

Theorem 2.13. *If T is a bounded linear operator on a Banach space X into itself and satisfies $\|T\| < 1$, then the operator $I - T$ has a unique bounded linear inverse $(I - T)^{-1}$ which is given by* C. Neumann's series

$$(I - T)^{-1} = \sum_{n=0}^{\infty} T^n.$$

Here I is the identity operator: $Ix = x$ for every $x \in X$, and $T^0 = I$.

2.4.2 Method of Continuity

In Chapter 15 we make use of the following method of continuity (see [33, Chapter 5, Theorem 5.2]) in order to prove the *existence theorem* for the Dirichlet problem:

Theorem 2.14 (the method of continuity)**.** *Let $\mathcal{B}$ be a Banach space and let $\mathcal{V}$ be a normed linear space. If $\mathcal{L}_0$ and $\mathcal{L}_1$ are two bounded linear operators from $\mathcal{B}$ into $\mathcal{V}$, we define a family of bounded linear operators*

$$\mathcal{L}_t = (1 - t)\mathcal{L}_0 + t\,\mathcal{L}_1 : \mathcal{B} \longrightarrow \mathcal{V}$$

for $0 \leq t \leq 1$. Assume that there exists a positive constant C, independent of x and t, such that

$$\|x\|_{\mathcal{B}} \leq C \|\mathcal{L}_t x\|_{\mathcal{V}} \quad \text{for all } x \in \mathcal{B}. \tag{2.12}$$

Then the operator $\mathcal{L}_1$ maps $\mathcal{B}$ onto $\mathcal{V}$ if and only if the operator $\mathcal{L}_0$ maps $\mathcal{B}$ onto $\mathcal{V}$.

Proof. Assume that $\mathcal{L}_s$ is surjective for some $s \in [0,1]$. By inequality (2.12), it follows that $\mathcal{L}_s$ is bijective, so that the inverse $\mathcal{L}_s^{-1} : \mathcal{V} \to \mathcal{B}$ exists. Here we remark that

$$\left\|\mathcal{L}_s^{-1}\right\| \leq C.$$

Now let t be an arbitrary point of the interval $[0,1]$. For any given $y \in \mathcal{V}$, the equation

$$\mathcal{L}_t x = y$$

is equivalent to the equation

$$\mathcal{L}_s x = \mathcal{L}_t x + (\mathcal{L}_s - \mathcal{L}_t)x = y + (s-t)\left(\mathcal{L}_1 x - \mathcal{L}_0 x\right).$$

Hence we have the equivalent assertions

$$\begin{aligned} \mathcal{L}_t x = y &\iff x = \mathcal{L}_s^{-1}\left(y + (s-t)\left(\mathcal{L}_1 x - \mathcal{L}_0 x\right)\right) \\ &\iff \big(I - (s-t)\mathcal{L}_s^{-1}\left(\mathcal{L}_1 - \mathcal{L}_0\right)\big)\, x = \mathcal{L}_s^{-1} y. \end{aligned}$$

However, if $|t-s|$ is so small that

$$|s-t| < \delta := \frac{1}{C\left(\|\mathcal{L}_1\| + \|\mathcal{L}_0\|\right)},$$

then it follows that

$$\begin{aligned} \left\|(s-t)\,\mathcal{L}_s^{-1}\left(\mathcal{L}_1 - \mathcal{L}_0\right)\right\| &\leq |s-t| \left\|\mathcal{L}_s^{-1}\right\| \|\mathcal{L}_1 - \mathcal{L}_0\| \\ &\leq C\, |s-t| \left(\|\mathcal{L}_1\| + \|\mathcal{L}_0\|\right) = \frac{|s-t|}{\delta} \\ &< 1. \end{aligned}$$

This proves that the operator

$$\big(I - (s-t)\,\mathcal{L}_s^{-1}\left(\mathcal{L}_1 - \mathcal{L}_0\right)\big)$$

has, as a Neumann series (Theorem 2.13), the inverse

$$\big(I - (s-t)\,\mathcal{L}_s^{-1}\left(\mathcal{L}_1 - \mathcal{L}_0\right)\big)^{-1} = \sum_{n=0}^{\infty} (s-t)^n \left(\mathcal{L}_s^{-1}\right)^n \left(\mathcal{L}_1 - \mathcal{L}_0\right)^n.$$

Therefore, we obtain that, for all $t \in [0,1]$ satisfying $|t-s| < \delta$,

$$\mathcal{L}_t x = y \iff x = \left(I - (s-t)\mathcal{L}_s^{-1}(\mathcal{L}_1 - \mathcal{L}_0)\right)^{-1} \mathcal{L}_s^{-1} y.$$

By dividing the interval $[0,1]$ into subintervals of length less than δ, we find that the mapping $\mathcal{L}_t$ is surjective for all $t \in [0,1]$, provided that $\mathcal{L}_s$ is surjective for some $s \in [0,1]$. In particular, this proves that $\mathcal{L}_1$ maps $\mathcal{B}$ onto $\mathcal{V}$ if and only if $\mathcal{L}_0$ maps $\mathcal{B}$ onto $\mathcal{V}$.

The proof of Theorem 2.14 is complete. □

2.4.3 Finite Dimensional Spaces

The next theorem asserts that there is no point in studying abstract finite dimensional normed linear spaces:

Theorem 2.15. *All n-dimensional normed linear spaces over the same scalar field $\mathbf{K}$ are isomorphic to $\mathbf{K}^n$ with the maximum norm*

$$\|\alpha\| = \max_{1 \le i \le n} |\alpha_i|, \quad \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbf{K}^n.$$

Topological properties of the space $\mathbf{K}^n$ applies to all finite dimensional normed linear spaces.

Corollary 2.16. *All finite dimensional normed linear spaces are complete.*

Corollary 2.17. *Every finite dimensional linear subspace of a normed linear space is closed.*

Corollary 2.18. *A subset of a finite dimensional normed linear space is compact if and only if it is closed and bounded.*

By Corollary 2.17, it follows that the closed unit ball in a finite dimensional normed linear space is compact. Conversely, this property characterizes finite dimensional spaces:

Theorem 2.19. *If the closed unit ball in a normed linear space X is compact, then X is finite dimensional.*

2.4.4 The Hahn–Banach Extension Theorem

The Hahn–Banach extension theorem asserts the existence of linear functionals dominated by norms (see [99, Chapter IV, Section 5, Theorem 1 and Corollary]):

Theorem 2.20 (Hahn–Banach). *Let X be a normed linear space over the real or complex number field* $\mathbf{K}$ *and let M be a linear subspace of X. If f is a continuous linear functional defined on M, then it can be extended to a continuous linear functional $\widetilde{f}$ on X so that*

$$\|\widetilde{f}\| = \|f\|.$$

Let X be a real or complex, normed linear space. A closed subset M of X is said to be *balanced* if it satisfies the condition

$$x \in M, \ |\alpha| \leq 1 \implies \alpha x \in M.$$

The next two theorems assert the existence of non-trivial continuous linear functionals (see [99, Chapter IV, Section 6, Theorem 3]):

Theorem 2.21 (Mazur). *Let X be a real or complex, normed linear space and let M be a closed convex, balanced subset of X. Then, for any element $x_0 \notin M$ there exists a continuous linear functional f_0 on X such that*

$$\begin{aligned} &f_0(x_0) > 1, \\ &|f_0(x)| \leq 1 \quad \text{on } M. \end{aligned}$$

Proof. Since M is closed and $x_0 \notin M$, it follows that

$$\operatorname{dist}(x_0, M) > 0.$$

If $0 < d < \operatorname{dist}(x_0, M)$, we let

$$\begin{aligned} &B\left(0, \frac{d}{2}\right) = \left\{x \in X : \|x\| \leq \frac{d}{2}\right\}, \\ &B\left(x_0, \frac{d}{2}\right) = x_0 + B\left(0, \frac{d}{2}\right) = \left\{x \in X : \|x - x_0\| \leq \frac{d}{2}\right\}, \\ &U = \left\{x \in X : \operatorname{dist}(x, M) \leq \frac{d}{2}\right\}. \end{aligned}$$

Then we have the assertions

$$\begin{aligned} &U \cap B\left(x_0, \frac{d}{2}\right) = \emptyset, \\ &B\left(0, \frac{d}{2}\right) \subset U, \end{aligned}$$

since $0 \in M$. Moreover, since M is convex and balanced, it is easy to verify the following three assertions:

(a) U is convex.

(b) U is balanced.

(c) U is absorbing, that is, for any $x \in X$, there exists a constant $\alpha > 0$ such that $\alpha^{-1}x \in U$.

Hence, we can define the *Minkowski functional* p_U of U by the formula

$$p_U(x) = \inf \left\{ \alpha > 0 : \alpha^{-1}x \in U \right\} \quad \text{for every } x \in X.$$

Since U is closed, it is easy to verify the following assertions:

$$\begin{cases} p_U(x) > 1 & \text{if } x \notin U, \\ p_U(x) \leq 1 & \text{if } x \in U. \end{cases}$$

Therefore, by applying [99, Chapter IV, Section 6, Corollary 1 to Theorem 1] to our situation we can find a continuous linear functional f_0 on X such that

$$f_0(x_0) = p_U(x_0) > 1,$$
$$|f_0(x)| \leq p_U(x) \quad \text{on } X.$$

In particular, we have the assertion

$$|f_0(x)| = p_U(x) \leq 1 \quad \text{on } M,$$

since $M \subset U$.

The proof of Theorem 2.21 is complete. □

Theorem 2.22. *Let X be a normed linear space and let M be a closed linear subspace of X. Then, for any element $x_0 \notin M$ there exists a continuous linear functional f_0 on X such that*

$$\begin{cases} f_0(x_0) > 1, \\ f_0(x) = 0 & \textit{on } M. \end{cases}$$

Proof. Indeed, it suffices to note that

$$|f_0(x)| \leq 1 \quad \text{on } M \implies f_0(x) = 0 \quad \text{on } M,$$

since M is a linear space.

The proof of Theorem 2.22 is complete. □

Finally, the next theorem asserts that, for each point $x_0 \neq 0$ there exists a continuous linear functional f_0 such that $f_0(x_0) \neq f_0(0) = 0$:

Theorem 2.23. *Let X be a normed linear space. For each non-zero element x_0 of X, there exists a continuous linear functional f_0 on X such that*

$$\begin{cases} f_0(x_0) = \|x_0\|, \\ \|f_0\| = 1. \end{cases}$$

2.4.5 Dual Spaces

Let X be a normed linear space over the real or complex number field $\mathbf{K}$. A continuous linear functional on X is usually called a *bounded linear functional* on X.

The space $L(X, \mathbf{K})$ of all bounded linear functionals on X is called the *dual space* of X, and is denoted by X'. We shall write

$$f(x) = \langle f, x \rangle$$

for the value of the functional $f \in X'$ and the vector $x \in X$. The bounded (resp. simple) convergence topology on X' is called the *strong* (resp. *weak**) *topology* on X' and the dual space X' equipped with this topology is called the *strong* (resp. *weak**) *dual space* of X.

It follows from an application of Theorem 2.11 with $Y := \mathbf{K}$ that the strong dual space X' is a Banach space with the norm

$$\|f\| = \sup_{x \in X \setminus \{0\}} \frac{|f(x)|}{\|x\|} = \sup_{\substack{x \in X \\ \|x\| \leq 1}} |f(x)| .$$

We remark that

$$|\langle f, x \rangle| \leq \|f\| \cdot \|x\| \quad \text{for all } x \in X.$$

Theorem 2.23 asserts that the dual space X' separates points of X, that is, for arbitrary two distinct points x_1, x_2 of X, there exists a functional $f \in X'$ such that $f(x_1) \neq f(x_2)$.

2.4.6 Annihilators

Let A be a subset of a normed linear space X. An element f of the dual space X' is called an annihilator of A if it satisfies the condition

$$f(x) = 0 \quad \text{for all } x \in A.$$

We let

$$A^0 = \{f \in X' : f(x) = 0 \text{ for all } x \in A\}$$

be the set of all annihilators of A. This is not a one way proposition. If B is a subset of X', we let

$$^0B = \{x \in X : f(x) = 0 \text{ for all } f \in B\}$$

be the set of all annihilators of B.

Here are some basic properties of annihilators:

(i) The sets A^0 and ^0B are closed linear subspaces of X and X', respectively.
(ii) If M is a closed linear subspace of X, then $^0(M^0) = M$.
(iii) If A is a subset of X and M is the closure of the subspace spanned by A, then $M^0 = A^0$ and $M = {}^0(A^0)$.

2.4.7 Dual Spaces of Normed Factor Spaces

Let M be a closed linear subspace of a normed linear space X. Then each element f of M^0 defines a bounded linear functional $\widetilde{f}$ on the normed factor space X/M by the formula

$$\widetilde{f}(\widetilde{x}) = f(x) \quad \text{for all } \widetilde{x} \in X/M.$$

Indeed, the value $f(x)$ on the right-hand side does not depend on the choice of a representative x of the equivalence class $\widetilde{x}$, and we have the formula

$$\|\widetilde{f}\| = \|f\|.$$

Furthermore, it is easy to see that the mapping

$$\pi\colon f \longmapsto \widetilde{f}$$

of M^0 into $(X/M)'$ is linear and surjective; hence we have the following:

Theorem 2.24. *The strong dual space $(X/M)'$ of the factor space X/M can be identified with the space M^0 of all annihilators of M by the linear isometry π.*

2.4.8 Bidual Spaces

Each element x of a normed linear space X defines a bounded linear functional Jx on the strong dual space X' by the formula

$$Jx(f) = f(x) \quad \text{for all } f \in X'. \tag{2.13}$$

Then Theorem 2.23 asserts that

$$\|Jx\| = \sup_{\substack{f \in X' \\ \|f\| \leq 1}} |Jx(f)| = \|x\|,$$

so that the mapping J is a linear isometry of X into the strong dual space $(X')'$ of X'. The space $(X')'$ is called the *strong bidual* (or *second dual*) *space* of X.

Summing up, we have the following:

Theorem 2.25. *A normed linear space X can be embedded into its strong bidual space $(X')'$ by the linear isometry J defined by formula* (2.13).

If the mapping J is surjective, that is, if $X = (X')'$, then we say that X is *reflexive.*

For example, we have the following:

Theorem 2.26. *The space $L^p(\Omega)$ is reflexive if and only if $1 < p < \infty$ (see [2, Theorem 2.46]).*

2.4.9 Weak Convergence

A sequence $\{x_n\}$ in a normed linear space X is said to be *weakly convergent* if a finite $\lim_{n\to\infty} f(x_n)$ exists for each f in the dual space X' of X. A sequence $\{x_n\}$ in X is said to *converge weakly* to an element x of X if $\lim_{n\to\infty} f(x_n) = f(x)$ for every $f \in X'$; we then write $w - \lim_{n\to\infty} x_n = x$ or simply $x_n \to x$ weakly. Since the space X' separates points of X, the limit x is uniquely determined. Theorem 2.25 asserts that X may be considered as a linear subspace of its bidual space $(X')'$; hence the *weak topology* on X is just the simple convergence topology on the bidual space $(X')' = L(X', \mathbf{K})$.

For weakly convergent sequences, we have the following:

Theorem 2.27. *(i) $s - \lim_{n\to\infty} x_n = x$ implies $w - \lim_{n\to\infty} x_n = x$.*

(ii) A weakly convergent sequence $\{x_n\}$ is bounded:

$$\sup_n \|x_n\| < +\infty.$$

Furthermore, if $w - \lim_{n\to\infty} x_n = x$, then the sequence $\{x_n\}$ is bounded and we have the inequality

$$\|x\| \leq \liminf_{n\to\infty} \|x_n\|.$$

Part (ii) of Theorem 2.27 has a converse:

Theorem 2.28. *A sequence $\{x_n\}$ in X converges weakly to an element x of X if the following two conditions (a) and (b) are satisfied:*

(a) The sequence $\{x_n\}$ is bounded.
(b) $\lim_{n\to\infty} f(x_n) = f(x)$ for every f in some strongly dense subset of X'.

For the strong and weak closures of a linear subspace, we have the following (see [99, Chapter V, Section 1, Theorem 11]):

Theorem 2.29 (Mazur)**.** *Let X be a normed linear space. If M is a closed linear subspace of X in the strong topology of X, then it is closed in the weak topology of X.*

Proof. Our proof is based on a reduction to absurdity. Assume, to the contrary, that M is not weakly closed. Then there exists a point $x_0 \in X \setminus M$ such that x_0 is an accumulation point of the set M in the weak topology of X. Namely, there exists a sequence $\{x_n\}$ of M such that x_n converges weakly to x_0. However, by applying Mazur'a theorem (Theorem 2.21) we can find a continuous linear functional f_0 on X such that

$$\begin{aligned} &f_0(x_0) > 1, \\ &|f_0(x)| \leq 1 \quad \text{on } M. \end{aligned}$$

Hence we have the assertion

$$1 < |f_0(x_0)| = \lim_{n\to\infty} |f_0(x_n)| \leq 1.$$

This is a contradiction.

The proof of Theorem 2.29 is complete. □

Finally, the next Eberlein–Shmulyan theorem gives a necessary and sufficient condition for *reflexivity* of a Banach space in terms of *sequential weak compactness* (see [99, Appendix to Chapter V, Section 4, Theorem]):

Theorem 2.30 (Eberlein–Shmulyan)**.** *A Banach space X is* reflexive *if and only if it is locally sequentially* weakly compact*, that is, X is* reflexive *if and only if every strongly bounded sequence of X contains a subsequence which converges* weakly *to an element of X.*

2.4.10 Weak Convergence*

A sequence $\{f_n\}$ in the dual space X' is said to be *weakly* convergent* if a finite $\lim_{n\to\infty} f_n(x)$ exists for every $x \in X$. A sequence $\{f_n\}$ in X' is said to *converge weakly** to an element f of X' if $\lim_{n\to\infty} f_n(x) = f(x)$ for every $x \in X$; we then write $w * -\lim_{n\to\infty} f_n = f$ or simply $f_n \to f$ weakly*. The weak* topology on X' is just the simple topology on the space $X' = L(X, \mathbf{K})$.

We have the following analogue of Theorem 2.27:

Theorem 2.31. *(i)* $s - \lim_{n\to\infty} f_n = f$ *implies* $w * -\lim_{n\to\infty} f_n = f$.

(ii) If X is a Banach space, then a weakly convergent sequence $\{f_n\}$ in X' converges weakly* to an element f of X' and we have the inequality*

$$\|f\| \leq \liminf_{n\to\infty} \|f_n\|.$$

One of the important consequences of Theorem 2.31 is the *sequential weak* compactness* of bounded sets:

Theorem 2.32. *Let X be a separable Banach space. Then every bounded sequence in the strong dual space X' has a subsequence which converges weakly* to an element of X'.*

2.4.11 Dual Operators

The notion of the *transposed matrix* may be extended to the notion of dual operators as follows: Let T be a linear operator from X into Y with domain $D(T)$ dense in X. Such operators are called *densely defined operators*. Each element g of the dual space Y' of Y defines a linear functional G on $D(T)$ by the formula

$$G(x) = g(Tx) \quad \text{for all } x \in D(T).$$

If this functional G is continuous everywhere on $D(T)$ in the strong topology on X', it follows from an application of Theorem 2.8 that G can be extended uniquely to a continuous linear functional g' on the closure

$$\overline{D(T)} = X,$$

that is, there exists a unique element g' of the dual space X' of X which is an extension of G. So we let

$$\begin{aligned} &D(T') \\ &= \text{the totality of those } g \in Y' \text{ such that the mapping} \end{aligned}$$

$$x \longmapsto g(Tx)$$

is continuous everywhere on $D(T)$ in the strong topology on X',

and define

$$T'g = g'.$$

In other words, the mapping T' is a linear operator from Y' into X' with domain $D(T')$ such that

$$g(Tx) = T'g(x) \quad \text{for all } x \in D(T) \text{ and } g \in D(T'). \tag{2.14}$$

The operator T' is called the *dual operator* or *transpose* of T.

Frequently, we write $\langle f, x\rangle$ or $\langle x, f\rangle$ for the value $f(x)$ of a functional f at a point x. For example, we write formula (2.14) as follows:

$$\langle Tx, g\rangle = \langle x, T'g\rangle \quad \text{for all } x \in D(T) \text{ and } g \in D(T'). \tag{2.14$'$}$$

The next theorem states that the continuity of operators is inherited by the transposes ([99, Chapter VII, Section 1, Theorem 2]):

Theorem 2.33. *Let X, Y be normed linear spaces and X', Y' their strong dual spaces, respectively. If T is a bounded linear operator on X into Y, then its transpose T' is a bounded linear operator on Y' into X', and we have the formula*

$$\|T'\| = \|T\|.$$

2.4.12 Adjoint Operators

We assume that a normed linear space X is equipped with a *conjugation*, that is, with a continuous, unitary operation

$$X \ni u \mapsto \overline{u} \in X$$

satisfying the following conditions:

$$\overline{u+v} = \overline{u} + \overline{v},$$
$$\overline{\alpha u} = \overline{\alpha}\,\overline{u},$$
$$\overline{(\overline{u})} = u$$

for all u, $v \in X$ and $\alpha \in \mathbf{K}$. For example, if X is a function space, then $\overline{u}$ is just the usual pointwise complex conjugate:

$$\overline{u}(x) = \overline{u(x)} \quad \text{for all } u \in X.$$

A conjugation on X induces a conjugation on the dual space X' by the formula

$$\langle \overline{f}, u \rangle = \overline{\langle f, \overline{u} \rangle} \quad \text{for all } f \in X' \text{ and } u \in X.$$

Hence we can define a bounded *sesquilinear* form $(\cdot, \cdot)$ on $X' \times X$ by the formula

$$(f, u) = \langle \overline{f}, u \rangle \quad \text{for all } f \in X' \text{ and } u \in X.$$

Now we consider the case where X and Y are each equipped with a conjugation. The notion of the *adjoint matrix* may be extended to the notion of adjoints as follows: Let $A : X \to Y$ be a linear operator with domain $D(A)$ dense in X. Each element v of Y' defines a linear functional V on $D(A)$ by the formula

$$V(u) = (Au, v) \quad \text{for every } u \in D(A).$$

If this functional V is continuous everywhere on $D(A)$ in the strong topology on X', by applying Theorem 2.8 we obtain that V can be extended uniquely to a continuous linear functional v^* on $\overline{D(A)} = X$. So we let

$$\begin{aligned}
&D(A^*)\\
={}&\text{the totality of those } v \in Y' \text{ such that the mapping}\\
&u \longmapsto (Au, v)\\
&\text{is continuous everywhere on } D(A) \text{ in the strong topology on } X',
\end{aligned}$$

and define

$$A^* v = v^*.$$

In other words, the mapping A^* is a linear operator from Y' into X' with domain $D(A^*)$ such that

$$(u, A^* v) = (Au, v) \quad \text{for all } u \in D(A) \text{ and } v \in D(A^*).$$

The operator A' is called the *adjoint operator* or *adjoint* of A.

2.5 Closed Operators

Let X and Y be normed linear spaces over the same scalar field $\mathbf{K}$. Let T be a linear operator from X into Y with domain $D(T)$. The *graph* $G(T)$ of T is the set

$$G(T) = \{\{x, Tx\} : x \in D(T)\}$$

in the product space $X \times Y$. Note that $G(T)$ is a linear subspace of $X \times Y$. We say that T is closed if its graph $G(T)$ is *closed* in $X \times Y$. This is equivalent to saying that

$$\begin{aligned} &\{x_n\} \subset D(T),\ x_n \longrightarrow x \text{ in } X,\ Tx_n \longrightarrow y \text{ in } Y \\ &\Longrightarrow\ x \in D(T),\, Tx = y. \end{aligned}$$

In particular, if T is continuous and its domain $D(T)$ is closed in X, then T is a closed linear operator.

We remark that if T is a closed linear operator which is also injective, then its inverse T^{-1} is a closed linear operator. Indeed, this follows from the fact that the mapping $\{x, y\} \longmapsto \{y, x\}$ is a homeomorphism of $X \times Y$ onto $Y \times X$.

A linear operator T is said to be *closable* if the closure $\overline{G(T)}$ in $X \times Y$ of $G(T)$ is the graph of a linear operator, say, $\overline{T}$, that is, $\overline{G(T)} = G(\overline{T})$.

A linear operator is called a *closed extension* of T if it is a closed linear operator which is also an extension of T. It is easy to see that if T is closable, then every closed extension of T is an extension of $\overline{T}$. Thus the operator $\overline{T}$ is called the *minimal closed extension* of T.

The next theorem gives a necessary and sufficient condition for a linear operator to be closable ([99, Chapter II, Section 6, Proposition 2]):

Theorem 2.34. *A linear operator T from X into Y with domain $D(T)$ is closable if and only if the following condition is satisfied:*

$$\{x_n\} \subset D(T),\ x_n \longrightarrow 0 \text{ in } X,\ Tx_n \longrightarrow y \text{ in } Y \implies y = 0.$$

The notion of adjoints introduced in Subsection 2.4.12 gives a very simple criterion for closability. In fact, we can prove the following ([38, Chapter 3, Section 5, Theorem 5.29]):

Theorem 2.35. *Let X and Y be reflexive Banach spaces. If $T : X \to Y$ is a densely defined, closable linear operator, then the adjoint $T^* : Y' \to X'$ is closed and densely defined. Moreover, $T^{**} = (T^*)^*$ is the minimal closed extension of T. Namely, we have the formula*

$$G(T^{**}) = \overline{G(T)}.$$

Now we can formulate three pillars of functional analysis – Banach's open mapping theorem ([14, Theorem 2.6]), Banach's closed graph theorem ([14, Theorem 2.9]) and Banach's closed range theorem for closed operators ([14, Theorem 2.19]):

Theorem 2.36 (Banach's open mapping theorem)**.** *Let X and Y be*

Banach spaces. Then every continuous linear operator on X onto Y is open, that is, it maps every open set in X onto an open set in Y.

Theorem 2.37 (Banach's closed graph theorem)**.** *Let X and Y be Banach spaces. Then every closed linear operator on X into Y is continuous.*

Corollary 2.38. *Let X and Y be Banach spaces. If T is a continuous, one-to-one linear operator on X onto Y, then its inverse T^{-1} is also continuous; hence T is an isomorphism.*

Indeed, the inverse T^{-1} is a closed linear operator, so that Theorem 2.37 applies.

We give useful characterizations of closed linear operators with closed range ([14, Exercise 2.14]):

Theorem 2.39. *Let X and Y be Banach spaces and T a closed linear operator from X into Y with domain $D(T)$. Then the range $R(T)$ of T is closed in Y if and only if there exists a constant $C > 0$ such that*

$$\operatorname{dist}(x, N(T)) \leq C\,\|Tx\| \quad \text{for all } x \in D(T).$$

Here

$$\operatorname{dist}(x, N(T)) = \inf_{z \in N(T)} \|x - z\|$$

is the distance from x to the null space $N(T)$ of T.

The key point of the next theorem is that if the range $R(T)$ is closed in Y, then a necessary and sufficient condition for the equation $Tx = y$ to be solvable is that the given right-hand side $y \in Y$ is annihilated by every solution $x' \in X'$ of the homogeneous transposed equation $T'x' = 0$ ([99, Chapter VII, Section 5, Theorem]):

Theorem 2.40 (Banach's closed range theorem)**.** *Let X and Y be Banach spaces and T a densely defined, closed linear operator from X into Y. Then the following four conditions are equivalent:*

(i) *The range $R(T)$ of T is closed in Y.*
(ii) *The range $R(T')$ of the transpose T' is closed in X'.*
(iii) $R(T) = {}^0N(T') = \{x \in X : \langle x, x' \rangle = 0 \text{ for all } x' \in N(T')\}$.
(iv) $R(T') = {}^0N(T) = \{x' \in X' : \langle x', x \rangle = 0 \text{ for all } x \in N(T)\}$.

2.6 Complemented Subspaces

Let X be a linear space. Two linear subspaces M and N of X are said to be *algebraic complements* in X if X is the direct sum of M and N, that is, if $X = M \dotplus N$. Algebraic complements M and N in a normed linear space X are said to be *topological complements* in X if the addition mapping

$$\{y, z\} \longmapsto y + z$$

is an isomorphism of $M \times N$ onto X. We then write

$$X = M \oplus N.$$

As an application of Corollary 2.38, we obtain the following:

Theorem 2.41. *Let X be a Banach space. If M and N are closed algebraic complements in X, then they are topological complements.*

A closed linear subspace of a normed linear space X is said to be *complemented* in X if it has a topological complement. By Theorem 2.41, this is equivalent in Banach spaces to the existence of a closed algebraic complement.

The next theorem gives two criteria for a closed subspace to be complemented ([14, Section 2.4]):

Theorem 2.42. *Let X be a Banach space and M a closed subspace of X. If M has either finite dimension or finite codimension, then it is complemented in X.*

2.7 Compact Operators

Let X and Y be normed linear spaces over the same scalar field $\mathbf{K}$. A linear operator T on X into Y is said to be *compact* or *completely continuous* if it maps every bounded subset of X onto a relatively compact subset of Y, that is, if the closure of $T(B)$ is compact in Y for every bounded subset B of X. This is equivalent to saying that, for every bounded sequence $\{x_n\}$ in X, the sequence $\{Tx_n\}$ has a subsequence which converges in Y.

We list some facts which follow at once:

(i) Every compact operator is bounded.
Indeed, a compact operator maps the unit sphere onto a bounded set.

(ii) Every bounded linear operator with finite dimensional range is compact.
This is an immediate consequence of Corollary 2.18.
(iii) No isomorphism between infinite dimensional spaces is compact.
This follows from an application of Theorem 2.19.
(iv) A linear combination of compact operators is compact.
(v) The product of a compact operator with a bounded operator is compact.

The next theorem states that if Y is a Banach space, then the compact operators on X into Y form a closed subspace of $L(X, Y)$ ([14, Theorem 6.1]):

Theorem 2.43. *Let X be a normed linear space and Y a Banach space. If $\{T_n\}$ is a sequence of compact linear operators which converges to an operator T in the space $L(X, Y)$ with the uniform topology, then T is compact.*

As for the transposes of compact operators, we have the following ([14, Theorem 6.4]):

Theorem 2.44 (Schauder)**.** *Let X and Y be normed linear spaces. If T is a compact linear operator on X into Y, then its transpose T' is a compact linear operator on Y' into X'.*

2.8 The Riesz–Schauder Theory

Now we state the most interesting results on compact linear operators, which are essentially due to F. Riesz in the Hilbert space setting. The results are extended to Banach spaces by Schauder:

Theorem 2.45. *Let X be a Banach space and T a compact linear operator on X into itself. Set*

$$S = I - T.$$

Then we have the following three assertions (i), (ii) and (iii):

(i) The null space $N(S)$ of S is finite dimensional and the range $R(S)$ of S is closed in X.
(ii) The null space $N(S')$ of the transpose S' is finite dimensional and the range $R(S')$ of S' is closed in X'.
(iii) $\dim N(S) = \dim N(S')$.

By combining Theorems 2.40 and 2.45, we can obtain an extension of the theory of linear mappings in finite dimensional linear spaces ([14, Theorem 6.6]):

Corollary 2.46 (the Fredholm alternative)**.** *Let T be a compact linear operator on a Banach space X into itself. If $S = I - T$ is either one-to-one or onto, then it is an isomorphism of X onto itself.*

Let T be a bounded linear operator on X into itself. The *resolvent set* of T, denoted $\rho(T)$, is defined to be the set of scalars $\lambda \in \mathbf{K}$ such that $\lambda I - T$ is an isomorphism of X onto itself. In this case, the inverse $(\lambda I - T)^{-1}$ is called the *resolvent* of T. The complement of $\rho(T)$, that is, the set of scalars $\lambda \in \mathbf{K}$ such that $\lambda I - T$ is not an isomorphism of X onto itself is called the *spectrum* of T, and is denoted by $\sigma(T)$.

The set $\sigma_p(T)$ of scalars $\lambda \in \mathbf{K}$ such that $\lambda I - T$ is not one-to-one forms a subset of $\sigma(T)$, and is called the *point spectrum* of T. A scalar $\lambda \in \mathbf{K}$ belongs to $\sigma_p(T)$ if and only if there exists a non-zero element $x \in X$ such that $Tx = \lambda x$. In this case, λ is called an *eigenvalue* of T and x an *eigenvector* of T corresponding to λ. Also the null space $N(\lambda I - T)$ of $\lambda I - T$ is called the *eigenspace* of T corresponding to λ, and the dimension of $N(\lambda I - T)$ is called the *multiplicity* of λ.

By using C. Neumann's series (Theorem 2.13), we find that the resolvent set $\rho(T)$ is open in $\mathbf{K}$ and that

$$\{\lambda \in \mathbf{K} : |\lambda| > \|T\|\} \subset \rho(T).$$

Hence the spectrum $\sigma(T) = \mathbf{K} \setminus \rho(T)$ is closed and bounded in $\mathbf{K}$.

If T is a compact operator and λ is a non-zero element of $\sigma(T)$, then, by applying Corollary 2.46 to the operator $\lambda^{-1}T$ we obtain that $\lambda I - T$ is not one-to-one, that is, $\lambda \in \sigma_p(T)$. Also note that if X is infinite dimensional, then T is not an isomorphism of X onto itself; hence $0 \in \sigma_p(T)$. Therefore the scalar field $\mathbf{K}$ can be decomposed as follows:

$$\mathbf{K} = (\sigma_p(T) \cup \{0\}) \cup \rho(T).$$

We can say rather more about the spectrum $\sigma(T)$ in terms of transpose operators ([99, Chapter X, Section 5, Theorems 1, 2 and 3]):

Theorem 2.47 (Riesz–Schauder)**.** *Let T be a compact linear operator on a Banach space X into itself. Then we have the following three assertions (i), (ii) and (iii):*

(i) The spectrum $\sigma(T)$ of T is either a finite set or a countable set

accumulating only at the zero 0*; and every non-zero element of* $\sigma(T)$ *is an eigenvalue of* T.

(ii) $\dim N(\lambda I - T) = \dim N(\overline{\lambda} - T') < +\infty$ *for all* $\lambda \neq 0$.

(iii) Let $\lambda \neq 0$. *The non-homogeneous equation*

$$(\lambda I - T)x = y$$

has a solution if and only if y *is orthogonal to the space* $N(\overline{\lambda} - T')$. *Similarly, the non-homogeneous transpose equation*

$$(\overline{\lambda} I - T')z = w$$

has a solution if and only if w *is orthogonal to the space* $N(\lambda I - T)$. *Moreover, the operator* $\lambda I - T$ *is onto if and only if it is one-to-one.*

2.9 Fredholm Operators

Throughout this section, the letters X, Y, Z denote Banach spaces over the same scalar field $\mathbf{K}$.

A linear operator $T : X \to Y$ is called a *Fredholm operator* if the following five conditions are satisfied:

(i) The domain $D(T)$ of T is dense in X.
(ii) T is a closed operator.
(iii) The null space $N(T) = \{x \in D(T) : Tx = 0\}$ of T has finite dimension; $\dim N(T) < \infty$.
(iv) The range $R(T)$ of T is closed in Y.
(v) The range $R(T)$ has finite codimension in Y; $\operatorname{codim} R(T) = \dim Y/R(T) < \infty$.

Then the *index* of T is defined by the formula

$$\operatorname{ind} T := \dim N(T) - \operatorname{codim} R(T).$$

For example, we find from Theorems 2.45 and 2.40 that if $X = Y$ and T is compact, then the operator $I - T$ is a Fredholm operator and $\operatorname{ind}(I - T) = 0$.

We give a characterization of Fredholm operators. First, we have the following ([48, Theorem 2.24]):

Theorem 2.48. *If* $T : X \to Y$ *is a Fredholm operator with domain* $D(T)$*, then there exist a bounded linear operator* $S : Y \to X$ *and compact linear operators* $P : X \to X$, $Q : Y \to Y$ *such that*

(a) $ST = I - P$ *on* $D(T)$,
(b) $TS = I - Q$ *on* Y.

Furthermore, we have the formulas

$$R(P) = N(T),$$
$$\dim R(Q) = \operatorname{codim} R(T).$$

Theorem 2.48 has a converse:

Theorem 2.49. *Let* T *be a closed linear operator from* X *into* Y *with domain* $D(T)$ *dense in* X. *Assume that there exist bounded linear operators* $S_1 : Y \to X$ *and* $S_2 : Y \to X$ *and compact linear operators* $K_1 : X \to X$, $K_2 : Y \to Y$ *such that*

(a) $S_1 T = I - K_1$ *on* $D(T)$,
(b) $T S_2 = I - K_2$ *on* Y.

Then T *is a Fredholm operator.*

Now we state some important properties of Fredholm operators ([48, Theorem 2.21]):

Theorem 2.50. *If* $T : X \to Y$ *is a Fredholm operator and if* $S : Y \to Z$ *is a Fredholm operator, then the product* $ST : X \to Z$ *is a Fredholm operator, and we have the formula*

$$\operatorname{ind}(ST) = \operatorname{ind} S + \operatorname{ind} T.$$

The next theorem states that the index is *stable* under compact perturbations or small perturbations ([48, Theorem 2.26]):

Theorem 2.51. *(i) If* $T : X \to Y$ *is a Fredholm operator and if* $K : X \to Y$ *is a compact linear operator, then the sum* $T + K : X \to Y$ *is a Fredholm operator, and we have the formula*

$$\operatorname{ind}(T + K) = \operatorname{ind} T.$$

(ii) The Fredholm operators form an open subset *of the space* $L(X, Y)$ *of bounded operators. More precisely, if* $E : X \to Y$ *is a bounded operator with* $\|E\|$ *sufficiently small, then the sum* $T + E : X \to Y$ *is a Fredholm operator, and we have the formula*

$$\operatorname{ind}(T + E) = \operatorname{ind} T.$$

As for the transposes of Fredholm operators, we have the following:

Theorem 2.52. *If $T : X \to Y$ is a Fredholm operator and if Y is reflexive, then the transpose $T' : Y' \to X'$ of T is a Fredholm operator, and we have the formula*

$$\operatorname{ind} T' = -\operatorname{ind} T.$$

Now we can state a generalization of the *Fredholm alternative* (Corollary 2.46) in terms of adjoint operators ([48, Theorem 2.27]):

Theorem 2.53 (the Fredholm alternative)**.** *Let $A : X \to Y$ be a Fredholm operator with* ind $A = 0$*. Then there are two, mutually exclusive possibilities (i) and (ii):*

(i) The homogeneous equation $Au = 0$ has only the trivial solution $u = 0$. In this case, we have the following two assertions:

- *(a) For each $f \in Y$, the non-homogeneous equation $Au = f$ has a unique solution $u \in X$.*
- *(b) For each $g \in X'$, the adjoint equation $A^* v = g$ has a unique solution $v \in Y'$.*

(ii) The homogeneous equation $Au = 0$ has exactly p linearly independent solutions $u_1, u_2, \ldots, u_p$ for some $p \geq 1$. In this case, we have the following three assertions:

- *(c) The homogeneous adjoint equation $A^* v = 0$ has exactly p linearly independent solutions $v_1, v_2, \ldots, v_p$.*
- *(d) The non-homogeneous equation $Au = f$ is solvable if and only if the right-hand side f satisfies the orthogonal conditions*
$$(v_j, f) = 0 \quad \text{for all } 1 \leq j \leq p.$$
- *(e) The non-homogeneous adjoint equation $A^* v = g$ is solvable if and only if the right-hand side g satisfies the orthogonal conditions*
$$(g, u_j) = 0 \quad \text{for all } 1 \leq j \leq p.$$

2.10 Hilbert Spaces

A complex (or real) linear space X is called a *pre-Hilbert space* or *inner product space* if, to each ordered pair of elements x and y of X, there is associated a complex (or real) number (x, y) in such a way that

(I1) $(y, x) = \overline{(x, y)}$.

(I2) $(\alpha x, y) = \alpha(x, y)$ for all $\alpha \in \mathbf{C}$ (or $\alpha \in \mathbf{R}$).
(I3) $(x + y, z) = (x, z) + (y, z)$ for all x, y and $z \in X$.
(I4) $(x, x) \geq 0$; $(x, x) = 0$ if and only if $x = 0$.

Here $\overline{(x, y)}$ denotes the complex conjugate of (x, y). In the real case condition (I1) becomes simply $(y, x) = (x, y)$. The number (x, y) is called the *inner product* or *scalar product* of x and y.

The following are immediate consequences of conditions (I1), (I2) and (I3):

(i) $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$ for all α, $\beta \in \mathbf{C}$.
(ii) $(x, \alpha y + \beta z) = \overline{\alpha}(x, y) + \overline{\beta}(x, z)$ for all α, $\beta \in \mathbf{C}$.

These properties (i) and (ii) are frequently called *sesquilinearity*. In the real case they are bilinearity.

We list some basic properties of the inner product:

(1) The *Schwarz inequality* holds true for all x, $y \in X$:

$$|(x, y)|^2 \leq (x, x)(y, y).$$

Here the equality holds true if and only if x and y are linearly dependent.

(2) The quantity

$$\|x\| = \sqrt{(x, x)} \quad \text{(the non-negative square root)}$$

satisfies axioms (N1), (N2) and (N3) of a norm; hence a pre-Hilbert space is a normed linear space by the norm

$$\|x\| = \sqrt{(x, x)}.$$

(3) The inner product (x, y) is a continuous function of x and y:

$$\|x_n - x\| \longrightarrow 0, \ \|y_n - y\| \longrightarrow 0 \implies (x_n, y_n) \longrightarrow (x, y).$$

(4) The *parallelogram law* holds true for all x, $y \in X$:

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2). \tag{2.15}$$

Conversely, we assume that X is a normed linear space whose norm satisfies condition (2.15). We let

$$(x, y) = \frac{1}{4}\left(\|x + y\|^2 - \|x - y\|^2\right)$$

if X is a real normed linear space, and let

$$(x, y) = \frac{1}{4}\left(\|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2\right)$$

if X is a complex normed linear space. Then it is easy to verify that the number (x, y) satisfies axioms (I1) through (I4) of an inner product; hence X is a pre-Hilbert space.

A pre-Hilbert space is called a *Hilbert space* if it is complete with respect to the norm derived from the inner product.

If X and Y are pre-Hilbert spaces over the same scalar field $\mathbf{K}$, then the product space $X \times Y$ is a pre-Hilbert space by the inner product

$$\begin{aligned}&(\{x_1, y_1\}, \{x_2, y_2\}) \\ &= (x_1, x_2) + (y_1, y_2) \quad \text{for all } x_1,\ x_2 \in X \text{ and } y_1,\ y_2 \in Y.\end{aligned}$$

Furthermore, if X and Y are Hilbert spaces, then so is $X \times Y$.

2.10.1 Orthogonality

Let X be a pre-Hilbert space. Two elements x, y of X are said to be *orthogonal* if $(x, y) = 0$; we then write $x \perp y$. We remark that

$$\begin{aligned} x \perp y &\iff y \perp x. \\ x \perp x &\iff x = 0. \end{aligned}$$

If A is a subset of X, we let

$$A^{\perp} = \{x \in X : (x, y) = 0 \text{ for all } y \in A\}.$$

In other words, $A^{\perp}$ is the set of all those elements of X which are orthogonal to every element of A.

We list some facts which follow at once:

(i) The set $A^{\perp}$ is a linear subspace of X.
(ii) $A \subset B \implies B^{\perp} \subset A^{\perp}$.
(iii) $A \cap A^{\perp} = \{0\}$.
(iv) The set $A^{\perp}$ is closed.
(v) $A^{\perp} = \overline{A}^{\perp} = \overline{[A]}^{\perp}$ where $\overline{A}$ is the closure of A and $[A]$ is the space spanned by A, that is, the space of all finite linear combinations of elements of A.

Facts (iv) and (v) follow from the continuity of the inner product.

2.10.2 The Closest-Point Theorem and Applications

Theorem 2.54 (the closest-point theorem)**.** *Let A be a closed convex subset of a Hilbert space X. If x is a point not in A, then there is a unique point a in A such that*

$$\|x - a\| = \operatorname{dist}(x, A).$$

Theorem 2.54 can be proved by using the parallelogram law.

One of the consequences of Theorem 2.54 is that every closed linear subspace of a Hilbert space is complemented:

Theorem 2.55. *Let M be a closed linear subspace of a Hilbert space X. Then every element x of X can be decomposed uniquely in the form*

$$x = y + z, \quad y \in M, \ z \in M^{\perp}. \tag{2.16}$$

Moreover, the mapping $x \mapsto \{y, z\}$ is an isomorphism of X onto $M \times M^{\perp}$.

We shall write the decomposition (2.16) in the form

$$X = M \oplus M^{\perp}, \tag{2.16$'$}$$

emphasizing that the mapping $x \mapsto \{y, z\}$ is an isomorphism of X onto $M \times M^{\perp}$. The space $M^{\perp}$ is called the *orthogonal complement* of M.

Corollary 2.56. *If M is a closed linear subspace of a Hilbert space X, then it follows that $M^{\perp\perp} = (M^{\perp})^{\perp} = M$. Furthermore, if A is a subset of X, then we have $A^{\perp\perp} = \overline{[A]}$, where $[A]$ is the space spanned by A.*

With the above notation (2.16), we define a mapping P_M of X into M by the formula

$$P_M x = y.$$

Since the decomposition (2.16) is unique, it follows that P_M is linear.

Furthermore, we easily obtain the following:

Theorem 2.57. *The operator P_M enjoys the following three properties (i), (ii) and (iii):*

(i) $P_M^2 = P_M$ *(idempotent property).*
(ii) $(P_M x, x') = (x, P_M x')$ *(symmetric property).*
(iii) $\|P_M\| \leq 1$.

The operator P_M is called the *orthogonal projection* onto M.

Similarly, we define a mapping $P_{M^\perp}$ of X into $M^\perp$ by the formula

$$P_{M^\perp} x = z.$$

Then Corollary 2.56 asserts that $P_{M^\perp}$ is the orthogonal projection onto $M^\perp$. It is clear that

$$P_M + P_{M^\perp} = I.$$

Now we give an important characterization of bounded linear functionals on a Hilbert space:

Theorem 2.58 (the Riesz representation theorem). *Every element y of a Hilbert space X defines a bounded linear functional $J_X y$ on X by the formula*

$$J_X y(x) = (x, y) \quad \text{for all } x \in X, \tag{2.17}$$

and we have the formula

$$\|J_X y\| = \sup_{\substack{x \in X \\ \|x\| \le 1}} |J_X y(x)| = \|y\|.$$

Conversely, for every bounded linear functional f on X, there exists a unique element y of X such that $f = J_X y$, that is,

$$f(x) = (x, y) \quad \text{for all } x \in X,$$

and so

$$\|f\| = \|y\|.$$

In view of formula (2.17), it follows that the mapping J_X enjoys the following property:

$$J_X(\alpha y + \beta z) = \overline{\alpha} J_X y + \overline{\beta} J_X z \quad \text{for all } y,\ z \in X \text{ and } \alpha,\ \beta \in \mathbf{C}.$$

We express this by saying that J_X is *conjugate linear* or *antilinear.* In the real case, J_X is linear.

Let X' be the strong dual space of a Hilbert space X, that is, the space of bounded linear functionals on X with the norm

$$\|f\| = \sup_{\substack{x \in X \\ \|x\| \le 1}} |f(x)|.$$

Then Theorem 2.58 may be restated as follows:

There is a conjugate linear, norm-preserving isomorphism (2.18)

J_X of X onto X'.

In this case, we say that X' is *antidual* to X.

Recall that a sequence $\{x_n\}$ in a normed linear space X is said to converge weakly to an element x of X if $f(x_n) \to f(x)$ for every $f \in X'$. Assertion (2.18) tells us that a sequence $\{x_n\}$ in a Hilbert space X converges weakly to an element x of X if and only if $(x_n, y) \to (x, y)$ for every $y \in X$.

Another important consequence of Theorem 2.58 is the *reflexivity* of Hilbert spaces:

Corollary 2.59. *Every Hilbert space can be identified with its strong bidual space.*

2.10.3 Orthonormal Sets

Let X be a pre-Hilbert space. A subset S of X is said to be *orthogonal* if every pair of distinct elements of S is orthogonal. Furthermore, if each element of S has norm one, then S is said to be *orthonormal*. We remark that if S is an orthogonal set of non-zero elements, we can construct an orthonormal set from S by normalizing each element of S. If $\{x_1, x_2, \dots, x_n\}$ is an orthonormal set and if $x = \sum_{i=1}^{n} \alpha_i x_i$, then we have the formula

$$\|x\|^2 = \sum_{i=1}^{n} |\alpha_i|^2, \quad \alpha_i = (x, x_i).$$

Therefore, every orthonormal set is linearly independent.

First, we state the Gram–Schmidt orthogonalization:

Theorem 2.60 (Gram–Schmidt)**.** *Let $\{x_i\}_{i \in I}$ be a finite or countable infinite set of linearly independent vectors of X. Then we can construct an orthonormal set $\{u_i\}_{i \in I}$ such that, for each $i \in I$,*

(a) u_i is a linear combination of $\{x_1, x_2, \dots, x_i\}$.
(b) x_i is a linear combination of $\{u_1, u_2, \dots, u_i\}$.

Corollary 2.61. *Every n-dimensional pre-Hilbert space over the scalar field $\mathbf{K}$ is isomorphic to the space $\mathbf{K}^n$ with the usual inner product.*

Let $\{u_\lambda\}_{\lambda \in \Lambda}$ be an orthonormal set of a pre-Hilbert space X. For each $x \in X$, we let

$$\hat{x}_\lambda = (x, u_\lambda), \quad \lambda \in \Lambda.$$

The scalars $\hat{x}_\lambda$ are called the *Fourier coefficients* of x with respect to $\{u_\lambda\}$.

Then we have the following:

Theorem 2.62. *For each $x \in X$, the set of those $\lambda \in \Lambda$ such that $\hat{x}_\lambda \neq 0$ is at most countable. Furthermore, we have the* Bessel inequality

$$\sum_{\lambda\in\Lambda} |\hat{x}_\lambda|^2 \leq \|x\|^2.$$

An orthonormal set S of X is called a *complete orthonormal system* if it is not contained in a larger orthonormal set of X.

As for the existence of such systems, we have the following:

Theorem 2.63. *Let X be a Hilbert space having a non-zero element. Then, for every orthonormal set S in X, there exists a complete orthonormal system that contains S.*

The next theorem gives useful criteria for the completeness of orthonormal sets:

Theorem 2.64. *Let $S = \{u_\lambda\}_{\lambda\in\Lambda}$ be an orthonormal set in a Hilbert space X. Then the following five conditions (i) through (v) are equivalent:*

(i) The set S is complete.
(ii) $S^\perp = \{0\}$.
(iii) The space $[S]$ spanned by S is dense in X: $\overline{[S]} = X$.
(iv) For every $x \in X$, we have the formula

$$\|x\|^2 = \sum_{\lambda\in\Lambda} |\hat{x}_\lambda|^2. \tag{2.19}$$

(v) For every $x \in X$, we have the formula

$$x = \sum_{\lambda\in\Lambda} \hat{x}_\lambda u_\lambda \quad \text{in } X. \tag{2.20}$$

Formula (2.19) is called the *Parseval identity* and formula (2.20) is called the *Fourier series expansion* of x with respect to $\{u_\lambda\}$.

2.10.4 Adjoint Operators

Throughout this subsection, the letters X, Y, Z denote Hilbert spaces over the same scalar field $\mathbf{K}$. Let T be a linear operator from X into

Y with domain $D(T)$ dense in X. Each element y of Y defines a linear functional f on $D(T)$ by the formula

$$f(x) = (Tx, y) \quad \text{for every } x \in D(T).$$

If this functional f is continuous everywhere on $D(T)$, by applying Theorem 2.8 we obtain that f can be extended uniquely to a continuous linear functional $\widetilde{f}$ on $\overline{D(T)} = X$. Therefore, Riesz's theorem (Theorem 2.63) asserts that there exists a unique element y^* of X such that

$$\widetilde{f}(x) = (x, y^*) \quad \text{for all } x \in X.$$

In particular, we have the formula

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

Hence we let

$$\begin{aligned} D(T^*) = {} & \text{the totality of those } y \in Y \text{ such that the mapping} \\ & x \longmapsto (Tx, y) \\ & \text{is continuous everywhere on } D(T), \end{aligned}$$

and define

$$T^* y = y^*.$$

In other words, the mapping T^* is a linear operator from Y into X with domain $D(T^*)$ such that

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

The operator T^* is called the *adjoint operator* or simply the *adjoint* of T.

We list some basic properties of adjoints:

(i) The operator T^* is closed.
(ii) If $T \in L(X, Y)$, then $T^* \in L(Y, X)$ and $\|T^*\| = \|T\|$.
(iii) If T, $S \in L(X, Y)$, then $(\alpha T + \beta S)^* = \overline{\alpha} T^* + \overline{\beta} S^*$ for all α, $\beta \in \mathbf{C}$.
(iv) If $T \in L(X, Y)$ and $S \in L(Y, Z)$, then $(ST)^* = T^* S^*$.

A densely defined linear operator T from X into itself is said to be *self-adjoint* if $T = T^*$. Note that every self-adjoint operator is closed.

As for the adjoints of closed operators, we have the following (see Theorem 2.35):

Theorem 2.65. *If T is a densely defined, closed linear operator from X into Y, then the adjoint T^* is a densely defined, closed linear operator from Y into X and we have the formula*

$$T^{**} = (T^*)^* = T.$$

Corollary 2.66. *If T is a densely defined, closable linear operator, then the adjoint T^* is densely defined and the operator T^{**} coincides with the minimal closed extension $\overline{T}$ of T. Namely, we have the formula*

$$G\left(T^{**}\right) = \overline{G(T)}.$$

2.11 Continuous Functions on Metric Spaces

In this last section we formulate two fundamental theorems concerning spaces of continuous functions defined on a metric space.

2.11.1 The Ascoli–Arzelà Theorem

Let X be a subset of a metric space (S, ρ) and F a Banach space. We let

$$C(X, F) = \text{ the space of continuous maps of } X \text{ into } F.$$

We say that a subset Φ of $C(X, F)$ is *equicontinuous* at a point x_0 of X if, for any given ε there exists a constant $\delta = \delta(x_0, \varepsilon) > 0$ such that we have, for all $f \in \Phi$,

$$x \in X, \ \rho(x, x_0) < \delta \implies |f(x) - f(x_0)| < \varepsilon.$$

We say that Φ is equicontinuous on X if it is equicontinuous at every point of X.

The next theorem provides a criterion for compactness of compact subsets of functions spaces:

Theorem 2.67 (Ascoli–Arzelà)**.** *Let X be a compact subset of a metric space (S, ρ) and F a Banach space. Let Φ be a subset of the space $C(X, F)$ of continuous maps with supremum norm. Then Φ is* relatively compact *in $C(X, F)$ if and only if the following two conditions are satisfied:*

(I) Φ *is equicontinuous.*
(II) *For each $x \in X$, the set $\Phi(x) = \{f(x) : f \in \Phi\}$ is relatively compact in F.*

The next corollary is a version of the Ascoli–Arzelà theorem in the case where F is the real numbers $\mathbf{R}$ or complex numbers $\mathbf{C}$.

Corollary 2.68. *Let X be a compact subset of a metric space, and let Φ be a subset of the space $C(X)$ of continuous functions on X with supremum norm. Then Φ is relatively compact in $C(X)$ if and only if it is equicontinuous and bounded for the supremum norm.*

We remark that a subset Φ is *relatively compact* if and only if it has the property that any sequence in Φ has a convergent subsequence, converging in its closure. The Ascoli–Arzelà theorem is used mostly when X is a σ-compact space. In that case, we have the following version of the Ascoli–Arzelà theorem:

Corollary 2.69 (Ascoli–Arzelà)**.** *Let X be a metric space that is a denumerable union of compact sets. Let $\{f_n\}$ be a sequence of continuous maps of X into a Banach space F. Assume that:*

(1) The sequence $\{f_n\}$ is equicontinuous as a family of maps.
(2) For each $x \in X$, the closure of the set $\{f_n(x) : n = 1, 2, \ldots\}$ is compact in F.

Then there exists a subsequence of $\{f_n\}$ that converges pointwise to a continuous map $f \in C(X, F)$, and this convergence is uniform on every compact subset of X.

2.11.2 The Stone–Weierstrass Theorem

Let (X, d) be a compact metric space. In the space $C(X)$ of continuous, real-valued functions on X, we define the *uniform metric*

$$\rho\,(f, g) = \max_{x \in X} |f(x) - g(x)| \quad \text{for } f,\, g \in C(X).$$

Moreover, we introduce the following operations on $C(X)$ for any f, $g \in C(X)$ and $\alpha \in \mathbf{R}$:

(S1) $(f + g)(x) = f(x) + g(x)$ for all $x \in X$.
(S2) $(fg)(x) = f(x) \cdot g(x)$ for all $x \in X$.
(S3) $(\alpha\, f)(x) = \alpha\, f(x)$ for all $x \in X$.

A subset $\mathcal{A}$ of $C(X)$ is called an *algebra* if it is a real linear subspace such that

$$f, g \in \mathcal{A} \implies fg \in \mathcal{A}.$$

Let $\mathcal{B}$ be a subset of $C(X)$. The intersection of all algebras containing $\mathcal{B}$ is an algbra. This algebra is called the *algebra generated by* $\mathcal{B}$.

Example 2.2. Let $X = [a, b]$ be the finite closed interval in $\mathbf{R}$. Let $C[a, b]$ be the space of continuous, real-valued functions on $[a, b]$. The algebra generated by functions 1 and x is the set P of all polynomials.

A subset $\mathcal{B}$ of $C(X)$ is said to *separate points* of X if, for any two distinct points x, $y \in X$ there exists a function $f \in \mathcal{B}$ such that

$$f(x) \neq f(y).$$

For example, the set P of all polynomials separates points of $X = [a, b]$.

The next theorem is a generalization of the classical *Weierstrass approximation theorem* (see [30, Chapter 3, Theorem 3.7.1], [29, Section 4.7, Theorem 4.45]):

Theorem 2.70 (Stone–Weierstrass)**.** *Let X be a compact metric space and let $\mathcal{A}$ be an algebra in the space $C(X)$ of continuous, real-valued functions on X. If $\mathcal{A}$ contains the constant function* 1 *and separates points of X, then $\mathcal{A}$ is dense in $C(X)$, that is, $\overline{\mathcal{A}} = C(X)$.*

The classical Weierstrass theorem is a special case of the Stone–Weierstrass theorem (Theorem 2.70) if we take

$$\begin{aligned} X &= C[a, b], \\ \mathcal{A} &= P. \end{aligned}$$

2.12 Notes and Comments

The topics in this chapter form a necessary background for what follows. For more thorough treatments of this subject, the reader might be referred to Brezis [14], Folland [29], Friedman [30], Kato [38], Kolmogorov–Fomin [39], Reed–Simon [58], Rudin [60], Schechter [66] and Yosida [99].

3

Measures, Integration and L^p Spaces

In this chapter we set forth the basic concepts of measure theory and develop the theory of integration on abstract measure spaces, paying particular attention to the Lebesgue integral on the Euclidean space $\mathbf{R}^n$. In particular, we prove Minkowski's inequality for integrals (Theorem3.16) and Hardy's inequality (Theorem 3.18)in L^p spaces. In Section 3.9, we prove the Marcinkiewicz interpolation theorem (Theorem 3.30) which plays an important role in the proof of Theorem 9.5 in Section 9.3. In Section 3.10, as an application of Marcinkiewicz's interpolation theorem we study Riesz potentials in the classical potential theory (Theorem 3.31).

3.1 Measure Theory

This section is a summary of the basic definitions and results about measure theory which will be used throughout the book. Most of the material are quite familiar to the reader and may be omitted. This section, included for the sake of completeness, should serve to settle questions of notation and such.

3.1.1 Measurable Spaces and Functions

Let X be a non-empty set. A collection $\mathcal{M}$ of subsets of X is said to be a *σ-algebra* in X if it has the following three properties (S1), (S2) and (S3):

(S1) The empty set $\emptyset$ belongs to $\mathcal{M}$.
(S2) If $A \in \mathcal{M}$, then its complement $A^c = X \setminus A$ belongs to $\mathcal{M}$.

(S3) If $\{A_n\}$ is an arbitrary countable collection of members of $\mathcal{M}$, then the union $\bigcup_{n=1}^{\infty} A_n$ belongs to $\mathcal{M}$.

The pair $(X, \mathcal{M})$ is called a *measurable space* and the members of $\mathcal{M}$ are called *measurable sets* in X.

For any collection $\mathcal{F}$ of subsets of X, there exists a smallest σ-algebra $\sigma(\mathcal{F})$ in X which contains $\mathcal{F}$. This $\sigma(\mathcal{F})$ is sometimes called the *σ-algebra generated by $\mathcal{F}$*.

We let

$$\overline{\mathbf{R}} = \{-\infty\} \cup \mathbf{R} \cup \{+\infty\}$$

with the obvious ordering. The topology on $\overline{\mathbf{R}}$ is defined by declaring that the open sets in $\overline{\mathbf{R}}$ are those which are unions of segments of the types

$$(a, b), \ [-\infty, a), \ (a, +\infty].$$

The elements of $\overline{\mathbf{R}}$ are called *extended real numbers.*

Let $(X, \mathcal{M})$ be a measurable space. An extended real-valued function f defined on a set $A \in \mathcal{M}$ is said to be *$\mathcal{M}$-measurable* or simply *measurable* if, for every $a \in \mathbf{R}$, the set

$$\{x \in A : f(x) > a\}$$

is in $\mathcal{M}$.

If A is a subset of X, we let

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

The function χ_A is called the *characteristic function* of A.

A real-valued function f on X is called a *simple function* if it takes on only a finite number of values. Thus, if a_1, a_2, ..., a_m are the distinct values of f, then f can be written as

$$f(x) = \sum_{j=1}^{m} a_j \chi_{A_j}(x)$$

where $A_j = \{x \in X : f(x) = a_j\}$. We remark that the function f is measurable if and only if each A_j is measurable.

The next theorem characterizes measurable functions in terms of simple functions:

Theorem 3.1. *An extended real-valued function defined on a measurable set is measurable if and only if it is a pointwise limit of a sequence* *of measurable simple functions. Furthermore, every non-negative measurable function is a pointwise limit of an increasing sequence of non-negative measurable simple functions.*

The next *monotone class theorem* will be useful for the study of measurability of functions:

Theorem 3.2 (the monotone class theorem)**.** *Let $\mathcal{F}$ be a π-system in X and let $\mathcal{H}$ be a linear space of real-valued functions on X. Assume that the following two conditions (i) and (ii) are satisfied:*

(i) $1 \in \mathcal{H}$ *and* $\chi_A \in \mathcal{H}$ *for all* $A \in \mathcal{F}$.
(ii) *If* $\{f_n\}$ *is an increasing sequence of non-negative functions in* $\mathcal{H}$ *such that* $f = \sup_n f_n$ *is bounded, then it follows that* $f \in \mathcal{H}$.

Then the linear space $\mathcal{H}$ contains all real-valued, bounded functions on X which are $\sigma(\mathcal{F})$-measurable.

3.1.2 Measures

Let $(X, \mathcal{M})$ be a measurable space. An extended real-valued function μ defined on $\mathcal{M}$ is called a *non-negative measure* or simply a *measure* if it has the following three properties (M1), (M2) and (M3):

(M1) $0 \leq \mu(A) \leq \infty$, $A \in \mathcal{M}$.
(M2) $\mu(\emptyset) = 0$.
(M3) The function μ is countably additive, that is,

$$\mu\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

for any disjoint countable collection $\{A_i\}$ of members of $\mathcal{M}$.

The triple $(X, \mathcal{M}, \mu)$ is called a *measure space*. In other words, a measure space is a measurable space which has a non-negative measure defined on the σ-algebra of its measurable sets. If $\mu(X) < \infty$, then the measure μ is called a *finite measure* and the space $(X, \mathcal{M}, \mu)$ is called a *finite measure space*. If X is a countable union of sets of finite measure, then the measure μ is said to be σ*-finite* on X. We also say that the measure space $(X, \mathcal{M}, \mu)$ is σ*-finite*.

Let $(X, \mathcal{M}, \mu)$ be a measure space. A set $E \in \mathcal{M}$ such that $\mu(E) = 0$ is called a *null set*. We remark that any countable union of null sets is a null set. A measure μ whose domain includes all subsets of null sets is said to be *complete*, that is, if it satisfies the condition

$$\mu(E) = 0 \text{ and } F \subset E \implies F \in \mathcal{M}.$$

It is known (see [29, Theorem 1.9]) that, for a measure space $(X, \mathcal{M}, \mu)$ there exist a σ-algebra $\overline{\mathcal{M}}$ and a unique extension $\overline{\mu}$ of μ to a complete measure on $\overline{\mathcal{M}}$. The measure $\overline{\mu}$ is called the *completion* of μ and the σ-algebra $\overline{\mathcal{M}}$ is called the *completion* of $\mathcal{M}$ with respect to μ.

3.1.3 Lebesgue Measures

The next theorem is one of the fundamental theorems in measure theory:

Theorem 3.3. *There exist a σ-algebra $\mathcal{M}$ in $\mathbf{R}^n$ and a non-negative measure μ on $\mathcal{M}$ having the following four properties (i), (ii), (iii) and (iv):*

(i) *Every open set in $\mathbf{R}^n$ is in $\mathcal{M}$.*
(ii) *If $A \subset B$, $B \in \mathcal{M}$ and $\mu(B) = 0$, then $A \in \mathcal{M}$ and $\mu(A) = 0$.*
(iii) *If $A = \{x \in \mathbf{R}^n : a_j \le x_j \le b_j (1 \le j \le n)\}$, then $A \in \mathcal{M}$ and $\mu(A) = \prod_{j=1}^n (b_j - a_j)$.*
(iv) *The measure μ is* translation invariant, *that is, if $x \in \mathbf{R}^n$ and $A \in \mathcal{M}$, then the set $x + A = \{x + y : y \in A\}$ is in $\mathcal{M}$ and $\mu(x + A) = \mu(A)$.*

The elements of $\mathcal{M}$ are called *Lebesgue measurable* sets in $\mathbf{R}^n$ and the measure μ is called the *Lebesgue measure* on $\mathbf{R}^n$.

3.1.4 Signed Measures

Let $(X, \mathcal{M})$ be a measurable space. A real-valued function μ defined on $\mathcal{M}$ is called a *signed measure* or *real measure* if it is countably additive, that is,

$$\mu\left(\sum_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

for any disjoint countable collection $\{A_i\}$ of members of $\mathcal{M}$. It should be noticed that every rearrangement of the series $\sum_i \mu(A_i)$ also converges, since the disjoint union $\sum_i A_i$ is not changed if the subscripts are permuted. A signed measure takes its values in $(-\infty, +\infty)$, but a non-negative measure may take $+\infty$; hence the non-negative measures do not form a subclass of the signed measures.

If μ is a signed measure, we define a function $|\mu|$ on $\mathcal{M}$ by the formula

$$|\mu|(A) = \sup \sum_{i=1}^{n} |\mu(A_i)|, \quad A \in \mathcal{M}, \tag{3.1}$$

where the supremum is taken over all countable partitions $\{A_i\}$ of A into members of $\mathcal{M}$. Then the function $|\mu|$ is a finite non-negative measure on $\mathcal{M}$. The measure $|\mu|$ is called the *total variation measure* of μ, and the quantity $|\mu|(X)$ is called the *total variation* of μ. We remark that

$$|\mu(A)| \leq |\mu|(A) \leq |\mu|(X), \quad A \in \mathcal{M}.$$

3.1.5 Borel Measures

Let X be a locally compact Hausdorff space. There exists a smallest σ-algebra $\mathcal{B}$ in X which contains all open sets in X. The members of $\mathcal{B}$ are called *Borel sets* in X. A signed measure defined on $\mathcal{B}$ is called a *real Borel measure* on X. A non-negative Borel measure μ is said to be *outer regular* on a Borel set B if it satisfies the condition

$$\mu(B) = \inf \{\mu(G) : B \subset G, \ G \text{ is open}\}.$$

A non-negative Borel measure μ is said to be *inner regular* on a Borel set B if it satisfies the condition

$$\mu(B) = \sup \{\mu(F) : F \subset B, \ F \text{ is compact}\}.$$

If μ is outer and inner regular on all Borel sets, it is called a *regular Borel measure.*

We give a useful criterion for the regularity of μ:

Theorem 3.4. *Let X be a locally compact Hausdorff space in which every open set is σ-compact. If μ is a non-negative Borel measure on X such that $\mu(K) < +\infty$ for every compact set $K \subset X$, then it is regular.*

Let (X, ρ) be a locally compact metric space. A *Radon measure* on X is a Borel measure that is finite on all compact sets in X, and is outer regular on all Borel sets in X and inner regular on all open sets in X.

3.1.6 Product Measures

Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be measurable spaces. We let

$$\mathcal{M} \otimes \mathcal{N} = \text{the smallest } \sigma\text{-algebra in } X \times Y \text{ which contains all rectangles } A \times B \text{ where } A \in \mathcal{M} \text{ and } B \in \mathcal{N}.$$

Then $\mathcal{M} \otimes \mathcal{N}$ is called the *product σ-algebra* on $X \times Y$, and $(X \times Y, \mathcal{M} \otimes \mathcal{N})$ is a measurable space.

For the product of measure spaces, we have the following ([29, Section 2.5]):

Theorem 3.5. *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two σ-finite measure spaces. Then there exists a unique σ-finite, non-negative measure λ on $\mathcal{M} \otimes \mathcal{N}$ such that*

$$\lambda(A \times B) = \mu(A)\nu(B) \quad \textit{for } A \in \mathcal{M} \textit{ and } B \in \mathcal{N}.$$

The measure λ is called the *product measure* of μ and ν, and is denoted by $\mu \times \nu$.

3.1.7 Direct Image of Measures

Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be measurable spaces. A mapping f of X into Y is said to be *measurable* if the inverse image $f^{-1}(B)$ of every $B \in \mathcal{N}$ is in $\mathcal{M}$.

Let $(X, \mathcal{M}, \mu)$ be a measure space and $(Y, \mathcal{N})$ a measurable space. If $f \to X \to Y$ is a measurable mapping, then we can define a measure ν on $(Y, \mathcal{N})$ by the formula

$$\nu(B) = \mu(f^{-1}(B)) \quad \text{for } B \in \mathcal{N}.$$

We then write $\nu = f_*\mu$. The measure $f_*\mu$ is called the *direct image* of μ under f.

3.1.8 Integrals

Let $(X, \mathcal{M}, \mu)$ be a measure space. If A is a measurable subset of X and if $f(x)$ is a non-negative measurable simple function on A of the form

$$f(x) = \sum_{j=1}^{m} a_j \chi_{A_j}(x), \quad a_j \geq 0,$$

then we let

$$\int_A f(x)\,d\mu(x) = \sum_{j=1}^{m} a_j \mu(A_j). \tag{3.2}$$

The convention: $0 \cdot \infty = 0$ is used here; it may happen that $a_j = 0$ and $\mu(A_j) = \infty$. If $f(x)$ is a non-negative measurable function on A, we let

$$\int_A f(x)\,d\mu(x) = \sup \int_A s(x)\,d\mu(x) \tag{3.3}$$

where the supremum is taken over all measurable simple functions s on A such that $0 \leq s(x) \leq f(x)$, $x \in A$. We remark that if $f(x)$ is a non-negative simple function, then the two definitions (3.2) and (3.3) of the integral $\int_A f(x)\,d\mu(x)$ coincide.

If $f(x)$ is a measurable function on A, we can write it in the form

$$f(x) = f^+(x) - f^-(x),$$

where

$$f^+(x) = \max\{f(x), 0\},$$
$$f^-(x) = \max\{-f(x), 0\}.$$

Both $f^+(x)$ and $f^-(x)$ are non-negative measurable functions on A. Then we define the integral of $f(x)$ by the formula

$$\int_A f(x)\,d\mu(x) = \int_A f^+(x)\,d\mu(x) - \int_A f^-(x)\,d\mu(x),$$

provided at least one of the integrals on the right-hand side is finite. If both integrals are finite, we say that $f(x)$ is *μ-integrable* or simply *integrable* on A.

For simplicity, we abbreviate

$$\int_A f\,d\mu = \int_A f(x)\,d\mu(x).$$

If μ is the Lebesgue measure on $\mathbf{R}^n$, we customarily write

$$\int_A f(x)\,dx,$$

instead of $\int_A f(x)\,d\mu(x)$.

A proposition concerning the points of a measurable set A is said to hold *μ-almost everywhere* (μ-a.e.) or simply *almost everywhere* (a.e.) in A if there exists a measurable set N of measure zero such that the proposition holds for all $x \in A \setminus N$. For example, if $f(x)$ and $g(x)$ are measurable functions on A and if $\mu(\{x \in A : f(x) \neq g(x)\}) = 0$, then we say that $f(x) = g(x)$ a.e. in A.

The next three theorems are concerned with the interchange of integration and limit process:

Theorem 3.6 (the monotone convergence theorem). *Let $\{f_n(x)\}$ be an increasing sequence of non-negative measurable functions defined on a measurable set A. Then we have the formula*

$$\lim_{n\to\infty} \int_A f_n(x)\, d\mu(x) = \int_A \left(\lim_{n\to\infty} f_n(x)\right) d\mu(x).$$

Theorem 3.7 (Fatou's lemma). *Let $\{f_n(x)\}$ be a sequence of non-negative measurable functions defined on a measurable set A. Then we have the inequality*

$$\int_A \left(\liminf_{n\to\infty} f_n(x)\right) d\mu(x) \leq \liminf_{n\to\infty} \int_A f_n(x)\, d\mu(x).$$

Theorem 3.8 (the Lebesgue dominated convergence theorem). *Let $\{f_n(x)\}$ be a sequence of measurable functions defined on a measurable set A which converges to some function $f(x)$ almost everywhere in A. Assume that there exists a non-negative integrable function $g(x)$ defined on A such that*

$$|f_n(x)| \leq g(x) \quad \textit{a.e. in } A, \ n = 1, 2, \ldots.$$

Then the limit function $f(x)$ is integrable on A and we have the formula

$$\int_A f(x)\, d\mu(x) = \lim_{n\to\infty} \int_A f_n(x)\, d\mu(x).$$

3.1.9 Fubini's Theorem

We consider integration on product spaces. Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two σ-finite measure spaces, and let $\mu \times \nu$ be the unique product measure of μ and ν on the product σ-algebra $\mathcal{M} \otimes \mathcal{N}$.

First, let E be a subset of $X \times Y$. For a point $x \in X$, we define the x-section E_x of E by the formula

$$E_x = \{y \in Y : (x, y) \in E\}.$$

For a point $y \in Y$, we define the y-section E^y of E by the formula

$$E^y = \{x \in X : (x, y) \in E\}.$$

For example, if $E = A \times B$ where $A \subset X$ and $B \subset Y$, then we have the formulas

$$E_x = B,$$
$$E^y = A.$$

Secondly, let $f(x, y)$ be a function defined on $X \times Y$. For a point $x \in X$, we define the x-section f_x of f by the formula

$$f_x(y) = f(x, y) \quad \text{for } y \in E_x.$$

For a point $y \in Y$, we define the y-section f^y of f by the formula

$$f^y(x) = f(x, y) \quad \text{for } x \in E^y.$$

For example, if $E = A \times B$ where $A \subset X$ and $B \subset Y$, then we have the formulas

$$(\chi_E)_x(y) = \chi_{E_x}(y) = \chi_B(y) \quad \text{for } y \in Y,$$
$$(\chi_E)^y(x) = \chi_{E^y}(x) = \chi_A(x) \quad \text{for } x \in X.$$

Now we assume that $f(x, y)$ is an $\mathcal{M} \otimes \mathcal{N}$-measurable function on $X \times Y$ such that its integral exists. Then we customarily write

$$\iint_{X \times Y} f(x, y)\, d(\mu \times \nu)(x, y).$$

This integral is called the *double integral* of f. If it happens that the function

$$g(x) = \int_Y f(x, y)\, d\nu(y), \quad x \in X$$

is defined and also its integral exists, then we denote the integral

$$\int_X g(x)\, d\mu(x)$$

by any one of the following notation:

$$\int_X \left(\int_Y f(x, y)\, d\nu(y) \right) d\mu(x), \quad \int_X d\mu(x) \int_Y f(x, y)\, d\nu(y),$$
$$\iint_{X \times Y} f(x, y)\, d\nu(y)\, d\mu(x), \quad \iint_{X \times Y} f\, d\nu\, d\mu.$$

Similarly, we write

$$\int_Y \left(\int_X f(x, y) d\mu(x) \right) d\nu(y), \quad \int_Y d\nu(y) \int_X f(x, y)\, d\mu(x),$$

$$\iint_{X\times Y} f(x,y)\, d\mu(x)\, d\nu(y), \qquad \iint_{X\times Y} f\, d\mu\, d\nu.$$

These integrals are called the *iterated integrals* of $f(x,y)$.

The next theorem describes the most important relationship between double integrals and iterated integrals:

Theorem 3.9 (Fubini)**.** *Assume that $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ are σ-finite measure spaces. Then we have the following three assertions:*

(i) If $E \in \mathcal{M}\otimes\mathcal{N}$, then it follows that $E_x \in \mathcal{N}$ for all $x \in X$ and that $E^y \in \mathcal{M}$ for all $y \in Y$. Moreover, $\nu(E_x)$ is an $\mathcal{M}$-measurable function of x and $\mu(E^y)$ is an $\mathcal{N}$-measurable function of y, respectively, and we have the formula

$$\int_X \nu(E_x)\, d\mu(x) = \int_Y \mu(E^y)\, d\nu(y) = (\mu\times\nu)(E).$$

In particular, if $(\mu\times\nu)(E) < \infty$, then it follows that $\nu(E_x) < \infty$ for μ-almost all $x \in X$ and that $\mu(E^y) < \infty$ for ν-almost all $y \in Y$.

(ii) If $f(x,y)$ is a non-negative, $\mathcal{M}\otimes mathcal N$-measurable function on $X\times Y$, then it follows that $f_x(y)$ is an $\mathcal{N}$-measurable function of y for any $x \in X$ and that $f^y(x)$ is an $\mathcal{M}$-measurable function of x for all $y \in Y$. Moreover, the functions

$$X \ni x \longmapsto g(x) = \int_Y f_x(y)\, d\nu(y) = \int_Y f(x,y)\, d\nu(y),$$
$$Y \ni y \longmapsto h(y) = \int_X f^y(x)\, d\mu(x) = \int_X f(x,y)\, d\mu(x)$$

are $\mathcal{M}$-measurable and $\mathcal{N}$-measurable, respectively, and we have the formula

$$\int_X g(x)\, d\mu(x) = \int_Y h(y)\, d\nu(y) = \iint_{X\times Y} f(x,y)\, d(\mu\times\nu).$$

(iii) If $f(x,y)$ is a $\mu\times\nu$-integrable function on $X\times Y$, then it follows that the function $f_x(y)$ is ν-integrable for μ-almost all $x \in X$ and that the function $f^y(x)$ is μ-integrable for ν-almost all $y \in Y$. Furthermore, the function $g(x)$ is μ-integrable on X and the function $h(y)$ is ν-integrable on Y, respectively, and we have the formula

$$\int_X g(x)\, d\mu(x) = \int_Y h(y)\, d\nu(y) = \iint_{X\times Y} f(x,y)\, d(\mu\times\nu).$$

The next theorem gives a useful version of Fubini's theorem:

Theorem 3.10 (Fubini)**.** *Assume that $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ-finite measure spaces. Then we have the following two assertions:*

(i) *If $f(x, y)$ is a $\mu \otimes \nu$-integrable function on $X \times Y$, then the function $f_x(y)$ on Y is ν-integrable for μ- almost all $x \in X$, and the function $f^y(x)$ is μ-integrable for ν-almost all $y \in Y$. Furthermore, the function $g(x)$, defined by the formula*

$$g(x) = \int_Y f_x(y)\, d\nu(y) = \int_Y f(x, y)\, d\nu(y)$$

for μ-almost all $x \in X$, is μ-integrable, and the function $h(y)$, defined by the formula

$$h(y) = \int_X f^y(x)\, d\mu(x) = \int_X f(x, y)\, d\mu(x)$$

for ν-almost all $y \in Y$, is ν-integrable, respectively, and we have the formula

$$\iint_{X \times Y} f(x, y)\, d(\mu \times \nu) = \int_X g(x)\, d\mu = \int_Y h(y)\, d\nu.$$

(ii) *Conversely, if $f(x, y)$ is an $\mathcal{M} \otimes \mathcal{N}$-measurable function on $X \times Y$, then the functions*

$$\varphi(x) = \int_Y |f(x, y)|\, d\nu(y), \quad x \in X,$$
$$\psi(y) = \int_X |f(x, y)|\, d\mu(x), \quad y \in Y$$

are $\mathcal{M}$-measurable and $\mathcal{N}$-measurable, respectively, and we have the formula

$$\iint_{X \times Y} |f(x, y)|\, d(\mu \times \nu) = \int_X \varphi(x)\, d\mu = \int_Y \psi(y)\, d\nu.$$

Furthermore, if either $\varphi(x)$ or $\psi(y)$ is integrable, then $f(x, y)$ is integrable, and part (i) applies.

Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be complete, σ-finite measure spaces, and let $(X \times Y, \mathcal{L}, \mu \times \nu)$ be the *completion* of $(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \times \nu)$. The next theorem is a version of Fubini's theorem (Theorem 3.9) for *complete* measures:

Theorem 3.11 (Fubini)**.** *Let $(X \times Y, \mathcal{L}, \mu \times \nu)$ be the completion of the product measure space $(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \times \nu)$. Then we have the following three assertions:*

(i) If $E \in \mathcal{L}$, then it follows that $E_x \in \mathcal{N}$ for μ-almost all $x \in X$ and that $E^y \in \mathcal{M}$ for ν-almost all $y \in Y$. Moreover, $\nu(E_x)$ is an $\mathcal{M}$-measurable function of x and $\mu(E^y)$ is an $\mathcal{N}$-measurable function of y, respectively, and we have the formula

$$\int_X \nu(E_x)\, d\mu(x) = \int_Y \mu(E^y)\, d\nu(y) = (\mu \times \nu)(E).$$

In particular, if $\mu(E) < \infty$, then it follows that $\nu(E_x) < \infty$ for μ-almost all $x \in X$ and that $\mu(E^y) < \infty$ for ν-almost all $y \in Y$.

(ii) If $f(x,y)$ is an $\mathcal{L}$-measurable function on the product space $X \times Y$ such that $f(x,y) \geq 0$ for $\mu \times \nu$-almost all $(x,y) \in X \times Y$, then it follows that $f_x(y)$ is an $\mathcal{N}$-measurable function of y for μ-almost all $x \in X$ and that $f^y(x)$ is an $\mathcal{M}$-measurable function of x for ν-almost all $y \in Y$. Moreover, the functions

$$X \ni x \longmapsto g(x) = \int_Y f(x,y)\, d\nu(y),$$
$$Y \ni y \longmapsto h(y) = \int_X f(x,y)\, d\mu(x)$$

are $\mathcal{M}$-measurable and $\mathcal{N}$-measurable, respectively, and the formula

$$\int_X g(x)\, d\mu(x) = \int_Y h(y)\, d\nu(y) = \iint_{X \times Y} f(x,y)\, d(\mu \times \nu)$$

holds true.

(iii) If $f(x,y)$ is an $\mathcal{L}$-measurable and $\mu \times \nu$-integrable function on $X \times Y$, then it follows that the function $f_x(y)$ is ν-integrable for μ-almost all $x \in X$ and that the function $f^y(x)$ is μ-integrable for ν-almost all $y \in Y$. Furthermore, the function $g(x)$ is μ-integrable on X and the function $h(y)$ is ν-integrable on Y, respectively, and we have the formula

$$\int_X g(x)\, d\mu(x) = \int_Y h(y)\, d\nu(y) = \iint_{X \times Y} f(x,y)\, d(\mu \times \nu).$$

The next corollary gives a useful version of Fubini's theorem for *complete* measures:

Corollary 3.12. *Let $(X \times Y, \mathcal{L}, \mu \times \nu)$ be the completion of the product measure space $(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \times \nu)$. If $f(x, y)$ is an $\mathcal{L}$-measurable function on $X \times Y$, then the functions*

$$\varphi(x) = \int_Y |f(x, y)| \, d\nu(y), \quad x \in X,$$
$$\psi(y) = \int_X |f(x, y)| \, d\mu(x), \quad y \in Y$$

are $\mathcal{M}$-measurable and $\mathcal{N}$-measurable, respectively, and we have the formula

$$\iint_{X \times Y} |f(x, y)| \, d(\mu \times \nu) = \int_X \varphi(x) \, d\mu(x) = \int_Y \psi(y) \, d\nu(y).$$

Furthermore, if either $\varphi(x)$ or $\psi(y)$ is integrable, then $f(x, y)$ is integrable, and Theorem 3.11 applies.

3.2 L^p Spaces

Let $(X, \mathcal{M}, \mu)$ be a measure space. Two $\mathcal{M}$-measurable functions f and g are said to be equivalent if they are equal μ-almost everywhere in X with respect to the measure μ, that is, if $f(x) = g(x)$ for all x outside of a set of μ-measure zero:

$$f \sim g \iff f(x) = g(x) \quad \text{for } \mu\text{-almost all } x \in X.$$

This is obviously an equivalence relation.

If $1 \leq p < \infty$, we let

$$L^p(X) = \text{the space of equivalence classes of } \mathcal{M}\text{-measurable functions } f \text{ on } X \text{ such that } |f|^p \text{ is } \mu\text{-integrable on } X.$$

We define

$$\|f\|_p := \left(\int_X |f(x)|^p \, d\mu(x) \right)^{1/p}.$$

An $\mathcal{M}$-measureable Lebesgue measurable function f on X is said to be *essentially bounded* if there exists a constant $C > 0$ such that

$$|f(x)| \leq C \quad \text{for } \mu\text{-almost all } x \in X.$$

For $p = \infty$, we let

$$L^\infty(X) = \text{the space of equivalence classes of essentially bounded,}$$

$\mathcal{M}$-measurable functions f on X.

We define

$$\begin{aligned}\|f\|_\infty &:= \operatorname{ess\,sup}_{x\in X} |f(x)| \\ &= \inf\{C : |f(x)| \leq C \text{ for } \mu\text{-almost all } x \in X\}.\end{aligned}$$

Then we have the following theorem:

Theorem 3.13. *(i) The space $L^p(X)$, $1 \leq p < \infty$, is a Banach space with the norm $\|\cdot\|_p$.*

(ii)The space $L^\infty(X)$ is a Banach space with the norm $\|\cdot\|_\infty$.

In order to show that $L^p(X)$ is a normed linear space, we need the following two inequalities:

Theorem 3.14 (Hölder's inequality)**.** *Let $1 \leq p$ and $q \leq \infty$ such that $1/p + 1/q = 1$. Then we have, for all $f \in L^p(X)$ and $g \in L^q(X)$,*

$$\left|\int_X f(x)g(x)\,d\mu(x)\right| \leq \left(\int_X |f(x)|^p\,d\mu(x)\right)^{1/p} \left(\int_X |g(x)|^q\,d\mu(x)\right)^{1/q}.$$

Theorem 3.15 (Minkowski's inequality)**.** *Let $1 \leq p \leq \infty$. Then we have, for all f, $g \in L^p(X)$,*

$$\begin{aligned}&\left(\int_X |f(x)+g(x)|^p\,d\mu\right)^{1/p} \qquad (3.4)\\ &\leq \left(\int_X |f(x)|^p\,d\mu\right)^{1/p} + \left(\int_X |g(x)|^p\,d\mu\right)^{1/p}.\end{aligned}$$

3.3 Minkowski's Inequality for Integrals

Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two σ-finite, complete measure spaces and let $(\mathcal{M} \times \mathcal{N})^*$ be the completion of $\mathcal{M} \otimes \mathcal{N}$ with respect to the product measure $\mu \times \nu$.

Then Minkowski's inequality for $1 \leq p \leq \infty$ (Theorem 3.15) can be generalized as follows:

Theorem 3.16 (Minkowski's inequality for integrals)**.** *Let $1 \leq p \leq \infty$ and assume that $f(x, y)$ is an $(\mathcal{M} \times \mathcal{N})^*$-measureble function on $X \times Y$, and satisfies the conditions*

$$f(\cdot, y) \in L^p(X) \quad \text{for } \nu\text{-almost all } y \in Y, \qquad (3.5a)$$

$$\int_Y \|f(\cdot, y)\|_p \, d\nu(y) < \infty. \tag{3.5b}$$

Then we have the following three assertions (3.6a), (3.6b) *and* (3.6c)*:*

$$f(x, \cdot) \in L^1(Y) \quad \textit{for } \mu\textit{-almost all } x \in X. \tag{3.6a}$$

$$F(x) = \int_Y f(x, y)\, d\nu(y) \in L^p(X). \tag{3.6b}$$

$$\|F\|_p \leq \int_Y \|f(\cdot, y)\|_p \, d\nu(y). \tag{3.6c}$$

Inequality (3.6c) *is called* Minkowski's inequality for integrals.

Proof. The proof is divided into two steps.

Step 1: First, we show that

The function $y \mapsto \|f(\cdot, y)\|_p$ is $\mathcal{N}$-measurable.

Step 1-a: The case where $1 \leq p < \infty$. Since the function $|f(x, y)|^p$ is $(\mathcal{M} \times \mathcal{N})^*$-measurable on $X \times Y$, it follows from an application of Fubini's theorem (Theorem 3.10) that the function

$$x \longmapsto |f(x, y)|^p$$

is $\mathcal{M}$-measurable for ν-almost all $y \in Y$, and that the function

$$y \longmapsto \int_X |f(x, y)|^p \, d\mu(x)$$

is $\mathcal{N}$-measurable. Hence we find that the function

$$y \longmapsto \|f(\cdot, y)\|_p = \left(\int_X |f(x, y)|^p \, d\mu(x) \right)^{1/p}$$

is $\mathcal{N}$-measurable.

Step 1-b: The case where $p = \infty$. We assume that

$$\mu(X) < \infty.$$

Then we have, part (ii) of Theorem 3.21 in Section 3.6,

$$\|f(\cdot, y)\|_\infty = \lim_{q \uparrow \infty} \|f(\cdot, y)\|_q \quad \text{for } \nu\text{-almost all } y \in Y.$$

This proves the $\mathcal{N}$-measurability of the function $\|f(\cdot, y)\|_\infty$, since the function $y \mapsto \|f(\cdot, y)\|_q$ is $\mathcal{N}$-measurable.

Now we assume that X is σ-finite, that is, it is a countable union of sets X_n of finite measure:

$$X = \bigcup_{n=1}^{\infty} X_n, \quad \mu(X_n) < \infty.$$

Since we have the assertion

$$\|f(\cdot, y)\|_\infty = \sup_{n\geq 1} \|f(\cdot, y)\|_\infty \quad \text{for } \nu\text{-almost all } y \in Y,$$

it follows that the function $y \mapsto \|f(\cdot, y)\|_\infty$ is $\mathcal{N}$-measurable.

Step 2: Now we prove inequality (3.6).

Step 2-a: The case where $p = \infty$. We remark that

$$|f(x, y)| \leq \|f(\cdot, y)\|_\infty$$
$$\text{for } \mu\text{-almost all } x \in X \text{ and for } \nu\text{-almost all } y \in Y.$$

Hence we have, by conditions (3.5a) and (3.5b),

$$\int_Y |f(x, y)|\, d\nu(y) \leq \int_Y \|f(\cdot, y)\|_\infty \, d\nu(y) < \infty \quad \text{for } \mu\text{-almost all } x \in X,$$

and so

$$f(x, \cdot) \in L^1(Y) \quad \text{for } \mu\text{-almost all } x \in X.$$

Furthermore, in view of Fubini's theorem (Theorem 3.10) it follows that the function

$$x \longmapsto F(x) = \int_Y f(x, y)\, d\nu(y)$$

is $\mathcal{M}$-measurable and that

$$|F(x)| \leq \int_Y \|f(\cdot, y)\|_\infty \, d\nu(y) \quad \text{for } \mu\text{-almost all } x \in X.$$

This proves the desired inequality (3.6c) for $p = \infty$:

$$\|F\|_\infty \leq \int_Y \|f(\cdot, y)\|_\infty \, d\nu(y).$$

Step 2-b: The case where $p = 1$. By Fubini's theorem (Theorem 3.10), we have the inequality

$$\begin{aligned} \|F\|_1 &= \int_X \left| \int_Y f(x, y)\, d\nu(y) \right| d\mu(x) \\ &\leq \int_X \int_Y |f(x, y)|\, d\nu(y)\, d\mu(x) = \int_Y \left(\int_X |f(x, y)|\, d\mu(x) \right) d\nu(y) \\ &= \int_Y \|f(\cdot, y)\|_1 \, d\nu(y). \end{aligned}$$

This proves the desired inequality (3.6c) for $p = 1$.

Step 2-c: The case where $1 < p < \infty$.

(i) First, we assume that

$$f(x,y) \geq 0 \quad \text{on } X \times Y. \tag{3.7}$$

Then, in view of Fubini's theorem (Theorem 3.10) it follows that the function

$$x \longmapsto F(x) = \int_Y f(x,y)\, d\nu(y)$$

is $\mathcal{M}$-measurable. Thus we have, by Hölder's inequality (Theorem 3.14),

$$\begin{aligned}
\int_X |F(x)|^p \, d\mu(x) &= \int_X |F(x)|^{p-1} \left(\int_Y f(x,y)\, d\nu(y) \right) d\mu(x) \\
&= \int_Y \left(\int_X |F(x)|^{p-1} f(x,y)\, d\mu(x) \right) d\nu(y) \\
&\leq \int_Y \left(\int_X \|F\|_p^{p-1} \|f(\cdot,y)\|_p \right) d\nu(y) \\
&= \|F\|_p^{p-1} \int_Y \|f(\cdot,y)\|_p \, d\nu(y).
\end{aligned}$$

Therefore, if we have the inequality

$$\|F\|_p < \infty,$$

it follows that

$$\|F\|_p \leq \int_Y \|f(\cdot,y)\|_p \, d\nu(y).$$

Now we assume that X and Y are σ-finite, that is, they are the countable unions of sets X_n and Y_n of finite measure, respectively:

$$X = \bigcup_{n=1}^{\infty} X_n, \quad \mu(X_n) < \infty.$$

$$Y = \bigcup_{n=1}^{\infty} Y_n, \quad \nu(Y_n) < \infty.$$

We let

$$f(x,y) = \begin{cases} \min\{f(x,y), n\} & \text{if } x \in X_n,\ y \in Y_n, \\ 0 & \text{otherwise.} \end{cases}$$

Then it follows that

$$f_n(x,y) \uparrow f(x,y) \quad \text{as } n \to \infty,$$

so that

$$F_n(x) = \int_Y f_n(x,y) d\nu(y) \uparrow F(x) = \int_Y f(x,y)\, d\nu(y) \quad \text{as } n \to \infty.$$

We remark that

$$\|F_n\|_p < \infty.$$

Hence it follows that

$$\|F_n\|_p \leq \int_Y \|f_n(\cdot,y)\|_p \, d\nu(y) \leq \int_Y \|f(\cdot,y)\|_p \, d\nu(y).$$

Therefore, by passing to the limit we obtain that

$$\|F\|_p \leq \int_Y \|f(\cdot,y)\|_p \, d\nu(y).$$

This proves the desired inequality (3.6c) for $1 < p < \infty$ under condition (3.7).

(ii) We consider the general case. By applying Step (i) to the function $|f(x,y)|$, we obtain that

$$G(x) = \int_Y |f(x,y)|\, d\nu(y) \in L^p(X),$$
$$\|G\|_p \leq \int_Y \|f(\cdot,y)\|_p \, d\nu(y).$$

Hence we have the assertion

$$G(x) < \infty \quad \text{for } \mu\text{-almost all } x \in X,$$

and so

$$f(x,\cdot) \in L^1(Y) \quad \text{for } \mu\text{-almost all } x \in X.$$

However, in view of Hölder's inequality (Theorem 3.14) it follows that

$$G(x) \in L^1(X_n), \quad n = 1, 2, \ldots.$$

By Fubini's theorem, this implies that

$$F(x) = \int_Y f(x,y)\, d\nu(y) \in L^1(X_n), \quad n = 1, 2, \ldots.$$

In particular, it follows that F is $\mathcal{M}$-measurable on each X_n, and so it is $\mathcal{M}$-measurable on $X = \bigcup_n X_n$.

Moreover, we have the inequality

$$\|F\|_p \leq \|G\|_p \leq \int_Y \|f(\cdot,y)\|_p \, d\nu(y).$$

This proves the desired inequality (3.6c) for $1 < p < \infty$.

The proof of Theorem 3.16 is now complete. □

3.4 Hardy's Inequality

First, we prove a general integral inequality on the interval $(0, \infty)$:

Theorem 3.17. *Let $K(x, y)$ be a Lebesgue measurable function defined on the interval $(0, \infty)$. Assume that $K(x, y)$ is positively homogeneous of degree -1, that is,*

$$K(\lambda x, \lambda y) = \lambda^{-1} K(x, y), \quad \lambda > 0,$$

and further that the integral

$$A_K := \int_0^\infty |K(1, y)| y^{-1/p} \, dy$$

is finite for some $1 \leq p \leq \infty$.

Then the operator Tf, defined by the formula

$$Tf(x) = \int_0^\infty K(x, y) f(y) \, dy,$$

is bounded from $L^p(0, \infty)$ into itself. More precisely, we have the inequality

$$\|Tf\|_p \leq A_K \|f\|_p, \quad f \in L^p(0, \infty).$$

Proof. Since we have the formula

$$\begin{aligned} Tf(x) &= \int_0^\infty K(x, y) f(y) \, dy \\ &= \int_0^\infty K(x \cdot 1, x \cdot z) f(zx) \, x \, dz = \int_0^\infty x^{-1} K(1, z) f(zx) \, x \, dz \\ &= \int_0^\infty K(1, z) f(zx) \, dz, \end{aligned}$$

by applying Minkowski's inequality for integrals (Theorem 3.16) we obtain that

$$\|Tf\|_p \leq \int_0^\infty |K(1, z)| \, \|f(z \cdot)\|_p \, dz.$$

However, it is easy to see that

$$\|f(z \cdot)\|_p = \left(\int_0^\infty |f(zx)|^p \, dx \right)^{1/p} = z^{-1/p} \left(\int_0^\infty |f(y)|^p \, dy \right)^{1/p}$$

$$= z^{-1/p} \|f\|_p .$$

Therefore, we have the inequality

$$\|Tf\|_p \leq \left(\int_0^\infty |K(1,z)| z^{-1/p}\, dz\right) \|f\|_p = A_K \|f\|_p .$$

The proof of Theorem 3.17 is complete. □

Now we can prove Hardy's inequality which will be used systematically:

Theorem 3.18 (Hardy's inequality)**.** *Let $1 \leq p \leq \infty$ and $\gamma \neq 0$. If $f(x)$ is a non-negative, Lebesgue measurable function on the interval $(0,\infty)$, we define a function $F(x)$ by the formula*

$$F(x) = \begin{cases} \int_0^x f(y)\, dy & \text{if } \gamma < 0, \\ \int_x^\infty f(y)\, dy & \text{if } \gamma > 0. \end{cases}$$

Then we have the inequality

$$\left(\int_0^\infty (x^\gamma F(x))^p \frac{dx}{x}\right)^{1/p} \leq \frac{1}{|\gamma|} \left(\int_0^\infty (x^{\gamma+1} f(x))^p \frac{dx}{x}\right)^{1/p} . \tag{3.8}$$

Proof. We only consider the case where $\gamma < 0$. The case where $\gamma > 0$ is proved similarly.

If we let

$$K(x,y) := \begin{cases} x^{\gamma-1/p} y^{-\gamma+1/p-1} & \text{if } 0 \leq y \leq x, \\ 0 & \text{if } x < y. \end{cases}$$

then it follows that $K(x,y)$ is positively homogeneous of degree -1 and satisfies the conditions

$$\int_0^\infty K(1,y) y^{-1/p}\, dy = \int_0^1 y^{-\gamma-1}\, dy = -\frac{1}{\gamma}.$$

If we introduce an integral operator

$$\begin{aligned} Tg(x) &:= \int_0^\infty K(x,y)\, g(y)\, dy \\ &= x^{\gamma-1/p} \int_0^x y^{-\gamma+1/p-1}\, g(y)\, dy, \quad g \in L^p(0,\infty), \end{aligned}$$

then, by applying Theorem 3.17 to our situation we obtain that

$$\left(\int_0^\infty \left(x^{\gamma-1/p} \int_0^x y^{-\gamma+1/p-1} g(y)\, dy\right)^p dx\right)^{1/p}$$

$$\leq \frac{1}{|\gamma|} \left(\int_0^\infty g(y)^p \, dy \right)^{1/p}.$$

In particular, if we let

$$g(y) := y^{\gamma - 1/p + 1} f(y),$$

then it follows that

$$\begin{aligned}
\left(\int_0^\infty (x^\gamma F(x))^p \frac{dx}{x} \right)^{1/p} &= \left(\int_0^\infty \left(x^{\gamma - 1/p} \int_0^x f(x) \, dy \right)^p dx \right)^{1/p} \\
&\leq \frac{1}{|\gamma|} \left(\int_0^\infty \left(y^{\gamma - 1/p + 1} f(y) \right)^p dy \right)^{1/p} \\
&= \frac{1}{|\gamma|} \left(\int_0^\infty (x^{\gamma+1} f(x))^p \frac{dx}{x} \right)^{1/p}.
\end{aligned}$$

The proof of Theorem 3.18 is complete. □

Example 3.1. If we let

$$\gamma =: \frac{1}{p} - 1, \quad 1 < p \leq \infty,$$

then we have, by inequality (3.10),

$$\left(\int_0^\infty \left(\frac{1}{x} \int_0^x f(y) \, dy \right)^p dx \right)^{1/p} \leq \frac{p}{p-1} \left(\int_0^\infty f(y)^p \, dy \right)^{1/p}.$$

3.5 The Generalized Hölder Inequality

Hölder's inequality (Theorem 3.14) can be generalized as follows:

Theorem 3.19 (the generalized Hölder inequality)**.** *Let $1 \leq p, q, r \leq \infty$ such that*

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{r}.$$

Then we have, for all $f \in L^p(X)$ and $g \in L^q(X)$,

$$\begin{aligned}
&\left(\int_X |f(x) g(x)|^r \, d\mu \right)^{1/r} \\
&\leq \left(\int_X |f(x)|^p \, d\mu \right)^{1/p} \left(\int_X |g(x)|^q \, d\mu \right)^{1/q}.
\end{aligned} \tag{3.9}$$

Proof. We only consider the case where $1 \leq p, q < \infty$. Since we have the formula

$$\frac{1}{\frac{p}{r}} + \frac{1}{\frac{q}{r}} = 1,$$

the desired inequality (3.9) follows by applying Hölder's inequality (Theorem 3.14) to the functions $|f|^r \in L^{p/r}(X)$ and $|g|^r \in L^{q/r}(X)$.

The proof of Theorem 3.19 is complete. □

Some basic properties of the L^p spaces are summarized in the following theorem:

Theorem 3.20. *Assume that*

$$0 < \mu(X) < \infty.$$

Then we have the following two assertions (i) and (ii):

(i) We have the inclusion

$$L^p(X) \subset L^q(X), \quad 1 \leq q < p \leq \infty,$$

with continuous injection.

More precisely, we have the inequality

$$\|f\|_q \leq \mu(X)^{1/q-1/p} \|f\|_p , \quad f \in L^p(X). \tag{3.10}$$

(ii) The norm $\|\cdot\|_p$ is continuous for $p \geq 1$, that is,

$$\lim_{q \uparrow p} \|f\|_q = \|f\|_p , \quad f \in L^p(X).$$

Proof. (i) By Hölder's inequality (Theorem 3.14), it follows that

$$\begin{aligned}\int_X |f(x)|^q \, d\mu &\leq \left(\int_X 1 \, d\mu\right)^{(p-q)/p} \left(\int_X |f(x)|^p \, d\mu\right)^{q/p} \\ &= \mu(X)^{1-p/q} \|f\|_p^q ,\end{aligned}$$

so that

$$\|f\|_q \leq \mu(X)^{1/q-1/p} \|f\|_p .$$

This proves the desired inequality (3.10).

(ii) First, we remark that, by inequality (3.10),

$$\limsup_{q \uparrow p} \|f\|_q \leq \|f\|_p .$$

(ii-a) The case $p = \infty$: It suffices to show that

$$\liminf_{q \uparrow \infty} \|f\|_q \geq \|f\|_\infty . \tag{3.11}$$

We may assume that

$$\|f\|_\infty > 0.$$

For each $\varepsilon > 0$, we can find a set $X_1 \subset X$ such that

$$X_1 \subset \{x \in X : |f(x)| \geq \|f\|_\infty - \varepsilon\},$$
$$\mu(X_1) > 0.$$

Then we have the inequality

$$\begin{aligned}\|f\|_q &= \left(\int_X |f(x)|^q \, d\mu(x)\right)^{1/q} \geq \left(\int_{X_1} |f(x)|^q \, d\mu(x)\right)^{1/q} \\ &\geq (\|f\|_\infty - \varepsilon)\mu(X_1)^{1/q},\end{aligned}$$

and so

$$\liminf_{q \uparrow \infty} \|f\|_q \geq \|f\|_\infty - \varepsilon.$$

This proves the desired inequality (3.13), since ε is arbitrary.

(ii-b) The case where $1 < p < \infty$: We let

$$X_1 = \{x \in X : |f(x)| > 1\},$$
$$X_2 = \{x \in X : |f(x)| \leq 1\}.$$

Then we have the formula

$$\|f\|_q^q = \int_{X_1} |f(x)|^q \, d\mu(x) + \int_{X_2} |f(x)|^q \, d\mu(x).$$

However, we have the inequalities

$$|f(x)|^q \leq |f(x)|^p, \quad x \in X_1,$$
$$|f(x)|^q \leq 1, \quad x \in X_2.$$

Hence, by applying the monotone convergence theorem (Theorem 3.6) to the first term and also the Lebesgue dominated convergence theorem (Theorem 3.8) to the second term we obtain that

$$\lim_{q \uparrow p} \|f\|_q^q = \int_{X_1} |f(x)|^p \, d\mu(x) + \int_{X_2} |f(x)|^p \, d\mu(x) = \|f\|_p^p.$$

Now the proof of Theorem 3.20 is complete. □

3.6 The Generelized Young Inequality

Now we prove a general theorem about integral operators on a measure space:

Theorem 3.21 (the generalized Young inequality). *Let* $(X, \mathcal{M}, \mu)$ *and* $(Y, \mathcal{N}, \nu)$ *be* σ*-finite, complete measure spaces and let* $1 \le p,\ q,\ r \le \infty$ *such that*

$$\frac{1}{p} + \frac{1}{r} = 1 + \frac{1}{q}.$$

Assume that $K(x, y)$ *is an* $(\mathcal{M} \times \mathcal{N})^*$*-measurable function on the product space* $X \times Y$ *such that*

$$\int_X |K(x,y)|^r \, d\mu(x) \le M \quad \text{for } \nu\text{-almost all } y \in Y$$

and that

$$\int_Y |K(x,y)|^r \, d\nu(y) \le N \quad \text{for } \mu\text{-almost all } x \in X,$$

where M *and* N *are positive constants.*

If $f \in L^p(X)$*, then the function* $Tf(x)$*, defined by the formula*

$$Tf(x) = \int_Y K(x,y) f(y) \, d\nu(y),$$

is well-defined for μ*-almost all* $x \in X$*, and is in the space* $L^q(X)$*.*

Furthermore, we have the inequality

$$\|Tf\|_q \le M^{1/q} N^{1-(1/p)} \|f\|_p \,. \tag{3.12}$$

Proof. The proof of Theorem 3.21 is divided into two steps.

Step 1: First, we assume that

$$\begin{aligned} &K(x,y) \ge 0 \quad \text{on } X \times Y, \\ &f(y) \ge 0 \quad \text{on } Y. \end{aligned}$$

Then we have, by Fubini's theorem (Theorem 3.10), the following two assertions (i) and (ii):

(i) The function $y \mapsto K(x,y) f(y)$ is $\mathcal{N}$-measurable for μ-almost all $x \in X$.

(ii) The function $Tf(x) = \int_Y K(x,y) f(y) \, d\nu(y)$ is $\mathcal{M}$-measurable on X.

By the generalized Hölder inequality (Theorem 3.19), it follows that

$$\begin{aligned} Tf(x) &= \int_Y K(x,y)^{1-r/q} K(x,y)^{r/q} f(y)^{p/q} f(y)^{1-p/q}\, d\nu(y) \qquad (3.13) \\ &\le \left(\int_Y K(x,y)^r\, d\nu(y)\right)^{1-1/p} \left(\int_Y K(x,y)^r f(y)^p\, d\nu(y)\right)^{1/q} \\ &\quad \times \left(\int_Y f(y)^p\, d\nu(y)\right)^{1/p-1/q} \\ &\le N^{1-1/p} \|f\|_p^{1-p/q} \left(\int_X K(x,y)^r f(y)^p\, d\nu(y)\right)^{1/q}. \end{aligned}$$

Step I-a: The case where $q = \infty$: In view of inequality (3.13), it follows that

$$Tf(x) \le N^{1-1/p} \|f\|_p \quad \text{for } \mu\text{-almost all } x \in X,$$

so that

$$\|Tf\|_\infty \le N^{1-1/p} \|f\|_p .$$

This proves the desired inequality (3.12) for $q = \infty$.

Step I-b: The case where $1 \le q < \infty$: We have, by Fubini's theorem (Theorem 3.10),

$$\begin{aligned} &\int_X Tf(x)^q\, d\mu(x) \\ &\le \left(N^{1-1/p} \|f\|_p^{1-p/q}\right)^q \int_X \int_Y K(x,y)^r f(y)^p\, d\nu(y)\, d\mu(x) \\ &\le \left(N^{1-1/p} \|f\|_p^{1-p/q}\right)^q M \|f\|_p^p \\ &= \left(M^{1/q} N^{1-1/p} \|f\|_p\right)^q, \end{aligned}$$

and so

$$\|Tf\|_q \le M^{1/q} N^{1-1/p} \|f\|_p .$$

This proves the desired inequality (3.12) for $1 < q < \infty$.

Moreover, it follows that

$$Tf(x) < \infty \quad \text{for } \mu\text{-almost all } x \in X.$$

Step II: We consider the general case. To do this, we remark that

$$\begin{aligned} K(x,y)f(y) &= \left[(\operatorname{Re} K(x,y)f(y))^+ - (\operatorname{Re} K(x,y)f(y))^-\right] \\ &\quad + \sqrt{-1}\left[(\operatorname{Im} K(x,y)f(y))^+ - (\operatorname{Im} K(x,y)f(y))^-\right], \end{aligned}$$

and that

$$(\operatorname{Re} K(x,y)f(y))^{\pm} \leq |K(x,y)f(y)|,$$
$$(\operatorname{Im} K(x,y)f(y))^{\pm} \leq |K(x,y)f(y)|.$$

Here

$$f^{+}(x) = \max\{f(x), 0\},$$
$$f^{-}(x) = \max\{-f(x), 0\}.$$

However, we have, by Step I,

$$\int_Y |K(x,y)||f(y)|\,d\nu(y) < \infty \quad \text{for } \mu\text{-almost all } x \in X.$$

Hence it follows that

(1) The function $y \mapsto K(x,y)f(y)$ is $\mathcal{N}$-integrable for μ-almost all $x \in X$.
(2) The function $Tf(x) = \int_Y K(x,y)f(y)\,d\nu(y)$ is $\mathcal{M}$-measurable on X.

Furthermore, since we have the inequality

$$|Tf(x)| \leq \int_Y |K(x,y)||f(y)|\,d\nu(y) \quad \text{for } \mu\text{-almost all } x \in X,$$

we find that the desired inequality (3.12) remains valid for this case.

The proof of Theorem 3.21 is complete. □

The case where $r := 1$ and $p := q$ is useful in applications:

Corollary 3.22 (Schur's lemma). *Assume that $K(x,y)$ is an $(\mathcal{M}\times\mathcal{N})^*$-measurable function on the product space $X \times Y$ such that*

$$\int_X |K(x,y)|\,d\mu(x) \leq M \quad \textit{for } \nu\textit{-almost all } y \in Y$$

and that

$$\int_Y |K(x,y)|\,d\nu(y) \leq N \quad \textit{for } \mu\textit{-almost all } x \in X,$$

where M and N are positive constants.

If $f \in L^p(X)$, then the function $Tf(x)$, defined by the formula

$$Tf(x) = \int_Y K(x,y)f(y)\,d\nu(y),$$

is well-defined for μ-almost all $x \in X$, and is in $L^p(X)$.

Furthermore, we have the inequality

$$\|Tf\|_p \leq M^{1/p} N^{1-(1/p)} \|f\|_p .$$

3.7 Convolutions

In this section we prove a useful inequality for convolutions:

Theorem 3.23 (Young's inequality)**.** *Let $1 \leq p,\ q,\ r \leq \infty$ such that*

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1.$$

If $f \in L^p(\mathbf{R}^n)$ and $g \in L^q(\mathbf{R}^n)$, then the function $(f * g)(x)$, defined by the formula

$$(f * g)(x) = \int_{\mathbf{R}^n} f(x - y)\, g(y)\, dy,$$

is well-defined for almost all $x \in \mathbf{R}^n$, and is in $L^r(\mathbf{R}^n)$. Furthermore, we have the inequality

$$\|f * g\|_r \leq \|f\|_p \|g\|_q . \tag{3.14}$$

The function $f * g$ is called the convolution *of f and g.*

Proof. **Step I**: First, we have to verify that

$$\text{The function } f(x \pm y) \text{ is Lebesgue measurable on } \mathbf{R}^n \times \mathbf{R}^n. \tag{3.15}$$

Step I-a: First, we prove assertion (3.15) under the condition that $f(x)$ is Borel measurable on $\mathbf{R}^n$, that is,

Claim 3.1. If $f(x)$ is Borel measurable on $\mathbf{R}^n$, then $f(x \pm y)$ is Borel measurable on $\mathbf{R}^n \times \mathbf{R}^n$.

The proof of Claim 3.1 is divided into five steps:

Step (1): If $f(x)$ is continuous on $\mathbf{R}^n$, then $f(x \pm y)$ is continuous on $\mathbf{R}^n \times \mathbf{R}^n$, and so it is Borel measurable on $\mathbf{R}^n \times \mathbf{R}^n$.

Step (2): Let G be an arbitrary open subset of $\mathbf{R}^n$, and let

$$f(x) := \text{the characteristic function } \chi_G(x) \text{ of } G.$$

If we let

$$f_N(x) := \min \{N \operatorname{dist}(x, G^c), 1\} \quad \text{for } N \in \mathbf{N},$$

then it follows that

$$f_N(x) \text{ is continuous on } \mathbf{R}^n,$$

$$f_N(x) \longrightarrow f(x), \quad x \in \mathbf{R}^n.$$

Hence we find from Step (1) that the function

$$f(x \pm y) = \lim_{N \to \infty} f_N(x \pm y)$$

is Borel measurable on $\mathbf{R}^n \times \mathbf{R}^n$, if $f = \chi_G$.

Step (3): We define

$$\mathfrak{A} = \{A \subset \mathbf{R}^n : \chi_A(x \pm y) \text{ is Borel measurable on } \mathbf{R}^n \times \mathbf{R}^n\}.$$

Then $\mathfrak{A}$ is a σ-algebra.

Indeed, we have the following three assertions (a), (b) and (c):

(a) The empty set $\emptyset$ is in $\mathfrak{A}$, since $\chi_\emptyset = 0$.

(b) If $A \in \mathfrak{A}$, then we have $\chi_{A^c} = 1 - \chi_A$, so that the function

$$\chi_{A^c}(x \pm y) = 1 - \chi_A(x \pm y)$$

is Borel measurable on $\mathbf{R}^n \times \mathbf{R}^n$.

(c) If $A_j \in \mathfrak{A}$ and $A = \bigcup_{j=1}^\infty A_j$, then the function

$$\chi_A(x \pm y) = \sup_{j \geq 1} \chi_{A_j}(x \pm y)$$

is Borel measurable on $\mathbf{R}^n \times \mathbf{R}^n$.

Step (4): By Steps (2) and (3), it follows that $\mathfrak{A}$ is a σ-algebra which contains all open subsets of $\mathbf{R}^n$; hence $\mathfrak{A}$ contains all Borel subsets of $\mathbf{R}^n$. Namely, we find that the function $\chi_E(x \pm y)$ is Borel measurable on $\mathbf{R}^n \times \mathbf{R}^n$, for each Borel subset E of $\mathbf{R}^n$.

Step (5): Let $f(x)$ be an arbitrary Borel measurable function on $\mathbf{R}^n$. We may assume that

$$f(x) \geq 0 \quad \text{on } \mathbf{R}^n.$$

Then we can find an increasing sequence of non-negative, Borel measurable simple functions $f_N(x)$ on $\mathbf{R}^n$ such that

$$f_N(x) \longrightarrow f(x) \quad \text{for } x \in \mathbf{R}^n.$$

Hence, by Step (4) it follows that the function

$$f(x \pm y) = \lim_{N \to \infty} f_N(x \pm y)$$

is Borel measurable on $\mathbf{R}^n \times \mathbf{R}^n$.

This completes the proof of Claim 3.1. □

Step I-b: Now we prove assertion (3.15).

Let $f(x)$ be an arbitrary Lebesgue measurable function on $\mathbf{R}^n$. We may assume that

$$f(x) \geq 0 \quad \text{on } \mathbf{R}^n.$$

Then we can find two Borel measurable functions $g(x)$ and $h(x)$ on $\mathbf{R}^n$ such that

$$\begin{aligned} &0 \leq g(x) \leq f(x) \leq h(x) \quad \text{on } \mathbf{R}^n, \\ &g(x) = f(x) = h(x) \quad \text{for almost all } x \in \mathbf{R}^n. \end{aligned}$$

This follows from the fact that the Lebesgue measure is the completion of the Borel measure.

Hence we have the inequalities

$$0 \leq g(x \pm y) \leq f(x \pm y) \leq h(x \pm y) \quad \text{for } (x, y) \in \mathbf{R}^n \times \mathbf{R}^n.$$

However, we know from Step I-a that the functions $g(x \pm y)$ and $h(x \pm y)$ are Borel measurable on $\mathbf{R}^n \times \mathbf{R}^n$. Thus, by applying Fubini's theorem (Theorem 3.10) we obtain that

$$\begin{aligned} &\int_{\mathbf{R}^n} \int_{\mathbf{R}^n} (h(x \pm y) - g(x \pm y)) \, dx \, dy \\ &= \int_{\mathbf{R}^n} \left(\int_{\mathbf{R}^n} (h(x \pm y) - g(x \pm y)) \, dx \right) dy \\ &= \int_{\mathbf{R}^n} \left(\int_{\mathbf{R}^n} (h(x) - g(x)) \, dx \right) dy \\ &= 0. \end{aligned}$$

This proves that

$$g(x \pm y) = h(x \pm y) \quad \text{for almost all } (x, y) \in \mathbf{R}^n \times \mathbf{R}^n,$$

so that

$$g(x \pm y) = f(x \pm y) = h(x \pm y) \quad \text{for almost all } (x, y) \in \mathbf{R}^n \times \mathbf{R}^n.$$

Therefore, we find that the function $f(x \pm y)$ is Lebesgue measurable on $\mathbf{R}^n \times \mathbf{R}^n$, since the functions $g(x \pm y)$ and $h(x \pm y)$ are Borel measurable on $\mathbf{R}^n \times \mathbf{R}^n$.

Step II: **End of Proof of Theorem 3.23**. The case where $1 \leq r < \infty$: It suffices to apply Theorem 3.21 with

$$\begin{aligned} &(X, \mu) := (\mathbf{R}^n, dx), \quad (Y, \nu) := (\mathbf{R}^n, dy), \\ &p := q, \; q := r, \; r := p, \end{aligned}$$

$$K(x,y) := f(x-y), \quad f(y) := g(y).$$

Indeed, since we have the formula

$$\int_{\mathbf{R}^n} |f(x-y)|^p \, dx = \int_{\mathbf{R}^n} |f(x-y)|^p \, dy = \|f\|_p^p,$$

it follows from an application of Theorem 3.21 with $M = N := \|f\|_p^p$ that

$$\|f * g\|_r \leq \left(\|f\|_p^p\right)^{1/r} \left(\|f\|_p^p\right)^{1-1/q} \|g\|_p = \|f\|_p \|g\|_q .$$

This proves the desired inequality (3.14) for $1 \leq r < \infty$.

The case where $r = \infty$: If $f \in L^p(\mathbf{R}^n)$ and $g \in L^q(\mathbf{R}^n)$ with $1/p + 1/q = 1$, then, by applying Hölder's inequality (Theorem 3.14) we obtain that

$$\begin{aligned} |(f * g)(x)| &= \left| \int_{\mathbf{R}^n} f(x-y) g(y) \, dy \right| \\ &\leq \left(\int_{\mathbf{R}^n} |f(x-y)|^p \, dy \right)^{1/p} \left(\int_{\mathbf{R}^n} |g(y)|^q dy \right)^{1/q} \\ &= \|f\|_p \|g\|_q . \end{aligned}$$

Hence we have the desired inequality (3.14) for $r = \infty$:

$$\|f * g\|_\infty \leq \|f\|_p \|g\|_q .$$

The proof of Theorem 3.23 is now complete. □

3.7.1 Approximations to the Identity

The next theorem 3.25 underlies one of the most important uses of convolutions. Before coming to it, we need a technical lemma:

Lemma 3.24. *Assume that $f \in L^p(\mathbf{R}^n)$ with $1 \leq p < \infty$. If we let*

$$f_x(y) = f(x+y) \quad \text{for } y \in \mathbf{R}^n,$$

then we have the assertion

$$f_x \longrightarrow f \quad \text{in } L^p(\mathbf{R}^n) \text{ as } x \to 0.$$

Proof. (1) If g is a continuous function with compact support, then we have the assertion

$$g_x \longrightarrow g \quad \text{uniformly as } x \to 0,$$

and so

$$\|g_x - g\|_p \longrightarrow 0 \quad \text{as } x \to 0.$$

(2) Let $f \in L^p(\mathbf{R}^n)$. For each $\varepsilon > 0$, we can find a continuous function g with compact support such that

$$\|f - g\|_p < \frac{\varepsilon}{2}.$$

Then we have the inequality

$$\|f_x - g_x\|_p = \|f - g\|_p < \frac{\varepsilon}{2},$$

and so

$$\|f_x - f\|_p \leq \|f_x - g_x\|_p + \|g_x - g\|_p + \|g - f\|_p < \varepsilon + \|g_x - g\|_p .$$

This implies that

$$\limsup_{x \to 0} \|f_x - f\|_p \leq \varepsilon.$$

The proof of Lemma 3.24 is complete. □

Remark 3.1. Lemma 3.24 is not true for $p = \infty$. Indeed, the assertion

$$\|f_x - f\|_\infty = \sup_{y \in \mathbf{R}^n} |f(x+y) - f(y)| \longrightarrow 0 \quad \text{as } x \to 0$$

implies the uniform continuity of f on $\mathbf{R}^n$.

Theorem 3.25. *Let $\varphi(x)$ be a function in $L^1(\mathbf{R}^n)$ such that*

$$\int_{\mathbf{R}^n} \varphi(x)\, dx = 1.$$

For any $\varepsilon > 0$, we define

$$\varphi_\varepsilon(x) = \frac{1}{\varepsilon^n} \varphi\left(\frac{x}{\varepsilon}\right).$$

Then we have the following two assertions (i) and (ii):

(i) If $f \in L^p(\mathbf{R}^n)$ with $1 \leq p < \infty$, then it follows that

$$f * \varphi_\varepsilon \longrightarrow f \quad \text{in } L^p(\mathbf{R}^n) \text{ as } \varepsilon \downarrow 0.$$

(ii) If $f \in L^\infty(\mathbf{R}^n)$ is uniformly continuous, then it follows that

$$f * \varphi_\varepsilon \longrightarrow f \quad \text{in } L^\infty(\mathbf{R}^n) \text{ as } \varepsilon \downarrow 0.$$

Proof. (i) Since we have the formula

$$\int_{\mathbf{R}^n} \varphi_\varepsilon(x)\,dx = \int_{\mathbf{R}^n} \varphi(x)\,dx = 1,$$

it follows that

$$\begin{aligned}(f * \varphi_\varepsilon)(x) - f(x) &= \int_{\mathbf{R}^n} (f(x-y) - f(x))\,\varphi_\varepsilon(y)\,dy \\ &= \int_{\mathbf{R}^n} (f(x-\varepsilon z) - f(x))\,\varphi(z)\,dz.\end{aligned}$$

Hence, by applying Minkowski's inequality for integrals (Theorem 3.16) we obtain that

$$\|f * \varphi_\varepsilon - f\|_p \leq \int_{\mathbf{R}^n} \|f_{-\varepsilon z} - f\|_p\,|\varphi(z)|\,dz.$$

However, we have the inequality

$$\|f_{-\varepsilon z} - f\|_p \leq \|f_{-\varepsilon z}\|_p + \|f\|_p = 2\,\|f\|_p\,,$$

and so

$$\|f_{-\varepsilon z} - f\|_p\,|\varphi(z)| \leq 2\,\|f\|_p\,|\varphi(z)|,$$

with

$$\varphi \in L^1(\mathbf{R}^n).$$

Furthermore, Lemma 3.24 tells us that

$$\lim_{\varepsilon \downarrow 0} \|f_{-\varepsilon z} - f\|_p = 0, \quad z \in \mathbf{R}^n.$$

Therefore, it follows from an application of the Lebesgue dominated convergence theorem (Theorem 3.8) that

$$\limsup_{\varepsilon \downarrow 0} \|f * \varphi_\varepsilon - f\|_p \leq \lim_{\varepsilon \downarrow 0} \int_{\mathbf{R}^n} \|f_{-\varepsilon z} - f\|_p\,|\varphi(z)|\,dz = 0.$$

(ii) Assume that $f \in L^\infty(\mathbf{R}^n)$ is uniformly continuous. Since $\varphi(x)$ is integrable on $\mathbf{R}^n$, it follows that, for each $\delta > 0$, there exists a compact subset W of $\mathbf{R}^n$ such that

$$\int_{\mathbf{R}^n \setminus W} |\varphi(z)|\,dz < \delta.$$

Then we have the inequality

$$\sup_{x \in \mathbf{R}^n} |f * \varphi_\varepsilon(x) - f(x)|$$

$$
\begin{aligned}
&\le \sup_{x\in\mathbf{R}^n}\left(\int_{\mathbf{R}^n} |f(x-\varepsilon y)-f(x)|\,|\varphi(y)|\,dy\right)\\
&= \sup_{x\in\mathbf{R}^n}\left(\int_{W} |f(x-\varepsilon y)-f(x)|\,|\varphi(y)|\,dy\right.\\
&\quad \left.+\int_{\mathbf{R}^n\setminus W} |f(x-\varepsilon y)-f(x)|\,|\varphi(y)|\,dy\right)\\
&\le \sup_{x\in\mathbf{R}^n}\left(f(x-\varepsilon y)-f(x)|\right)\cdot\int_W |\varphi(y)|\,dy + 2\,\|f\|_\infty\cdot\int_{\mathbf{R}^n\setminus W}|\varphi(y)|\,dy\\
&\le \sup_{x\in\mathbf{R}^n}\left(|f(x-\varepsilon y)-f(x)|\right)\|\varphi\|_1 + 2\delta\,\|f\|_\infty .
\end{aligned}
$$

However, we have the assertion

$$
\lim_{\varepsilon\downarrow 0}\left[\sup_{\substack{x\in\mathbf{R}^n\\ x\in W}} |f(x-\varepsilon y)-f(x)|\right] = 0,
$$

since f is uniformly continuous and since W is compact.

Summing up, we obtain that

$$
\limsup_{\varepsilon\downarrow 0} \|f*\varphi_\varepsilon - f\|_\infty \le 2\delta\,\|f\|_\infty .
$$

This proves part (ii), since δ is arbitrary.

The proof of Theorem 3.25 is complete. □

The family of functions $\{\varphi_\varepsilon\}$ defined above is called an *approximation to the identity*. What makes these useful is that, by choosing φ appropriately we can obtain the functions $f*\varphi_\varepsilon$ to have nice properties.

Example 3.2 (the heat kernel)**.** Let

$$
K(x) := \frac{1}{(4\pi)^{n/2}}\, e^{-\frac{|x|^2}{4t}}, \quad t>0,
$$

and

$$
K_t(x) = \frac{1}{(\sqrt{t}\,)^n}\, K\left(\frac{x}{\sqrt{t}}\right) = \frac{1}{(4\pi t)^{n/2}}\, e^{-\frac{|x|^2}{4t}}, \quad t>0.
$$

We define

$$
\begin{aligned}
u(x,t) &= f*K_t(x)\\
&= \frac{1}{(4\pi t)^{n/2}}\int_{\mathbf{R}^n} e^{-\frac{|x-y|^2}{4t}}\, f(y)\,dy, \quad t>0.
\end{aligned}
$$

Then we have the following two assertions (i) and (ii):

(i) If $f \in L^p(\mathbf{R}^n)$ with $1 \le p < \infty$, then it follows that

$$u(\cdot, t) \longrightarrow f \quad \text{in } L^p(\mathbf{R}^n) \text{ as } t \downarrow 0.$$

(ii) If f is bounded and continuous on $\mathbf{R}^n$, then it follows that $u(x,t)$ is continuous on $\mathbf{R}^n \times [0, \infty)$ and further that

$$u(\cdot, t) \longrightarrow f \quad \text{in } L^\infty(\mathbf{R}^n) \text{ as } t \downarrow 0.$$

Proof. (i) Note that

$$\int_{\mathbf{R}^n} K(x)\,dx = \frac{1}{(4\pi)^{n/2}} \int_{\mathbf{R}^n} e^{-\frac{|x|^2}{4t}}\,dx = 1.$$

Therefore, part (i) follows by applying Theorem 3.25 with

$$\varphi(x) := K(x),$$
$$\varepsilon := \sqrt{t}.$$

(ii) Similarly, part (ii) follows by applying Theorem 3.25 with

$$\varphi(x) := K(x),$$
$$\varepsilon := \sqrt{t}.$$

The proof of Example 3.2 is complete. □

Remark 3.2. It should be emphasized that the function $u(x,t) = f * K_t(x)$ is a solution of the following initial value problem for the heat equation:

$$\begin{cases} \frac{\partial u}{\partial t}(x,t) - \Delta u(x,t) = 0 & \text{for all } x \in \mathbf{R}^n \text{ and } t > 0, \\ u(x,0) = f(x) & \text{for all } x \in \mathbf{R}^n. \end{cases}$$

3.7.2 Friedrichs' Mollifiers

Let $\rho(x)$ be a non-negative, bell-shaped C^∞ function on $\mathbf{R}^n$ satisfying the following two conditions:

$$\operatorname{supp} \rho = \{x \in \mathbf{R}^n : |x| \le 1\}. \tag{3.16a}$$

$$\int_{\mathbf{R}^n} \rho(x)\,dx = 1. \tag{3.16b}$$

For example, we may take

$$\rho(x) = \begin{cases} k \exp[-1/(1 - |x|^2)] & \text{if } |x| < 1, \\ 0 & \text{if } |x| \ge 1, \end{cases}$$

where the constant factor k is so chosen that condition (3.16) is satisfied.

For each $\varepsilon > 0$, we define

$$\rho_\varepsilon(x) = \frac{1}{\varepsilon^n}\rho\left(\frac{x}{\varepsilon}\right),$$

then $\rho_\varepsilon(x)$ is a non-negative, C^∞ function on $\mathbf{R}^n$, and satisfies the conditions

$$\operatorname{supp}\rho_\varepsilon = \{x \in \mathbf{R}^n : |x| \leq \varepsilon\}; \tag{3.17a}$$

$$\int_{\mathbf{R}^n} \rho_\varepsilon(x)\,dx = 1. \tag{3.17b}$$

The functions $\{\rho_\varepsilon\}$ are called *Friedrichs' mollifiers* (see Figure 3.1).

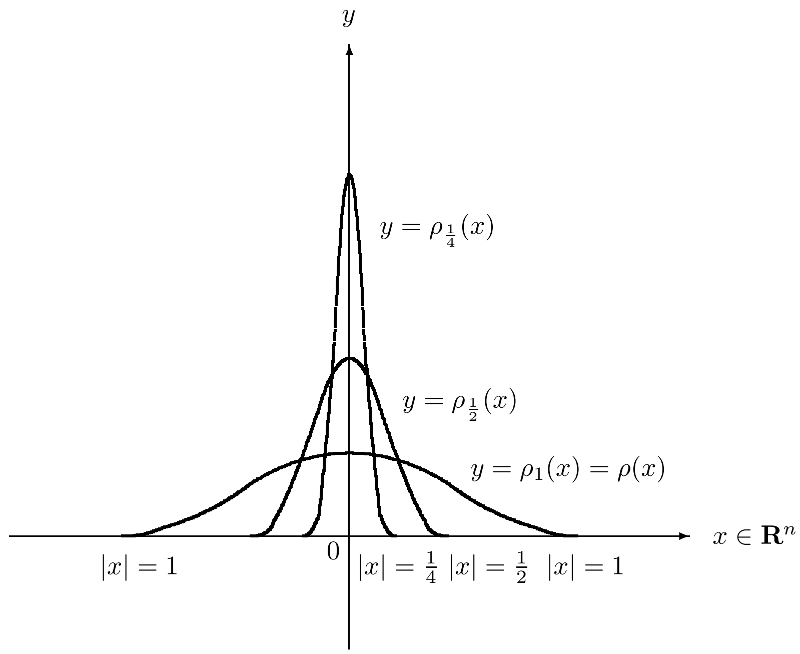

Fig. 3.1. Friedrichs' mollifiers $\{\rho_\varepsilon\}$

The next local version of Theorem 3.25 shows how mollifiers can be used to approximate a function by smooth functions in a domain of $\mathbf{R}^n$:

Theorem 3.26. *Let Ω be an open subset of $\mathbf{R}^n$. Then we have the following two assertions:*

(i) If $u \in L^p(\Omega)$ with $1 \leq p < \infty$ and vanishes outside a compact

subset K of Ω, then it follows that $\rho_\varepsilon * u \in C_0^\infty(\Omega)$ provided that $\varepsilon < \operatorname{dist}(K, \partial\Omega)$, and further that $\rho_\varepsilon * u \to u$ in $L^p(\Omega)$ as $\varepsilon \downarrow 0$.

(ii) If $u \in C_0^m(\Omega)$ with $0 \leq m < \infty$, then it follows that $\rho_\varepsilon * u \in C_0^\infty(\Omega)$ provided that $\varepsilon < \operatorname{dist}(\operatorname{supp} u, \partial\Omega)$, and further that $\rho_\varepsilon * u \to u$ in $C_0^m(\Omega)$ as $\varepsilon \downarrow 0$.

Here

$$\operatorname{dist}(K, \partial\Omega) = \inf\{|x-y| : x \in K,\ y \in \partial\Omega\}.$$

The functions $\rho_\varepsilon * u$ are called *regularizations* of the function u.

Corollary 3.27. *The space $C_0^\infty(\Omega)$ is dense in $L^p(\Omega)$ for each $1 \leq p < \infty$.*

Proof. Indeed, Corollary 3.27 is an immediate consequence of part (i) of Theorem 3.26, since L^p functions with compact support are dense in $L^p(\Omega)$. □

The next result gives another useful construction of smooth functions that vanish outside compact sets:

Corollary 3.28. *Let K be a compact subset of $\mathbf{R}^n$. If Ω is an open subset of $\mathbf{R}^n$ such that $K \subset \Omega$, then there exists a function $f \in C_0^\infty(\Omega)$ such that*

$$\begin{aligned} &0 \leq f(x) \leq 1 \quad \text{in } \Omega, \\ &f(x) = 1 \quad \text{on } K. \end{aligned}$$

Proof. Let

$$\delta = \operatorname{dist}(K, \partial\Omega),$$

and define a relatively compact subset U of Ω, containing K, as follows:

$$U = \left\{ x \in \Omega : |x-y| < \frac{\delta}{2} \text{ for some } y \in K \right\}.$$

Then it is easy to verify that the function

$$f(x) = \rho_\varepsilon * \chi_U(x) = \frac{1}{\varepsilon^n} \int_U \rho\left(\frac{x-y}{\varepsilon}\right) dy, \quad 0 < \varepsilon < \frac{\delta}{2},$$

satisfies all the conditions. □

3.8 Distribution Functions

In this section we are interested in giving a concise expression for the L^p-norm of a function. Let $(\Omega, \mathcal{B}, \mu)$ be a measure space. Let $f(x)$ be a measurable function defined on Ω. If $\tau > 0$, we define the *distribution function* of f by the formula

$$\sigma(\tau) := \mu\left(\{x \in \Omega : |f(x)| > \tau\}\right).$$

The next theorem asserts that a measurable function is L^p-integrable over Ω with respect to μ if and only if its distribution function is integrable over the interval $[0, \infty)$ with respect to $p\tau^{p-1}\,d\tau = d\tau^p$:

Theorem 3.29. *Let $f(x)$ be a measurable function defined on Ω, and let $1 \leq p < \infty$. Then $f \in L^p(\Omega)$ if and only if it satisfies the condition*

$$\int_0^\infty \tau^{p-1}\sigma(\tau)\,d\tau < \infty. \tag{3.18}$$

Moreover, in this case we have the formula

$$\int_\Omega |f(x)|^p \, d\mu = p\int_0^\infty \tau^{p-1}\sigma(\tau)\,d\tau = \int_0^\infty \sigma(\tau)\,d\tau^p. \tag{3.19}$$

Proof. First, it should be noticed that the function $\sigma(\tau)$ is monotone decreasing and right-continuous in τ, for $\tau > 0$.

The proof of Theorem 3.29 is divided into two steps.

Step (I): If condition (3.20) holds true, then we have the following claim:

Claim 3.2. $\tau^p\,\sigma(\tau) \to 0$ as $\tau \to 0$ and $\tau \to \infty$.

Proof. (a) Our proof is based on a reduction to absurdity. Assume, to the contrary, that we have, as $\tau \to 0$,

$$\tau^p\,\sigma(\tau) \nrightarrow 0.$$

Then we can find a sequence $\{\tau_j\}$ and a positive constant δ such that

$$\tau_j^p\,\sigma(\tau_j) \geq \delta, \quad 0 < \tau_{j+1} < \frac{1}{2}\tau_j.$$

Hence it follows that

$$\begin{aligned} p\int_{\tau_{j+1}}^{\tau_j} \tau^{p-1}\sigma(\tau)\,d\tau &\geq p\sigma(\tau_j)\int_{\tau_{j+1}}^{\tau_j} \tau^{p-1}\,d\tau = \sigma(\tau_j)\left(\tau_j^p - \tau_{j+1}^p\right) \\ &\geq \frac{\delta}{\tau_j^p}\left(\tau_j^p - \tau_{j+1}^p\right) = \delta\left(1 - \left(\frac{\tau_{j+1}}{\tau_j}\right)^p\right) \end{aligned}$$

$$\geq \delta \left(1 - \frac{1}{2^p}\right).$$

This contradict condition (3.20), since we have

$$p \int_0^\infty \tau^{p-1} \sigma(\tau)\, d\tau = \infty, \quad 1 \leq p < \infty.$$

(b) Our proof is based on a reduction to absurdity. Assume, to the contrary, that we have, as $\tau \to \infty$,

$$\tau^p\, \sigma(\tau) \nrightarrow 0.$$

Then we can find a sequence $\{\tau_j\}$ and a positive constant δ such that

$$\tau_{j+1}^p\, \sigma(\tau_{j+1}) \geq \delta, \quad 0 < \tau_j < \frac{1}{2}\tau_{j+1}.$$

Hence it follows that

$$\begin{aligned} p \int_{\tau_j}^{\tau_{j+1}} \tau^{p-1} \sigma(\tau)\, d\tau \geq p\sigma(\tau_{j+1}) \int_{\tau_j}^{\tau_{j+1}} \tau^{p-1}\, d\tau &= \sigma(\tau_{j+1}) \left(\tau_{j+1}^p - \tau_j^p\right) \\ &\geq \frac{\delta}{\tau_{j+1}^p} \left(\tau_{j+1}^p - \tau_j^p\right) = \delta \left(1 - \left(\frac{\tau_j}{\tau_{j+1}}\right)^p\right) \\ &\geq \delta \left(1 - \frac{1}{2^p}\right). \end{aligned}$$

This contradict condition (3.18), since we have the assertion

$$p \int_0^\infty \tau^{p-1} \sigma(\tau)\, d\tau = \infty, \quad 1 \leq p < \infty.$$

The proof of Claim 3.2 is complete. □

In view of Claim 3.2, by integration by parts we obtain that

$$\begin{aligned} p \int_0^\infty \tau^{p-1} \sigma(\tau)\, d\tau &= \left[\tau^p\, \sigma(\tau)\right]_0^\infty - \int_0^\infty \tau^p\, d\sigma(\tau) \\ &= -\int_0^\infty \tau^p\, d\sigma(\tau) = \int_\Omega |f(x)|^p\, d\mu. \end{aligned}$$

This proves the desired formula (3.19).

Step (II): Conversely, we assume that $f \in L^p(\Omega)$. Then we have, for all $\tau > 0$,

$$\tau^p\, \sigma(\tau) \leq \int_{E(\tau)} |f(x)|^p\, d\mu, \tag{3.20}$$

where

$$E(\tau) := \{x \in \Omega : |f(x)| > \tau\}.$$

However, since $f \in L^p(\Omega)$ it follows that

$$\begin{aligned}\int_\Omega |f(x)|^p \, d\mu &= \lim_{\tau\to\infty} \int_{\Omega\setminus E(\tau)} |f(x)|^p \, d\mu \\ &= \int_\Omega |f(x)|^p \, d\mu - \lim_{\tau\to\infty} \int_{E(\tau)} |f(x)|^p \, d\mu,\end{aligned}$$

so that

$$\lim_{\tau\to\infty} \int_{E(\tau)} |f(x)|^p \, d\mu = 0.$$

Hence we have, by inequality (3.20),

$$\lim_{\tau\to\infty} \tau^p \, \sigma(\tau) \leq \lim_{\tau\to\infty} \int_{E(\tau)} |f(x)|^p \, d\mu = 0.$$

Therefore, by integration by parts it follows that

$$\begin{aligned}\int_\Omega |f(x)|^p \, d\mu &= -\lim_{\varepsilon\downarrow 0} \int_\varepsilon^\infty \tau^p \, d\sigma(\tau) \\ &= \lim_{\varepsilon\downarrow 0} \left\{ -\left[\tau^p \, \sigma(\tau)\right]_\varepsilon^\infty + \int_\varepsilon^\infty p\tau^{p-1} \, \sigma(\tau) \, d\tau \right\} \\ &= \lim_{\varepsilon\downarrow 0} \left\{ \varepsilon^p \, \sigma(\varepsilon) + \int_\varepsilon^\infty p\tau^{p-1} \, \sigma(\tau) \, d\tau \right\} \\ &\geq p \int_0^\infty \tau^{p-1} \, \sigma(\tau) \, d\tau.\end{aligned}$$

This proves the desired condition (3.18), since $1 \leq p < \infty$.

The proof of Theorem 3.29 is now complete. □

3.9 Marcinkiewicz's Interpolation Theorem

In this section we prove the Marcinkiewicz interpolation theorem (Theorem 3.30) for which we need some terminology. Let $(\Omega, \mathcal{B}, \phi)$ and $(\widehat{\Omega}, \widehat{\mathcal{B}}, \widehat{\phi})$ be two measure spaces, and the norms of $L^p(\Omega)$ and $L^p(\widehat{\Omega})$ will be denoted by the same notation $\|\cdot\|_p$, but it is clear from the context which is which.

(i) Let T be a not necessarily linear mapping of $L^p(\Omega)$ into $L^q(\widehat{\Omega})$ with $1 \leq p \leq \infty$ and $1 \leq q \leq \infty$. The mapping T is said to be *of type* (p, q) if there exists a positive constant K such that we have, for all $f \in L^p(\Omega)$,

$$\|Tf\|_q \leq K \, \|f\|_p \, .$$

(ii) Let T be a not necessarily linear mapping of $L^p(\Omega)$ into a set of measurable functions defined on $\widehat{\Omega}$ with $1 \leq p \leq \infty$. In the case where $q < \infty$, the mapping T is said to be *of weak type* (p, q) if there exists a positive constant K such that we have, for all $f \in L^p(\Omega)$ and all $s > 0$,

$$\widehat{\phi}\left(\left\{y \in \widehat{\Omega} : |(Tf)(y)| > s\right\}\right) \leq \left(\frac{K \|f\|_p}{s}\right)^q .$$

In the case where $q = \infty$ a mapping of type (p, ∞) is said to be *of weak type* (p, ∞).

It is easy to see that a mapping of type (p, q) is of weak type (p, q).

We denote the totality of functions which are expressed as a sum of a function in $L^p(\Omega)$ and a function in $L^q(\Omega)$ by $L^p(\Omega) + L^q(\Omega)$, that is,

$$L^p(\Omega) + L^q(\Omega) := \{f + g : f \in L^p(\Omega),\ g \in L^q(\Omega)\} .$$

Let $1 \leq p < q < r < \infty$. If $f(x) \in L^q(\Omega)$, we let

$$f_1(x) := \begin{cases} 0 & \text{if } |f(x)| \leq 1, \\ f(x) & \text{if } |f(x)| > 1, \end{cases}$$

and

$$f_2(x) := \begin{cases} f(x) & \text{if } |f(x)| \leq 1, \\ 0 & \text{if } |f(x)| > 1. \end{cases}$$

Then it follows that $f(x)$ can be written in the form

$$f(x) = f_1(x) + f_2(x),$$

with

$$f_1(x) \in L^p(\Omega), \qquad f_2(x) \in L^r(\Omega),$$

Indeed, it suffices to note that

$$|f_1(x)|^p \leq |f(x)|^q,$$
$$|f_2(x)|^r \leq |f(x)|^q.$$

Therefore, for $1 \leq p < q < r < \infty$, we have the inclusion

$$L^q(\Omega) \subset L^p(\Omega) + L^r(\Omega).$$

Marcinkiewicz's interpolation theorem reads as follows:

Theorem 3.30 (Marcinkiewicz)**.** *Let $1 \le p < q < r$ and let T be a linear mapping of the space $L^p(\Omega) + L^r(\Omega)$ into a set of measurable functions defined on $\widehat{\Omega}$. If T is of weak type (p, p) and also of weak type (r, r), then it is of type (q, q).*

Proof. If $f(x)$ is a measurable function defined on Ω and if $\tau > 0$, we let

$$(f)_r := \phi(\{x \in \Omega : |f(x)| > \tau\}).$$

We shall use the same notation for a measurable function defined on $\widehat{\Omega}$ with ϕ replaced by $\widehat{\phi}$.

First, since T is of weak type (p, p) and of weak type (r, r), we can find two constants $K_1 > 0$ and $K_2 > 0$ such that we have, for any $\tau > 0$ and for any $f \in L^p(\Omega)$ or for any $f \in L^r(\Omega)$,

$$(Tf)_\tau \le \left(\frac{K_1 \|f\|_p}{\tau}\right)^p \quad \text{or} \quad (Tf)_\tau \le \left(\frac{K_2 \|f\|_r}{\tau}\right)^r, \tag{3.21}$$

respectively. If $f(x)$ is a function in $L^q(\Omega)$, we let

$$f_1(x) := \begin{cases} f(x) & \text{if } |f(x)| \le \tau, \\ \tau \operatorname{sign} f(x) & \text{if } |f(x)| > \tau, \end{cases}$$

and

$$f_2(x) := f(x) - f_1(x).$$

Then, since $Tf(x) = Tf_1(x) + Tf_2(x)$, we have the inclusion

$$\{y : |(Tf)(y)| > \tau\} \subset \{y : |(Tf_1)(y)| > \tau/2\} \cup \{y : |(Tf_2)(y)| > \tau/2\}\,.$$

Hence, by condition (3.21) with $\tau := \tau/2$ it follows that

$$\begin{aligned}(Tf)_\tau &\le (Tf_1)_{\tau/2} + (Tf_2)_{\tau/2} \\ &\le \left(\frac{2K_1 \|f_1\|_r}{\tau}\right)^r + \left(\frac{2K_2 \|f_2\|_p}{\tau}\right)^p.\end{aligned} \tag{3.22}$$

However, we remark that

$$(f_1)_\sigma = \begin{cases} (f)_\sigma & \text{for } \sigma < \tau, \\ 0 & \text{for } \sigma \ge \tau, \end{cases}$$

and that

$$(f_2)_\sigma = (f)_{\sigma+\tau} \quad \text{for } \sigma \ge 0.$$

Thus, it follows from an application of Theorem 3.29 that

$$\|f_1\|_r^r = r \int_0^\infty \sigma^{r-1} (f_1)_\sigma \, d\sigma = r \int_0^\tau \sigma^{r-1} (f)_\sigma \, d\sigma, \tag{3.23}$$

and further that

$$\|f_2\|_p^p = p \int_0^\infty \sigma^{p-1} (f_2)_\sigma \, d\sigma \tag{3.24}$$
$$= p \int_0^\infty \sigma^{p-1} (f)_{\sigma+\tau} \, d\sigma = p \int_\tau^\infty (\sigma - \tau)^{p-1} (f)_\sigma \, d\sigma.$$

By combining inequality (3.22) and two formulas (3.23) and (3.24), we obtain that

$$(Tf)_\tau \leq \left(\frac{2K_1}{\tau} \right)^r r \int_0^\tau \sigma^{r-1} (f)_\sigma \, d\sigma + \left(\frac{2K_2}{\tau} \right)^p p \int_\tau^\infty (\sigma - \tau)^{p-1} (f)_\sigma \, d\sigma.$$

Therefore, we have, by this inequality and Theorem 3.29,

$$\begin{aligned}
&\|Tf\|_q^q \\
&= q \int_0^\infty \tau^{q-1} (Tf)_\tau \, d\tau \\
&\leq q \int_0^\infty \tau^{q-1} \left(\frac{2K_1}{\tau} \right)^r r \left(\int_0^\tau \sigma^{r-1} (f)_\sigma \, d\sigma \right) d\tau \\
&\quad + q \int_0^\infty \tau^{q-1} \left(\frac{2K_2}{\tau} \right)^p p \left(\int_\tau^\infty (\sigma - \tau)^{p-1} (f)_\sigma \, d\sigma \right) d\tau \\
&= qr(2K_1)^r \int_0^\infty \tau^{q-r-1} \left(\int_0^\tau \sigma^{r-1} (f)_\sigma \, d\sigma \right) d\tau \\
&\quad + pq(2K_2)^p \int_0^\infty \tau^{q-p-1} \left(\int_\tau^\infty (\sigma - \tau)^{p-1} (f)_\sigma \, d\sigma \right) d\tau \\
&= qr(2K_1)^r \int_0^\infty \left(\int_\sigma^\infty \tau^{q-r-1} \, d\tau \right) \sigma^{r-1} (f)_\sigma \, d\sigma \\
&\quad + pq(2K_2)^p \int_0^\infty \left(\int_0^\sigma \tau^{q-p-1} (\sigma - \tau)^{p-1} \, d\tau \right) (f)_\sigma \, d\sigma \\
&= \frac{qr(2K_1)^r}{r-q} \int_0^\infty \sigma^{q-1} (f)_\sigma \, d\sigma + pq(2K_2)^p B(q-p, p) \int_0^\infty \sigma^{q-1} (f)_\sigma \, d\sigma \\
&= \left\{ \frac{r(2K_1)^r}{r-q} + p(2K_2)^p B(q-p, p) \right\} \|f\|_q^q .
\end{aligned}$$

This proves that the operator T is of type (q, q), with

$$K := \left(\frac{r(2K_1)^r}{r-q} + p(2K_2)^p B(q-p, p) \right)^{1/q}.$$

The proof of Theorem 3.30 is complete. □

3.10 Riesz Potentials

In this section, as an application of Marcinkiewicz's interpolation theorem (Theorem 3.30) we study Riesz potentials in the classical potential theory.

Let Δ be the usual Laplacian

$$\Delta = \sum_{j=i}^{n} \frac{\partial^2}{\partial x_j^2}.$$

Then we have, for every $f \in \mathcal{S}(\mathbf{R}^n)$,

$$\widehat{(-\Delta f)}(\xi) = |\xi|^2 \widehat{f}(\xi),$$

and also

$$\mathcal{F}((-\Delta)^{\alpha/2} f)(\xi) = |\xi|^\alpha \widehat{f}(\xi), \quad \alpha > 0.$$

If $0 < \alpha < n$, we define a *Riesz potential* I_α by the formula

$$I_\alpha(f)(x) = (-\Delta)^{-\alpha/2}(f)(x) = \frac{1}{\gamma(\alpha)} \int_{\mathbf{R}^n} |x-y|^{\alpha-n} f(y)\, dy,$$

where

$$\gamma(\alpha) = \frac{\pi^{n/2}\, 2^\alpha\, \Gamma(\alpha/2)}{\Gamma\left((n-\alpha)/2\right)}.$$

The function

$$\frac{1}{\gamma(\alpha)} |x|^{\alpha-n}$$

is called a *Riesz kernel*.

The purpose of this secion is to prove the following:

Theorem 3.31. *Let $0 < \alpha < n$, $1 < p < q$ and $1/q = 1/p - \alpha/n$. Then we have, for all $f \in L^p(\Omega)$,*

$$\|I_\alpha(f)\|_q \leq A_{p,q} \|f\|_p, \tag{3.25}$$

with some constant $A_{p,q} > 0$.

Proof. We have only to prove the theorem for the integral kernel

$$K(x) = |x|^{\alpha-n}.$$

We shall show that the mapping

$$f \longmapsto K * f$$

is of weak type (p, q). Namely, we have, for all $\lambda > 0$,

$$m\{x : |K * f(x)| > \lambda\} \leq \left(A_{p,q}\,\frac{\|f\|_p}{\lambda}\right)^q, \tag{3.26}$$

where m is the Lebesgue measure on $\mathbf{R}^n$.

To do so, it suffices to prove that

$$m\{x : |K * g(x)| > \lambda\} \leq \left(\frac{A_{p,q}}{\lambda}\right)^q \quad \text{for all } \|g\|_p = 1. \tag{3.27}$$

Indeed, by letting

$$g(x) = \frac{f(x)}{\|f\|_p},$$

we obtain from inequality (3.27) that

$$m\left\{x : |K * f(x)| > \lambda\,\|f\|_p\right\} = m\{x : |K * g(x)| > \lambda\} \leq \left(\frac{A_{p,q}}{\lambda}\right)^q.$$

This proves the desired inequality (3.26) if we take

$$\lambda := \frac{\lambda}{\|f\|_p}.$$

In order to prove inequality (3.27), we let, for a positive constant μ,

$$K(x) := K_1(x) + K_\infty(x),$$

where

$$K_1(x) = \begin{cases} K(x) & \text{for } |x| \leq \mu, \\ 0 & \text{for } |x| > \mu, \end{cases}$$

and

$$K_2(x) = \begin{cases} 0 & \text{for } |x| \leq \mu, \\ K(x) & \text{for } |x| > \mu. \end{cases}$$

The constant μ will be chosen later on.

Since we have the formula

$$K * g = K_1 * g + K_\infty * g,$$

it follows that

$$\begin{aligned} &m\{x : |K * g(x)| > 2\lambda\} \\ &\leq m\{x : |K_1 * g(x)| > \lambda\} + m\{x : |K_\infty * g(x)| > \lambda\}. \end{aligned} \tag{3.28}$$

(1) First, we have the inequality

$$\begin{aligned} & m\{x : |K_1 * g(x)| > \lambda\} \\ & \leq \int_{\mathbf{R}^n} \left(\frac{|K_1 * g(x)|}{\lambda} \right)^p dx = \frac{\| |K_1 * g(x)| \|_p^p}{\lambda^p} \\ & \leq \frac{1}{\lambda^p} \|K_1\|_1^p \|g\|_p^p = \frac{1}{\lambda^p} \|K_1\|_1^p . \end{aligned} \tag{3.29}$$

However, it follows that

$$\begin{aligned} \|K_1\|_1 &= \int_{|x| \leq \mu} |x|^{\alpha - n} \, dx = \int_{\Sigma_{n-1}} \int_0^\mu r^{\alpha - n} r^{n-1} \, dr \, d\sigma \\ &= \frac{\omega_n}{\alpha} \mu^\alpha , \end{aligned} \tag{3.30}$$

where ω_n is the surface area of the unit sphere Σ_{n-1} in $\mathbf{R}^n$

$$\omega_n := |\Sigma_{n-1}| = \frac{2\pi^{n/2}}{\Gamma(n/2)}.$$

Therefore, by combining inequality (3.29) and formula (3.30) we obtain that

$$m\{x : |K_1 * g(x)| > \lambda\} \leq C_1^p \frac{\mu^{\alpha p}}{\lambda^p}, \quad C_1 := \frac{\omega_n}{\alpha}. \tag{3.31}$$

(2) Secondly, by using Young's inequality (Theorem 3.23) we have the inequality

$$\|K_\infty * g\|_\infty \leq \|K_\infty\|_{p'} \|g\|_p = \|K_\infty\|_{p'} , \quad p' = \frac{p}{p-1}. \tag{3.32}$$

However, it follows that

$$\begin{aligned} & \|K_\infty\|_{p'} \\ &= \left(\int_{|x| \geq \mu} \left(|x|^{\alpha - n} \right)^{p'} dx \right)^{1/p'} \\ &= \left(\int_{\Sigma_{n-1}} \int_\mu^\infty r^{(\alpha - n)p'} r^{n-1} \, dr \, d\sigma \right)^{1/p'} \\ &= \left(\omega_n \left[\frac{1}{(\alpha - n)p' + n} r^{(\alpha - n)p' + n} \right]_\mu^\infty \right)^{1/p'} \\ &= \omega_n^{1/p'} \left(\frac{\mu^{n-(n-\alpha)p'}}{(n-\alpha)p' - n} \right)^{1/p'} \end{aligned} \tag{3.33}$$

$$= C_2\,\mu^{(\alpha-n)+n(1-1/p)} = C_2\,\mu^{-n/q}, \quad C_2 := \left(\omega_n\,\frac{p-1}{p(n-\alpha)}\right)^{1/p'}.$$

If we choose the constant μ as

$$\mu := \left(\frac{C_2}{\lambda}\right)^{q/n},$$

or equivalently

$$\lambda = C_2\,\mu^{-n/q},$$

then it follows from inequalities (3.32) and (3.33) that

$$\|K_\infty * g\|_\infty \le \|K_\infty\|_{p'} = \lambda.$$

This proves that

$$m\,\{x : |K_\infty * g(x)| > \lambda\} = 0. \tag{3.34}$$

(3) By combining inequalities (3.28), (3.31) and assertion (3.34), we have proved that

$$m\,\{x : |K * g(x)| > 2\lambda\} \le C_1^p \left(\frac{1}{\lambda}\right)^q,$$

so that (by replacing 2λ by λ)

$$m\,\{x : |K * g(x)| > \lambda\} \le (A_{p,q}\lambda)^q, \quad A_{p,q} := 2\,C_1^{p/q}.$$

Therefore, the desired inequality (3.27) (and hence (3.25)) follows from an application of Marcinkiewicz's theorem 3.30. □

We can prove a more precise estimate for Riesz potentials with $\alpha := n\mu$ (see [45, Chapter 1, Lemma 1.34], [33, Chapter 7, Lemma 7.12]):

Theorem 3.32. *If $0 < \mu \le 1$, we define the Riesz potential by the formula*

$$(I_{n\mu} f)(x) := \int_\Omega |x-y|^{n(\mu-1)}\, f(y)\, dy.$$

Then, for any p, q satisfying the conditions

$$1 \le p \le q \le \infty,$$
$$0 \le \delta = \delta(p,q) := \frac{1}{p} - \frac{1}{q} < \mu,$$

the operator $I_{n\mu}$ *maps* $L^p(\Omega)$ *continuously into* $L^q(\Omega)$*. More precisely, we have, for all* $f \in L^p(\Omega)$*,*

$$\|I_{n\mu} f\|_{L^q(\Omega)} \le \left(\frac{1-\delta}{\mu-\delta}\right)^{1-\delta} V_n^{1-\mu}\, |\Omega|^{\mu-\delta}\, \|f\|_{L^p(\Omega)}\,.$$

Here V_n *is the volume of the unit ball in* $\mathbf{R}^n$

$$V_n = \frac{\omega_n}{n} = \frac{\pi^{n/2}}{\Gamma(n/2+1)}.$$

3.11 Notes and Comments

For more thorough treatments of the subject in this chapter, the reader might be referred to Duoandikoetxea [23], Folland [29], Friedman [30], Malý–Ziemer [45], Rudin [60], Stein–Shakarchi [71] and Torchinsky [91].

4
Elements of Real Analysis

This chapter is devoted to the precise definitions and statements of real analytic tools such as BMO and VMO functions, the Calderón–Zygmund decomposition (Theorem 4.7), the John–Nirenberg inequality (Theorem 4.10), the Hardy–Littlewood maximal function (Theorem 4.4), sharp functions (Theorem 4.14) and spherical harmonics (Theorem 4.31).

4.1 BMO **Functions**

In this section we recall some basic definitions and results concerning BMO functions from real analysis.

First, we let

$$L^1_{\mathrm{loc}}(\mathbf{R}^n) = \text{the space of equivalence classes of Lebesgue measurable functions on } \mathbf{R}^n \text{ which are integrable on every compact subset of } \mathbf{R}^n.$$

The elements of $L^1_{\mathrm{loc}}(\mathbf{R}^n)$ are called *locally integrable functions* on $\mathbf{R}^n$.

A function $f \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$ is said to be of *bounded mean oscillation*, $f \in \mathrm{BMO}$, if it satisfies the condition (see [37])

$$\|f\|_* := \sup_B \frac{1}{|B|} \int_B |f(x) - f_B| \, dx < \infty,$$

where the supremum is taken over all balls B in $\mathbf{R}^n$, $|B|$ is the Lebesgue measure of B and f_B is the integral average of f over B

$$f_B := \frac{1}{|B|} \int_B f(y) \, dy.$$

The quantity $\|\cdot\|_*$ is called the BMO *norm*. This is not properly a norm, since any function which is constant almost everywhere has zero oscillation. However, it is easy to see that these are the only functions having zero oscillation. Therefore, we view the class BMO as the quotient space of the above space by the space of constant functions. In other words, two functions which differ by a constant coincide as functions in the class BMO. Then it should be emphasized (see [52]) that the class BMO is a Banach space equipped with the BMO norm $\|\cdot\|_*$.

It is not important that we subtract exactly f_B in the definition of BMO. More precisely, we can obtain the following claim:

Claim 4.1. Let $f \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$. Assume that, for each ball B, there exists a constant α_B such that we have the inequality

$$\frac{1}{|B|}\int_B |f(x)-\alpha_B|\,dx \leq C,$$

with some positive constant C independent of B. Then it follows that $f \in \mathrm{BMO}$ with $\|f\|_* \leq 2C$.

Proof. Indeed, we have, for all balls B,

$$|f_B-\alpha_B| = \frac{1}{|B|}\left|\int_B (f(y)-f_B)\,dy\right| \leq \frac{1}{|B|}\int_B |f(x)-\alpha_B|\,dx \leq C.$$

Hence it follows that

$$\begin{aligned}\frac{1}{|B|}\int_B |f(x)-f_B|\,dx &\leq \frac{1}{|B|}\int_B |f(x)-\alpha_B|\,dx + \frac{1}{|B|}\int_B |\alpha_B-f_B|\,dx \\ &\leq 2C.\end{aligned}$$

This proves that $f \in \mathrm{BMO}$ with $\|f\|_* \leq 2C$.

The proof of Claim 4.1 is complete. □

Sometimes we will define the BMO norm with cubes in place of balls. If Q is a cube with sides parallel to the coordinate axes, then we can define an equivalent BMO norm $\|\cdot\|_{*'}$ by the formula

$$\|f\|_{*'} := \sup_Q \frac{1}{|Q|}\int_Q |f(x)-f_Q|\,dx,$$

where the supremum is taken over all cubes Q in $\mathbf{R}^n$, and f_Q is the integral average of f over Q

$$f_Q := \frac{1}{|Q|}\int_Q f(y)\,dy.$$

In fact, we can obtain the following lemma:

Lemma 4.1. *For any function $f \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$, we have the inequalities*

$$\|f\|_{*'} \le \frac{n^{n/2} V_n}{2^{n-1}} \|f\|_{*}, \tag{4.1}$$

$$\|f\|_{*} \le \frac{2^{n+1}}{V_n} \|f\|_{*'}, \tag{4.2}$$

where

$$V_n := \frac{\omega_n}{n} = \frac{\pi^{n/2}}{\Gamma(n/2+1)}$$

is the volume of the unit ball in $\mathbf{R}^n$.

Proof. (1) Let $Q = Q(x,r)$ be a cube with side length $2r$ and with center x, and let $B = B(x, \sqrt{n}\, r)$ be a ball of radius $\sqrt{n}\, r$ about x (see Figure 4.1). Then it follows that

$$Q(x,r) \subset B(x, \sqrt{n}\, r),$$
$$|Q(x,r)| = (2r)^n, \quad |B(x, \sqrt{n}\, r)| = n^{n/2} r^n V_n.$$

Hence we have the inequality

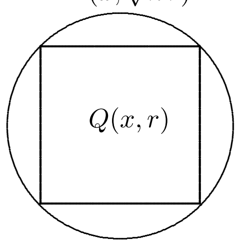

Fig. 4.1. The cube $Q(x,r)$ and the ball $B(x, \sqrt{n}\, r)$

$$\begin{aligned}
&|f_Q - f_B| \\
&= \left| \frac{1}{|Q|} \int_Q (f(x) - f_B)\, dx \right| \le \frac{1}{|Q|} \int_Q |f(x) - f_B|\, dx \\
&\le \left(\frac{|B|}{|Q|} \right) \left(\frac{1}{|B|} \int_B |f(x) - f_B|\, dx \right) = \frac{n^{n/2} V_n}{2^n} \left(\frac{1}{|B|} \int_B |f(x) - f_B|\, dx \right) \\
&\le \frac{n^{n/2} V_n}{2^n} \|f\|_{*}.
\end{aligned}$$

Therefore, we obtain that

$$\begin{aligned}
&\frac{1}{|Q|}\int_Q |f(x)-f_Q|\,dx \\
&\le \frac{1}{|Q|}\int_Q |f(x)-f_B|\,dx + \frac{1}{|Q|}\int_Q |f_B-f_Q|\,dx \\
&\le \left(\frac{|B|}{|Q|}\right)\frac{1}{|B|}\int_B |f(x)-f_B|\,dx + |f_Q-f_B| \\
&\le \frac{n^{n/2}V_n}{2^n}\left(\frac{1}{|B|}\int_B |f(x)-f_B|\,dx\right) + \frac{n^{n/2}V_n}{2^n}\,\|f\|_* \\
&\le \frac{n^{n/2}V_n}{2^{n-1}}\,\|f\|_* .
\end{aligned}$$

This proves the desired inequality (4.1).

(2) Similarly, if $Q = Q(x,r)$ is a cube with side length $2r$ and with center x and if $B = B(x,r)$ is a ball of radius r about x (see Figure 4.2), then it follows that

$$\begin{aligned}
&B(x,r) \subset Q(x,r), \\
&|B(x,r)| = r^n V_n, \quad |Q(x,r)| = (2r)^n.
\end{aligned}$$

Hence we have the inequality

$$\begin{aligned}
|f_B - f_Q| &= \left|\frac{1}{|B|}\int_B (f(x)-f_Q)\,dx\right| \le \frac{1}{|B|}\int_B |f(x)-f_Q|\,dx \\
&\le \left(\frac{|Q|}{|B|}\right)\left(\frac{1}{|Q|}\int_Q |f(x)-f_Q|\,dx\right) \\
&= \frac{2^n}{V_n}\left(\frac{1}{|Q|}\int_Q |f(x)-f_Q|\,dx\right) \\
&\le \frac{2^n}{V_n}\,\|f\|_{*'} .
\end{aligned}$$

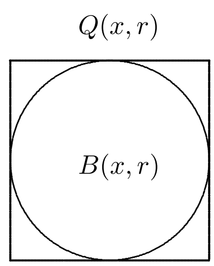

Fig. 4.2. The cube $Q(x,r)$ and the ball $B(x,r)$

Therefore, we obtain that

$$\begin{aligned}
&\frac{1}{|B|}\int_B |f(x)-f_B|\,dx \\
&\le \frac{1}{|B|}\int_B |f(x)-f_Q|\,dx + \frac{1}{|B|}\int_B |f_Q-f_B|\,dx \\
&\le \left(\frac{|Q|}{|B|}\right)\frac{1}{|Q|}\int_Q |f(x)-f_Q|\,dx + |f_B-f_Q| \\
&\le \frac{2^n}{V_n}\left(\frac{1}{|Q|}\int_Q |f(x)-f_Q|\,dx\right) + \frac{2^n}{V_n}\,\|f\|_{*'} \\
&\le \frac{2^{n+1}}{V_n}\,\|f\|_{*'}\,.
\end{aligned}$$

This proves the desired inequality (4.2).

The proof of Lemma 4.1 is complete. □

If Q is a cube with sides parallel to the coordinate axes, then we denote by δ_Q its side length and by x_Q its center, respectively. For each $\lambda > 0$, we denote by λQ the cube centered at x_Q with side length $\lambda\delta_Q$ (see Figure 4.3 below).

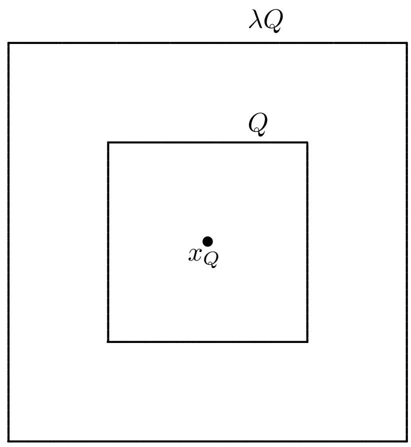

Fig. 4.3. The cubes Q and λQ

Lemma 4.2. *Let* $f \in \mathrm{BMO}$. *For any positive integer* j, *there exists a constant* $c(n) > 0$ *such that*

$$|f_{2^jQ} - f_Q| \le c(n)j\,\|f\|_*\,. \tag{4.3}$$

For example, we may take

$$c(n) = 2\, n^{n/2}\, V_n = 2\, \frac{(n\pi)^{n/2}}{\Gamma(n/2+1)}.$$

Proof. First, it follows that

$$\begin{aligned} |f_{2Q} - f_Q| &= \left| \frac{1}{|Q|} \int_Q (f_{2Q} - f(x))\, dx \right| \leq \frac{1}{|Q|} \int_Q |f_{2Q} - f(x)|\, dx \\ &\leq \frac{1}{|Q|} \int_{2Q} |f_{2Q} - f(x)|\, dx = 2^n \left(\frac{1}{|2Q|} \int_{2Q} |f_{2Q} - f(x)|\, dx \right) \\ &\leq 2^n \, \|f\|_{*'} \, . \end{aligned}$$

Similarly, we have, for any positive integer j,

$$\left| f_{2^j Q} - f_{2^{j-1}Q} \right| \leq 2^n \, \|f\|_{*'} \, .$$

Hence it follows that

$$\begin{aligned} \left| f_{2^j Q} - f_Q \right| &\leq |f_{2^j Q} - f_{2^{j-1}Q}| + |f_{2^{j-1}Q} - f_{2^{j-2}Q}| + \dots \\ &\quad + |f_{4Q} - f_{2Q}| + |f_{2Q} - f_Q| \\ &\leq j 2^n \, \|f\|_{*'} \, . \end{aligned} \tag{4.4}$$

However, we have, by Lemma 4.1,

$$\|f\|_{*'} \leq \frac{n^{n/2} V_n}{2^{n-1}} \, \|f\|_* \, . \tag{4.5}$$

Therefore, the desired inequality (4.3) follows by combining inequalities (4.4) and (4.5):

$$\left| f_{2^j Q} - f_Q \right| \leq j 2^n \, \|f\|_{*'} \leq j 2^n \frac{n^{n/2} V_n}{2^{n-1}} \, \|f\|_* = 2 j n^{n/2} V_n \, \|f\|_* \, .$$

The proof of Lemma 4.2 is complete. □

4.2 VMO Functions

Next we introduce a subspace of BMO functions whose BMO norm over a ball vanishes as the radius of the ball tends to zero. More precisely, if $f \in$ BMO and $r > 0$, then we let

$$\eta(r) := \sup_{\rho \leq r} \frac{1}{|B|} \int_B |f(x) - f_B|\, dx,$$

where the supremum is taken over all balls B with radius $\rho \leq r$. A function $f \in \mathrm{BMO}$ has *vanishing mean oscillation*, $f(x) \in \mathrm{VMO}$, if it satisfies the condition (see [61])

$$\lim_{r\downarrow 0} \eta(r) = 0.$$

The function $\eta(r)$ will be referred as the VMO modulus of f. The assumption that $f \in \mathrm{VMO}$ means a kind of continuity in the average sense, not in the pointwise sense.

First, we prove the following claim:

Claim 4.2. (i) Uniformly continuous functions which belong to BMO are VMO functions.

(ii) VMO is a closed subspace of BMO.

Proof. (i) Assume that $f(x)$ is a uniformly continuous function which belongs to BMO. Then, for any $\varepsilon > 0$ there exists a constant $r = r(\varepsilon) > 0$ such that, for each ball B with radius r,

$$x, y \in B \implies |f(x) - f(y)| < \varepsilon,$$

or equivalently,

$$\sup_{x,y\in B} |f(x) - f(y)| \leq \varepsilon.$$

Hence it follows that we have, for all $x \in B$,

$$\begin{aligned} |f(x) - f_B| &= \frac{1}{|B|}\left|\int_B (f(x) - f(y))\, dy\right| \leq \frac{1}{|B|}\int_B |f(x) - f(y)|\, dy \\ &\leq \sup_{x,y\in B} |f(x) - f(y)| \\ &\leq \varepsilon. \end{aligned}$$

Therefore, we obtain that, for all balls B with radius $\rho \leq r$,

$$\frac{1}{|B|}\int_B |f(x) - f_B|\, dx \leq \varepsilon,$$

so that

$$\eta(r) \leq \varepsilon.$$

This proves that $\lim_{r\downarrow 0} \eta(r) = 0$, that is, $f \in \mathrm{VMO}$.

(ii) Assume that f is the limit of a sequence $\{f_j\}$ of VMO functions in the BMO norm

$$\lim_{j\to\infty} \|f_j - f\|_* = 0.$$

Then it follows that, for all balls B with radius $\rho \leq r$,

$$\begin{aligned} &\frac{1}{|B|}\int_B |f(x) - f_B|\, dx \\ &\leq \frac{1}{|B|}\int_B |f_j(x) - f(x) - (f_j)_B + f_B|\, dx + \frac{1}{|B|}\int_B |f_j(x) - (f_j)_B|\, dx \\ &\leq \|f_j - f\|_* + \eta_j(r), \end{aligned}$$

where

$$\eta_j(r) = \sup_{\rho \leq r} \frac{1}{|B|}\int_B |f_j(x) - (f_j)_B|\, dx.$$

Hence we have the inequality

$$\eta(r) = \sup_{\rho \leq r} \frac{1}{|B|}\int_B |f(x) - f_B|\, dx \leq \|f_j - f\|_* + \eta_j(r). \tag{4.6}$$

However, for any given $\varepsilon > 0$ there exists a positive integer $N = N(\varepsilon)$ such that

$$\|f_N - f\|_* < \frac{\varepsilon}{2}. \tag{4.7}$$

Moreover, we can find a positive number $\delta = \delta(N, \varepsilon)$ such that

$$\eta_N(r) = \sup_{\rho \leq r} \frac{1}{|B|}\int_B |f_N(x) - (f_N)_B|\, dx < \frac{\varepsilon}{2} \quad \text{for all } 0 < r < \delta. \tag{4.8}$$

Therefore, by combining inequalities (4.6) with $j := N$, (4.7) and (4.8) we obtain that

$$\begin{aligned} \eta(r) &= \sup_{\rho \leq r} \frac{1}{|B|}\int_B |f(x) - f_B|\, dx \leq \|f_N - f\|_* + \eta_N(r) \\ &\leq \varepsilon \quad \text{for all } 0 < r < \delta. \end{aligned}$$

This proves that $\lim_{r \downarrow 0} \eta(r) = 0$, that is, $f \in \text{VMO}$.

The proof of Claim 4.2 is complete. □

The relationship between BMO and its subspace VMO is quite similar to the relationship between $L^\infty(\mathbf{R}^n)$ and its subspace BUC of bounded uniformly continuous functions on $\mathbf{R}^n$. This can be visualized as follows:

$$\boxed{\begin{array}{ccc} L^\infty(\mathbf{R}^n) & \longrightarrow & \text{BMO} \\ \uparrow & & \uparrow \\ \text{BUC} & \longrightarrow & \text{VMO} \end{array}}$$

In fact, the next theorem collects some important results concerning VMO functions (see [32, Chapter VI, Theorem 5.1], [91, Chapter VIII]):

Theorem 4.3. *For a function* $f \in \mathrm{BMO}$*, the following three conditions (i), (ii) and (iii) are equivalent:*

(i) f *is in* VMO.
(ii) f *is in the* BMO *closure of uniformly continuous functions that belong to* BMO.
(iii) $\lim_{y \to 0} \|f(\cdot - y) - f(\cdot)\|_* = 0$*, where* $f(x-y)$ *is the translation of* $f(x)$ *by* y*-units. More precisely, there exists a positive constant* $C = C(n)$ *such that*

$$\|f(\cdot - y) - f(\cdot)\|_* \leq C\eta(r) \quad \textit{for } |y| < r. \tag{4.9}$$

Examples 4.1. (i) $\ln|x| \in \mathrm{BMO}$, but $\ln|x| \notin \mathrm{VMO}$ $(n = 1)$.
(ii) $\ln|\ln|x|| \in \mathrm{VMO}$ $(n = 1)$.
(iii) $L^\infty(\mathbf{R}^n) \subset \mathrm{BMO}$.
(iv) $W^{\theta, n/\theta}(\mathbf{R}^n) \subset \mathrm{VMO}$ for $0 < \theta \leq 1$ (see Proposition 7.7).
(v) The Sobolev space $W^{1,n}(\mathbf{R}^n)$ is a proper subspace of VMO. For example, $|\ln|x||^\alpha \in W^{1,n}(\mathbf{R}^n)$ if $0 < \alpha < 1 - 1/n$, while $|\ln|x||^\alpha \in \mathrm{VMO} \setminus W^{1,n}(\mathbf{R}^n)$ if $1 - 1/n \leq \alpha < 1$.

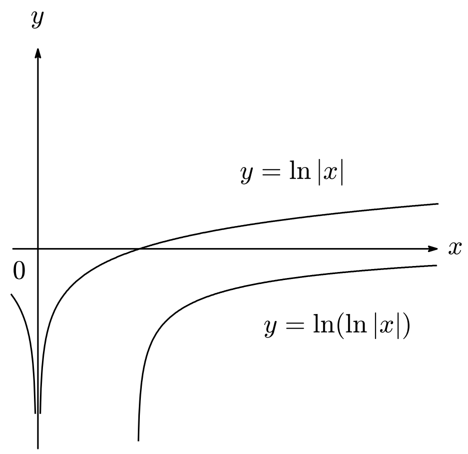

Fig. 4.4. The functions $\ln|x|$ and $\ln|\ln|x||$

Let Ω be a bounded domain of $\mathbf{R}^n$ with $n \geq 3$. Then it should be emphasized that, by replacing the ball B above by the intersection $B \cap \Omega$

we obtain the definitions of BMO(Ω) and VMO(Ω). Given a function defined on Ω that belongs to BMO(Ω) (resp. VMO(Ω)), we can extend it to the whole $\mathbf{R}^n$ preserving its BMO (resp. VMO) character (see [1, Proposition 1.3]).

4.3 The Calderón–Zygmund Decomposition

Let $f \in L^1(\mathbf{R}^n)$. The purpose of this section is to describe the splitting of the space $\mathbf{R}^n$ into a subset Ω made up of non-overlapping cubes over each of which the average of $|f|$ is between t and $2^n t$ and the complement $\mathbf{R}^n \setminus \Omega$ where $|f(x)| \leq t$ (Theorem 4.7). This is known as the *Calderón–Zygmund decomposition.* The Calderón–Zygmund decomposition is a key step in real analysis. The idea behind this decomposition is that it is often useful to split an arbitrary integrable function into its "small" and "large" parts, and then we use different techniques to analyze each part.

In the Euclidean space $\mathbf{R}^n$ we define the unit cube, open on the right, to be the set $[0, 1)^n$, and we let

$$\mathcal{D}_0 = \text{the collection of cubes in } \mathbf{R}^n \text{ which are congruent to } [0,1)^n \text{ whose vertices lie on the lattice } \mathbf{Z}^n.$$

If we dilate this family $\mathcal{Q}_0$ of cubes by a factor 2^{-k} for each integer $k \in \mathbf{Z}$, we obtain the family of cubes $\mathcal{D}_k$ as follows:

$$\mathcal{D}_k = \text{the collection of cubes, open on the right, whose vertices are adjacent points of the lattice } (2^{-k}\,\mathbf{Z})^n.$$

The cubes in $\mathcal{D} = \bigcup_{k \in \mathbf{Z}} \mathcal{D}_k$ are called *dyadic cubes.*

From this construction, we obtain the following three assertions (1), (2) and (3):

(1) Given $x \in \mathbf{R}^n$, there exists a unique cube in each family $\mathcal{D}_k$ which contains x.
(2) Any two dyadic cubes are either disjoint or one is contained in the other.
(3) A dyadic cube in $\mathcal{D}_k$ is contained in a unique cube of each family $\mathcal{D}_j$, $j < k$, and contains 2^n dyadic cubes of $\mathcal{D}_{k+1}$.

If $f \in L^1_{\text{loc}}(\mathbf{R}^n)$, we define the Hardy–Littlewood *maximal function*

Mf by the formula

$$Mf(x) := \sup_{x \in Q} \frac{1}{|Q|} \int_Q |f(y)|\, dy,$$

where the supremum is taken over all cubes Q containing x and $|Q|$ is the Lebesgue measure of Q. The Hardy–Littlewood maximal function will be studied in detail in the next Section 4.4.

In order to study the size of Mf, we shall look at its distribution function $|E_t|$. First, instead of looking at Mf, we try to obtain a family $\{Q_j\}$ of cubes such that the average of $|f|$ over each Q_j is comparable to t. This can be done most effectively by considering only dyadic cubes.

More precisely, we can prove the following theorem:

Theorem 4.4. *Let $f \in L^1(\mathbf{R}^n)$. Then, for every $t > 0$, there exists a family $\{Q_j\}$ of disjoint dyadic cubes which satisfies the following three conditions* (4.10)*,* (4.11) *and* (4.12)*:*

$$E_t = \{x \in \mathbf{R}^n : Mf(x) > t\} \subset \bigcup_j 3Q_j. \tag{4.10}$$

$$\frac{t}{4^n} < \frac{1}{|Q_j|} \int_{Q_j} |f(y)|\, dy \leq \frac{t}{2^n} \quad \textit{for every } Q_j. \tag{4.11}$$

$$|E_t| \leq \frac{12^n}{t} \int_{\mathbf{R}^n} |f(y)|\, dy. \tag{4.12}$$

Here we recall that if Q is a cube with center x_Q and with side length δ_Q, then αQ, $\alpha > 0$, denotes the cube with the same center x_Q as Q and with side length $\alpha\delta_Q$.

Proof. The proof is divided into two steps.

Step I: If $f \in L^1(\mathbf{R}^n)$ and $t > 0$, we let

$\mathcal{C}_t =$ the collection of cubes $Q \in \mathcal{D}$ which satisfy the condition

$$t < \frac{1}{|Q|} \int_Q |f(y)|\, dy \tag{4.13}$$

and are *maximal* among those which satisfy condition (4.13).

Then it is easy to verify the following three assertions (i), (ii) and (iii):

(i) Every cube $Q \in \mathcal{D}$ satisfying condition (4.13) is contained in some cube $Q' \in \mathcal{C}_t$, since condition (4.13) imposes an upper bound on the size of Q, that is, since we have the inequality

$$|Q| \leq \frac{1}{t} \int_{\mathbf{R}^n} |f(y)|\, dy.$$

(ii) The cubes in $\mathcal{C}_t$ are non-overlapping.

(iii) For any cube $Q \in \mathcal{C}_t$, we have the inequalities

$$t < \frac{1}{|Q|} \int_Q |f(y)|\, dy \leq 2^n t.$$

Indeed, if $Q \in \mathcal{D}_k$ is in $\mathcal{C}_t$ and if Q' is the only cube in $\mathcal{D}_{k-1}$ containing Q, then it follows that

$$\frac{1}{|Q'|} \int_{Q'} |f(y)|\, dy \leq t.$$

However, since $|Q'| = 2^n |Q|$, we obtain that

$$\frac{1}{|Q|} \int_Q |f(y)|\, dy \leq \frac{2^n}{|Q'|} \int_{Q'} |f(y)|\, dy \leq 2^n t.$$

Therefore, we have constructed a family $\mathcal{C}_t = \{Q_j\}$ of disjoint dyadic cubes which satisfies the condition

$$t < \frac{1}{|Q_j|} \int_{Q_j} |f(y)|\, dy \leq 2^n t. \tag{4.14}$$

Step II: (1) Assertion (4.10): If $\mathcal{C}_{t/4^n} = \{Q_j\}$, we show that

$$E_t = \{x \in \mathbf{R}^n : Mf(x) > t\} \subset \bigcup_j 3Q_j.$$

To do this, let x_0 be an arbitrary point of E_t. Then we can find a cube R containing x_0 in its interior and satisfying the condition

$$t < \frac{1}{|R|} \int_R |f(y)|\, dy.$$

We shall look for a dyadic cube of comparable size over which the average of $|f|$ is comparably large. If we take the only integer k such that

$$2^{-(k+1)n} < |R| \leq 2^{-kn},$$

then there exists at most one point of the lattice $\Lambda_k = 2^{-k}\mathbf{Z}^n$ which is an interior point of R. Here it should be noticed that there exists a cube in $\mathcal{D}_k$, and at most 2^n cubes in $\mathcal{D}_k$, meeting the interior of R. Hence we can find a cube $Q \in \mathcal{D}_k$ which meets the interior of R and satisfies the condition

$$\int_{R \cap Q} |f(y)|\, dy > \frac{t|R|}{2^n}.$$

However, it follows that

$$|R| \leq |Q| < 2^n |R|.$$

Hence we obtain that

$$\int_{R\cap Q} |f(y)|\,dy > \frac{t|R|}{2^n} > \frac{t}{2^n}\,\frac{|Q|}{2^n} = \frac{t|Q|}{4^n},$$

so that

$$\frac{1}{|Q|}\int_Q |f(y)|\,dy > \frac{t}{4^n}.$$

This implies that Q is contained in some cube $Q_j \in \mathcal{C}_{t/4^n}$.

Moreover, since R and Q meet and $|R| \leq |Q|$, it follows that

$$x_0 \in R \subset 3Q \subset 3Q_j.$$

This proves the desired assertion (4.10).

(2) Assertion (4.11): This follows from inequalities (4.14) with $t := t/4^n$

$$\frac{t}{4^n} < \frac{1}{|Q_j|}\int_{Q_j} |f(y)|\,dy \leq 2^n \cdot \frac{t}{4^n} = \frac{t}{2^n}.$$

(3) Assertion (4.12): By combining assertions (4.10) and (4.11), we obtain that

$$\begin{aligned}
|E_t| &\leq \sum_j |3Q_j| = 3^n \sum_j |Q_j| \\
&\leq 3^n \sum_j \frac{4^n}{t}\int_{Q_j} |f(y)|\,dy = \frac{12^n}{t}\sum_j \int_{Q_j} |f(y)|\,dy \\
&\leq \frac{12^n}{t}\int_{\mathbf{R}^n} |f(y)|\,dy.
\end{aligned}$$

The proof of Theorem 4.4 is complete. □

As an application of inequality (4.12), we can obtain an extension of *Lebesgue's differentiation theorem*:

Theorem 4.5 (Lebesgue's differentiation theorem). *Let* $f \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$. *If* $x \in \mathbf{R}^n$ *and* $r > 0$, *we let*

$$Q(x,r) := \prod_{k=1}^{n} [-r + x_k, r + x_k].$$

Then we have, for almost every point x *of* $\mathbf{R}^n$,

$$\lim_{r\downarrow 0} \frac{1}{|Q(x,r)|}\int_{Q(x,r)} |f(y) - f(x)|\,dy = 0. \tag{4.15}$$

Proof. Without loss of generality, we may assume that $f \in L^1(\mathbf{R}^n)$. Since we have the formula

$$\begin{aligned}&\{x \in \mathbf{R}^n : \text{assertion (4.15) does not hold}\}\\ &= \bigcup_{j=1}^{\infty} \left\{ x \in \mathbf{R}^n : \limsup_{r \downarrow 0} \frac{1}{|Q(x,r)|} \int_{Q(x,r)} |f(y) - f(x)|\, dy > \frac{1}{j} \right\},\end{aligned}$$

it suffices to show that, for every $t > 0$, the set

$$A_t := \left\{ x \in \mathbf{R}^n : \limsup_{r \downarrow 0} \frac{1}{|Q(x,r)|} \int_{Q(x,r)} |f(y) - f(x)|\, dy > t \right\}$$

has measure *zero*.

For any given $\varepsilon > 0$, we can find functions $g \in C_0(\mathbf{R}^n)$ and $h \in L^1(\mathbf{R}^n)$ such that

$$f(x) = g(x) + h(x),$$
$$\int_{\mathbf{R}^n} |h(x)|\, dx < \varepsilon.$$

Then we have the inequality

$$\begin{aligned}&\frac{1}{|Q(x,r)|} \int_{Q(x,r)} |f(y) - f(x)|\, dy\\ &\leq \frac{1}{|Q(x,r)|} \int_{Q(x,r)} |g(y) - g(x)|\, dy + |h(x)| + \frac{1}{|Q(x,r)|} \int_{Q(x,r)} |h(y)|\, dy\\ &\leq \frac{1}{|Q(x,r)|} \int_{Q(x,r)} |g(y) - g(x)|\, dy + |h(x)| + Mh(x).\end{aligned}$$

By the uniform continuity of g, it follows that

$$\lim_{r \downarrow 0} \frac{1}{|Q(x,r)|} \int_{Q(x,r)} |g(y) - g(x)|\, dy = 0.$$

Hence, by passing to the limit we obtain that

$$\limsup_{r \downarrow 0} \frac{1}{|Q(x,r)|} \int_{Q(x,r)} |f(y) - f(x)|\, dy \leq |h(x)| + Mh(x).$$

This implies that the set A_t can be decomposed as follows:

$$A_t \subset \left\{ x \in \mathbf{R}^n : Mh(x) > \frac{t}{2} \right\} \bigcup \left\{ x \in \mathbf{R}^n : |h(x)| > \frac{t}{2} \right\}.$$

However, we have the inequality

$$\left| \left\{ x \in \mathbf{R}^n : |h(x)| > \frac{t}{2} \right\} \right| = \int_{\{x \in \mathbf{R}^n : |h(x)| > t/2\}} dy$$

$$\leq \int_{\{x\in\mathbf{R}^n:|h(x)|>t/2\}} \frac{2}{t}|h(y)|\,dy$$
$$\leq \int_{\mathbf{R}^n} \frac{2}{t}|h(y)|\,dy = \frac{2}{t}\int_{\mathbf{R}^n} |h(y)|\,dy$$
$$< \frac{2}{t}\varepsilon,$$

and also, by inequality (4.12) with $t := t/2$,

$$\left|\left\{x \in \mathbf{R}^n : Mh(x) > \frac{t}{2}\right\}\right| \leq \frac{12^n}{t/2}\int_{\mathbf{R}^n} |h(y)|\,dy = \frac{2\cdot 12^n}{t}\int_{\mathbf{R}^n} |h(y)|\,dy$$
$$< \frac{2\cdot 12^n}{t}\varepsilon.$$

Summing up, we have proved that, for any given $\varepsilon > 0$,

$$|A_t| \leq \frac{2}{t}\varepsilon + \frac{2\cdot 12^n}{t}\varepsilon = \frac{2}{t}(1+12^n)\varepsilon.$$

This proves that $|A_t| = 0$ for every $t > 0$.

The proof of Theorem 4.5 is complete. □

Definition 4.1. Let $f \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$. A point x of $\mathbf{R}^n$ is called a *Lebesgue point* for f if it satisfies the condition

$$\lim_{r\downarrow 0} \frac{1}{|Q(x,r)|}\int_{Q(x,r)} |f(y) - f(x)|\,dy = 0. \tag{4.15}$$

Rephrased, Theorem 4.5 asserts that almost every point of $\mathbf{R}^n$ is a Lebesgue point for $f \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$.

Furthermore, we can prove the following corollary:

Corollary 4.6. *Let $f \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$. Then, for every Lebesgue point x for f, and thus, for almost every point $x \in \mathbf{R}^n$, we have the following two assertions* (4.16) *and* (4.17)*:*

$$f(x) = \lim_{r\downarrow 0} \frac{1}{|Q(x,r)|}\int_{Q(x,r)} f(y)\,dy. \tag{4.16}$$

$$|f(x)| \leq Mf(x). \tag{4.17}$$

Proof. (i) Let x be an arbitrary Lebesgue point for f. Then it follows from condition (4.15) that we have, as $r \downarrow 0$,

$$\left|\frac{1}{|Q(x,r)|}\int_{Q(x,r)} f(y)\,dy - f(x)\right| = \left|\frac{1}{|Q(x,r)|}\int_{Q(x,r)} (f(y) - f(x))\,dy\right|$$
$$\leq \frac{1}{|Q(x,r)|}\int_{Q(x,r)} |f(y) - f(x)|\,dy$$

$$\longrightarrow 0.$$

This proves the desired assertion (4.16).

(ii) We remark that

$$\frac{1}{|Q(x,r)|}\left|\int_{Q(x,r)} f(y)\,dy\right| \leq \frac{1}{|Q(x,r)|}\int_{Q(x,r)} |f(y)|\,dy \leq Mf(x). \tag{4.18}$$

However, by assertion (4.16) it follows that we have, as $r \downarrow 0$,

$$\left|\frac{1}{|Q(x,r)|}\left|\int_{Q(x,r)} f(y)\,dy\right| - |f(x)|\right| \leq \left|\frac{1}{|Q(x,r)|}\int_{Q(x,r)} f(y)\,dy - f(x)\right| \longrightarrow 0.$$

Therefore, by passing to the limit in inequality (4.18) we obtain that

$$|f(x)| = \lim_{r\downarrow 0}\frac{1}{|Q(x,r)|}\left|\int_{Q(x,r)} f(y)\,dy\right| \leq Mf(x).$$

This proves the desired assertion (4.17).

The proof of Corollary 4.6 is complete. □

Summing up, we can obtain the following Calderón–Zygmund decomposition theorem:

Theorem 4.7 (the Calderón–Zygmund decomposition)**.** *If $f(x)$ is an arbitrary function in the space $L^1(\mathbf{R}^n)$, then, for every $t > 0$ we can construct a family $\mathcal{C}_t = \{Q_j\}$ of disjoint maximal dyadic cubes over which the average of $|f|$ is greater than t. This family $\mathcal{C}_t$ satisfies the following two conditions* (4.14) *and* (4.19)*:*

$$t < \frac{1}{|Q_j|}\int_{Q_j} |f(y)|\,dy \leq 2^n t \quad \text{for every } Q_j. \tag{4.14}$$

$$|f(x)| \leq t \quad \text{for almost every } x \in \mathbf{R}^n \setminus \textstyle\bigcup_j Q_j. \tag{4.19}$$

Moreover, we have, for every $t > 0$,

$$E_t = \{x \in \mathbf{R}^n : Mf(x) > t\} \subset \bigcup_j 3Q_j, \tag{4.10}$$

where the Q_j range over $\mathcal{C}_{t/4^n}$.

Proof. By Theorem 4.4, we have only to prove assertion (4.19).

(i) Let x_0 be a Lebesgue point for f. If $\{Q_j\}$ is a sequence of cubes such that

$$Q_1 \supset Q_2 \supset \cdots \supset Q_j \supset \cdots,$$
$$\bigcap_j Q_j = \{x_0\},$$

then it follows that

$$f(x_0) = \lim_{j\to\infty} \frac{1}{|Q_j|} \int_{Q_j} f(y)\,dy. \tag{4.20}$$

Indeed, if the cube Q_j has side length r_j, then we have the assertion

$$Q_j \subset Q(x_0, r_j), \quad |Q_j| = r_j^n \downarrow 0,$$

Therefore, we obtain from condition (4.15) that, as $j \to \infty$,

$$\begin{aligned}\left|\frac{1}{|Q_j|}\int_{Q_j} f(y)\,dy - f(x_0)\right| &= \left|\frac{1}{|Q_j|}\int_{Q_j} (f(y) - f(x_0))\,dy\right| \\ &\le \frac{1}{|Q_j|}\int_{Q_j} |f(y) - f(x)|\,dy \\ &\le \frac{2^n}{|Q(x_0, r_j)|}\int_{Q(x,r_j)} |f(y) - f(x_0)|\,dy \\ &\longrightarrow 0.\end{aligned}$$

This proves the desired assertion (4.20).

(ii) Now, let $f \in L^1(\mathbf{R}^n)$, and let $\mathcal{C}_t = \{Q_j\}$ be a family of disjoint maximal dyadic cubes over which the average of $|f|$ is greater than t:

$$t < \frac{1}{|Q_j|}\int_{Q_j} |f(y)|\,dy \le 2^n t \quad \text{for every } Q_j.$$

If x_0 is a point of $\mathbf{R}^n \setminus \bigcup_j Q_j$, then we have, for any dyadic cube Q containing x_0,

$$\frac{1}{|Q|}\int_Q |f(y)|\,dy \le t.$$

Hence, by taking a sequence $\{R_k\}$ of dyadic cubes of decreasing size such that

$$\bigcap_k R_k = \{x_0\},$$

we obtain that

$$\frac{1}{|R_k|}\int_{R_k} |f(y)|\,dy \le t. \tag{4.21}$$

If, in addition, x_0 is a Lebesgue point for f and so for $|f|$, then we have, by inequality (4.21) and assertion (4.20),

$$|f(x_0)| = \lim_k \frac{1}{|R_k|}\int_{R_k} |f(y)|\,dy \le t.$$

Therefore, we have proved the desired assertion (4.19)

$$|f(x)| \le t \quad \text{for almost every } x \in \mathbf{R}^n \setminus \textstyle\bigcup_j Q_j.$$

The proof of Theorem 4.7 is complete. □

4.4 The Hardy–Littlewood Maximal Function

If $f \in L^1_{\text{loc}}(\mathbf{R}^n)$, we define the Hardy–Littlewood *maximal function* Mf by the formula

$$Mf(x) := \sup_{x\in Q} \frac{1}{|Q|}\int_Q |f(y)|\,dy,$$

where the supremum is taken over all cubes Q containing x. It should be noticed that we can take only those cubes Q for which x is an interior point. This implies that the function $x \mapsto Mf(x)$ is lower semicontinuous, since the set

$$E_t := \{x \in \mathbf{R}^n : Mf(x) > t\}$$

is open for every $t > 0$. Indeed, it suffices to note that

$$E_t = \{x \in \mathbf{R}^n : Mf(x) > t\} = \bigcup \left\{ Q^o : \frac{1}{|Q^o|}\int_{Q^o} |f(y)|\,dy > t \right\},$$

where the Q^o range over all open cubes containing x.

If we define the *maximal function* $M'f$ by the formula

$$M'f(x) := \sup_{r>0} \frac{1}{|Q(x,r)|}\int_{Q(x,r)} |f(y)|\,dy,$$

where

$$Q(x,r) := \left\{ y \in \mathbf{R}^n : \max_{1\le j\le n} |y_j - x_j| \le r \right\} = \prod_{j=1}^n [x_j - r, x_j + r],$$

then we have the inequalities

$$M'f(x) \le Mf(x) \le 2^n M'f(x), \quad x \in \mathbf{R}^n.$$

Indeed, if Q is a cube with side length ℓ and containing x, then it follows that

$$Q \subset Q(x,\ell),$$
$$|Q(x,\ell)| = (2\ell)^n = 2^n|Q|.$$

Hence we obtain that

$$\frac{1}{|Q|}\int_Q |f(y)|\,dy \leq \frac{2^n}{|Q(x,\ell)|}\int_{Q(x,\ell)} |f(y)|\,dy \leq 2^n M'f(x), \quad x \in \mathbf{R}^n.$$

This proves that

$$Mf(x) \leq 2^n M'f(x).$$

Furthermore, if we define the *maximal function* $M''f$ by the formula

$$M''f(x) := \sup_{x\in B}\frac{1}{|B|}\int_B |f(y)|\,dy,$$

where the supremum is taken over all balls B containing x, then, by arguing as in the proof of Lemma 4.1 we can obtain the inequalities

$$\frac{V_n}{2^n}M''f(x) \leq Mf(x) \leq \frac{V_n}{n^{n/2}}M''f(x), \quad x \in \mathbf{R}^n,$$

where $V_n = \omega_n/n = \pi^{n/2}/\Gamma(n/2+1)$ is the volume of the unit ball in $\mathbf{R}^n$.

The next theorem gives a fundamental norm estimate for the maximal function $Mf(x)$ for $1 < p < \infty$ (cf. [67, Chapter I, Section 1.3, Theorem 1]):

Theorem 4.8. *If $f \in L^p(\mathbf{R}^n)$ with $1 < p < \infty$, then it follows that $Mf \in L^p(\mathbf{R}^n)$. More precisely, we have the inequality*

$$\|Mf\|_{L^p(\mathbf{R}^n)} \leq C_p \|f\|_{L^p(\mathbf{R}^n)}, \tag{4.22}$$

with a positive constant C_p. For example, we may take

$$C_p = 2\left(\frac{12^n \cdot p}{p-1}\right)^{1/p}.$$

Proof. Our proof is divided into two steps.

Step I: The essential step in the proof is the following theorem:

Theorem 4.9. *Let $f \in L^1_{\text{loc}}(\mathbf{R}^n)$. If $t > 0$, we let*

$$E_t = \{x \in \mathbf{R}^n : Mf(x) > t\}.$$

Then we have the following two estimates for the Lebesgue measure $|E_t|$ of E_t:

$$|E_t| \leq \frac{2 \cdot 12^n}{t} \int_{\{x \in \mathbf{R}^n : |f(x)| > t/2\}} |f(y)|\, dy. \tag{4.23a}$$

$$|E_t| \geq \frac{1}{2^n t} \int_{\{x \in \mathbf{R}^n : |f(x)| > t\}} |f(y)|\, dy. \tag{4.23b}$$

Proof. (i) First, we decompose the function $f(x)$ as follows:

$$f(x) = f_1(x) + f_2(x),$$

where

$$f_1(x) = \begin{cases} f(x) & \text{if } |f(x)| > \frac{t}{2}, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$f_2(x) = \begin{cases} f(x) & \text{if } |f(x)| \leq \frac{t}{2}, \\ 0 & \text{otherwise.} \end{cases}$$

Then it follows that

$$\begin{aligned} Mf(x) &= \sup_{x \in Q} \frac{1}{|Q|} \int_Q |f(y)|\, dy \\ &\leq \sup_{x \in Q} \frac{1}{|Q|} \int_Q |f_1(y)|\, dy + \sup_{x \in Q} \frac{1}{|Q|} \int_Q |f_2(y)|\, dy \\ &= Mf_1(x) + Mf_2(x). \end{aligned}$$

However, since we have the inequality

$$|f_2(x)| \leq \frac{t}{2} \quad \text{in } \mathbf{R}^n,$$

it follows that

$$Mf_2(x) = \sup_{x \in Q} \frac{1}{|Q|} \int_Q |f_2(y)|\, dy \leq \frac{t}{2}.$$

Summing up, we have proved that

$$Mf(x) \leq Mf_1(x) + Mf_2(x) \leq \frac{t}{2}.$$

This implies that

$$Mf(x) > t \implies Mf_1(x) > \frac{t}{2}.$$

Therefore, by applying inequality (4.12) to the function f_1 we obtain that

$$\begin{aligned}|E_t| &= |\{x \in \mathbf{R}^n : Mf(x) > t\}| \\ &\leq \left|\left\{x \in \mathbf{R}^n : Mf_1(x) > \frac{t}{2}\right\}\right| \\ &\leq \frac{12^n}{t/2} \int_{\mathbf{R}^n} |f_1(y)|\, dy = \frac{2 \cdot 12^n}{t} \int_{\mathbf{R}^n} |f_1(y)|\, dy \\ &= \frac{2 \cdot 12^n}{t} \int_{\{x \in \mathbf{R}^n : |f(x)| > t/2\}} |f(y)|\, dy.\end{aligned}$$

This proves the desired estimate (4.23a).

(ii) Without loss of generality, we may assume that $f \in L^1(\mathbf{R}^n)$. Indeed, otherwise, we truncate and apply a limiting process.

Then, by using the Calderón–Zygmund decomposition for f (Theorem 4.7) we can construct a family $\{Q_j\}$ of disjoint dyadic cubes such that

$$t < \frac{1}{|Q_j|} \int_{Q_j} |f(y)|\, dy \leq 2^n t \quad \text{for every } Q_j.$$
$$|f(x)| \leq t \quad \text{for almost every } x \in \mathbf{R}^n \setminus \textstyle\bigcup_{j=1}^{\infty} Q_j.$$

Then it follows that

$$x \in Q_j \implies Mf(x) \geq \frac{1}{|Q_j|} \int_{Q_j} |f(y)|\, dy > t \implies x \in E_t.$$

This proves that

$$\bigcup_{j=1}^{\infty} Q_j \subset E_t.$$

Hence, we obtain that

$$|E_t| \geq \sum_{j=1}^{\infty} |Q_j| \geq \frac{1}{2^n t} \sum_{j=1}^{\infty} \int_{Q_j} |f(y)|\, dy = \frac{1}{2^n t} \int_{\bigcup_{j=1}^{\infty} Q_j} |f(y)|\, dy.$$

However, we remark that

$$x \notin \bigcup_{j=1}^{\infty} Q_j \implies |f(x)| \leq t,$$

or equivalently,

$$\{x \in \mathbf{R}^n : |f(x)| > t\} \subset \bigcup_{j=1}^{\infty} Q_j.$$

Therefore, we obtain that

$$\begin{aligned}|E_t| &\geq \sum_{j=1}^{\infty} |Q_j| \geq \frac{1}{2^n t} \int_{\bigcup_{j=1}^{\infty} Q_j} |f(y)|\, dy \\ &\geq \frac{1}{2^n t} \int_{\{x \in \mathbf{R}^n : |f(x)| > t\}} |f(y)|\, dy.\end{aligned}$$

This proves the desired estimate (4.23b).

The proof of Theorem 4.9 is complete. □

Step II: By applying Theorem 3.29 and Fubini's theorem (Theorem 3.10) to our situation, we obtain from estimate (4.23a) that

$$\begin{aligned}\|Mf\|^p_{L^p(\mathbf{R}^n)} &= \int_{\mathbf{R}^n} (Mf(x))^p\, dx = \int_0^\infty p t^{p-1}\, |E_t|\, dt \\ &\leq 2p \cdot 12^n \int_0^\infty t^{p-2} \left(\int_{\{y \in \mathbf{R}^n : |f(y)| > t/2\}} |f(x)|\, dx \right) dt \\ &= 2p \cdot 12^n \int_{\mathbf{R}^n} \left(\int_0^{2|f(x)|} t^{p-2}\, dt \right) |f(x)|\, dx \\ &= \frac{2p \cdot 12^n}{p-1} \int_{\mathbf{R}^n} (2|f(x)|)^{p-1}\, |f(x)|\, dx \\ &= \frac{2^p p \cdot 12^n}{p-1} \int_{\mathbf{R}^n} |f(x)|^p\, dx \\ &= \frac{2^p p \cdot 12^n}{p-1} \|f\|^p_{L^p(\mathbf{R}^n)}\,.\end{aligned}$$

This proves the desired estimate (4.22).

Now the proof of Theorem 4.8 is complete. □

Remark 4.1. Theorem 4.8 remains valid if the norm is taken on the half-space $\mathbf{R}^n_+$. In this case, the definition of BMO functions need to be modified by taking only cubes Q contained in $\mathbf{R}^n_+$ (see Figure 4.4).

4.5 The John–Nirenberg Inequality

In this section we study the rate of growth of functions in BMO. The next theorem asserts that logarithmic growth is the maximum possible for BMO functions:

Theorem 4.10 (John–Nirenberg)**.** *Let* $f \in$ BMO*. Then we can find*

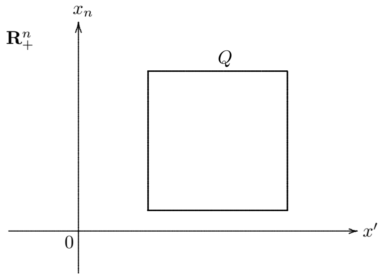

Fig. 4.5. The cube Q contained in the half-space $\mathbf{R}^n_+$

constants $C_1 > 0$ and $C_2 > 0$, depending only on the dimension, such that we have, for any cube Q and any $t > 0$,

$$|\{x \in Q : |f(x) - f_Q| > t\}| \leq C_1 \, e^{-(C_2/\|f\|_*)t} \, |Q|. \tag{4.24}$$

For example, we may take

$$C_1 = \sqrt{e}\,, \quad C_2 = \frac{1}{2^{n+1}e}.$$

Proof. First, we may assume that $\|f\|_* = 1$, since the John–Nirenberg inequality (4.24) is homogeneous. Indeed, if we let

$$g(x) := \frac{f(x)}{\|f\|_*}, \quad s := \frac{t}{\|f\|_*},$$

then we can rewrite inequality (4.24) in the form

$$|\{x \in Q : |g(x) - g_Q| > s\}| \leq C_1 \, e^{-C_2 s} \, |Q|, \quad s > 0. \tag{4.24$'$}$$

The proof of inequality (4.24′) is divided into three steps.

Step I: We fix a cube Q and take the constant e. Then it follows that

$$\frac{1}{|Q|} \int_Q |f(x) - f_Q| \, dx \leq 1 < e.$$

We make use of the Calderón–Zygmund decomposition of Q for the function

$$f(x) - f_Q$$

relative to e (see Theorem 4.7). This is done just as in Theorem 4.7 except that we begin by bisecting the sides of Q to form 2^n equal cubes $\{Q_{1,j}\}$. Then we have the following assertions:

$$e < \frac{1}{|Q_{1,j}|} \int_{Q_{1,j}} |f(x) - f_Q| \, dx \leq 2^n e.$$

$|f(x) - f_Q| \le e$ for almost every $x \in Q \setminus \bigcup_j Q_{1,j}$.

$|f_{Q_{1,j}} - f_Q| \le 2^n e$.

$$\sum_j |Q_{1,j}| \le \frac{1}{e}|Q|.$$

Indeed, it suffices to note that

$$|f_{Q_{1,j}} - f_Q| = \frac{1}{|Q_{1,j}|}\left|\int_{Q_{1,j}} (f(y) - f_Q)\,dy\right| \le \frac{1}{|Q_{1,j}|}\int_{Q_{1,j}} |f(y) - f_Q|\,dy$$
$$\le 2^n e,$$

and that

$$\sum_j |Q_{1,j}| \le \frac{1}{e}\sum_j \int_{Q_{1,j}} |f(y) - f_Q|\,dy \le \frac{1}{e}\int_Q |f(y) - f_Q|\,dy$$
$$\le \frac{1}{e}|Q|.$$

Step II: On each $Q_{1,j}$, we make use of the Calderón–Zygmund decomposition of $Q_{1,j}$ for the function

$$f(x) - f_{Q_{1,j}}$$

relative to e (see Theorem 4.7). This is done just as in Theorem 4.7 except that we begin by bisecting the sides of $Q_{1,j}$ to form 2^n equal cubes $\{Q_{2,k}\}$. Then we have the following assertions:

$$e < \frac{1}{|Q_{2,k}|}\int_{Q_{2,k}} |f(x) - f_{Q_{1,j}}|\,dx \le 2^n e.$$

$|f(x) - f_{Q_{1,j}}| \le e$ for almost every $x \in Q_{1,j} \setminus \bigcup_k Q_{2,k}$.

$|f_{Q_{2,k}} - f_{Q_{1,j}}| \le 2^n e$.

$$\sum_k |Q_{2,k}| \le \frac{1}{e}|Q_{1,j}|.$$

Indeed, it suffices to note that

$$\sum_k |Q_{2,k}| \le \frac{1}{e}\sum_k \int_{Q_{2,k}} |f(y) - f_{Q_{1,j}}|\,dy \le \frac{1}{e}\int_{Q_{1,j}} |f(y) - f_{Q_{1,j}}|\,dy$$
$$\le \frac{1}{e}|Q_{1,j}|.$$

Now we gather the cubes $Q_{2,k}$ corresponding to all the $Q_{1,j}$ and collectively rename them $Q_{2,k}$. Then we have the following assertions:

$$\sum_k |Q_{2,k}| \leq \frac{1}{e} \sum_j |Q_{1,j}| \leq \left(\frac{1}{e}\right)^2 |Q|.$$

$$\begin{aligned} |f(x) - f_Q| &\leq |f(x) - f_{Q_{1,j}}| + |f_{Q_{1,j}} - f_Q| \leq e + 2^n e \\ &\leq 2 \cdot 2^n e \quad \text{for almost every } x \in Q \setminus \textstyle\bigcup_k Q_{2,k}. \end{aligned}$$

Step III: By repeating this process for all cubes formed, we have, after N-steps, a family $\{Q_{N,j}\}$ of subcubes of Q which satisfies the following two conditions:

$$\sum_j |Q_{N,j}| \leq \left(\frac{1}{e}\right)^N |Q|.$$

$$|f(x) - f_Q| \leq N \cdot 2^n e \quad \text{for almost every } x \in Q \setminus \textstyle\bigcup_j Q_{N,j}.$$

(a) If $N \cdot 2^n e \leq t < (N+1) \cdot 2^n e$ with $N \in \mathbf{N}$, then it follows that

$$x \in Q,\ |f(x) - f_Q| > t \implies |f(x) - f_Q| > N \cdot 2^n e \implies x \in \bigcup_j Q_{N,j}.$$

This proves that

$$\{x \in Q : |f(x) - f_Q| > t\} \subset \bigcup_{j=1}^{\infty} Q_{N,j}.$$

Hence we have, for all $t \geq 2^n e$,

$$|\{x \in Q : |f(x) - f_Q| > t\}| \leq \sum_{j=1}^{\infty} |Q_{N,j}| \leq e^{-N}|Q| \leq e^{-C_2 t}|Q|,$$

with

$$C_2 := \frac{1}{2^{n+1}\, e}.$$

(b) On the other hand, if $0 < t < 2^n e$, then it follows that

$$C_2 t = \frac{1}{2^{n+1} e} \cdot t < \frac{1}{2}.$$

Hence we have, for all $0 < t < 2^n e$,

$$|\{x \in Q : |f(x) - f_Q| > t\}| \leq |Q| < e^{1/2 - C_2 t}\, |Q| = C_1\, e^{-C_2 t}\, |Q|\,,$$

with

$$C_1 := \sqrt{e}\,, \quad C_2 := \frac{1}{2^{n+1} e}.$$

Summing up, we have proved that

$$|\{x \in Q : |f(x) - f_Q| > t\}| \leq \begin{cases} C_1 \, e^{-C_2 t} \, |Q| & \text{for all } 0 < t < 2^n e, \\ e^{-C_2 t} \, |Q| & \text{for all } t \geq 2^n e. \end{cases}$$

This proves the desired inequality (4.24′).

The proof of Theorem 4.10 is complete. □

The next theorem asserts that if $f \in \mathrm{BMO}$, then it is locally in L^p for any $1 < p < \infty$ (see [70, Chapter IV, Section 1.3, Corollary]):

Theorem 4.11. *Let $f \in \mathrm{BMO}$ and $1 < p < \infty$. Then we have, for all cubes Q,*

$$\left(\frac{1}{|Q|} \int_Q |f(y) - f_Q|^p \, dy \right)^{1/p} \leq c_p \, \|f\|_* \,, \tag{4.25}$$

with a positive constant c_p. For example, we may take

$$c_p = \frac{(C_1 \, p \, \Gamma(p))^{1/p}}{C_2}.$$

Proof. We let

$$E_t := \{x \in Q : |f(x) - f_Q| > \lambda\} \,, \quad t > 0.$$

Then we have, by Theorem 3.29 and the John–Nirenberg inequality (4.24),

$$\begin{aligned} \int_Q |f(y) - f_Q|^p \, dy &= \int_0^\infty p \, t^{p-1} \, |E_t| \, dt \\ &\leq C_1 \, |Q| \int_0^\infty p \, t^{p-1} \, e^{-C_2 t / \|f\|_*} \, dt. \end{aligned}$$

Therefore, by making the change of variables

$$s = \frac{C_2 t}{\|f\|_*},$$

we obtain that

$$\begin{aligned} \frac{1}{|Q|} \int_Q |f(y) - f_Q|^p \, dy &\leq C_1 \, p \left(\frac{\|f\|_*}{C_2} \right)^p \int_0^\infty s^{p-1} \, e^{-s} \, ds \\ &= C_1 \, p \, \Gamma(p) \left(\frac{\|f\|_*}{C_2} \right)^p. \end{aligned}$$

This proves inequality (4.25) with

$$c_p := \frac{(C_1 \, p \, \Gamma(p))^{1/p}}{C_2}.$$

The proof of Theorem 4.11 is complete. □

Remark 4.2. Theorem 4.11 remains valid if the norm is taken on the half-space $\mathbf{R}^n_+$. In this case, the definition of BMO functions need to be modified by taking only cubes Q contained in $\mathbf{R}^n_+$ (see Figure 4.4).

Rephrased, Theorem 4.11 asserts that, for each $1 < p < \infty$, the quantity

$$\|f\|_{*,p} := \sup_Q \left(\frac{1}{|Q|} \int_Q |f(y) - f_Q|^p \, dy \right)^{1/p}$$

is a norm on the space BMO equivalent to $\| \cdot \|_*$.

Indeed, it suffices to note that we have, by Hölder's inequality (Theorem 3.14),

$$\begin{aligned} \frac{1}{|Q|} \int_Q |f(y) - f_Q| \, dy &\le \frac{1}{|Q|} \left(\int_Q |f(y) - f_Q|^p \, dy \right)^{1/p} \left(\int_Q 1^q \, dy \right)^{1/q} \\ &= \left(\frac{1}{|Q|} \int_Q |f(y) - f_Q|^p \, dy \right)^{1/p} \\ &\le \|f\|_{*,p} \, . \end{aligned}$$

This proves that the reverse inequality

$$\|f\|_* = \sup_Q \frac{1}{|Q|} \int_Q |f(y) - f_Q| \, dy \le \|f\|_{*,p}$$

holds true.

As a consequence of the proof of Theorem 4.11, we obtain the following two additional results:

Corollary 4.12. *Let $f \in$ BMO. Then we have, for all $0 < \lambda < C_2 / \|f\|_*$,*

$$\frac{1}{|Q|} \int_Q e^{\lambda |f(y) - f_Q|} \, dy < \infty.$$

Proof. If we expand the exponential function as a power series, we obtain from Theorem 3.29 and the John–Nirenberg inequality (4.24) that

$$\begin{aligned} \int_Q e^{\lambda |f(y) - f_Q|} \, dy &= |Q| + \sum_{p=1}^{\infty} \frac{\lambda^p}{p!} \int_Q |f(y) - f_Q|^p \, dy \\ &= |Q| + \sum_{p=1}^{\infty} \frac{\lambda^p}{p!} \int_0^{\infty} p \, t^{p-1} \, |E_t| \, dt \end{aligned}$$

$$
\begin{aligned}
&= |Q| + \sum_{p=1}^{\infty} \lambda \int_0^\infty \frac{(\lambda t)^{p-1}}{(p-1)!} \, |E_t| \, dt \\
&= |Q| + \int_0^\infty \lambda \, e^{\lambda t} \, |E_t| \, dt \\
&\le |Q| \left(1 + \int_0^\infty \lambda \, e^{\lambda t} \left(C_1 \, e^{-C_2 t/\|f\|_*} \right) dt \right) \\
&= |Q| \left(1 + \frac{\lambda C_1}{C_2/\|f\|_* - \lambda} \right).
\end{aligned}
$$

This proves that

$$
\frac{1}{|Q|} \int_Q e^{\lambda |f(y) - f_Q|} \, dy \le 1 + \frac{\lambda C_1}{C_2/\|f\|_* - \lambda}.
$$

The proof of Corollary 4.12 is complete. □

The proof of Theorem 4.11 with $p := 1$ can be used to prove that the converse to the John–Nirenberg inequality (Theorem 4.10) holds true:

Corollary 4.13. *Let $f \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$. Assume that there exist positive constants C_1, C_2 and K such that we have, for any cube Q and any $t > 0$,*

$$
|\{x \in Q : |f(x) - f_Q| > t\}| \le C_1 \, e^{-C_2 t/K} \, |Q|. \tag{4.26}
$$

Then it follows that $f \in \mathrm{BMO}$.

Proof. Indeed, we have, by Theorem 3.29 and inequality (4.26),

$$
\begin{aligned}
\int_Q |f(y) - f_Q| \, dy = \int_0^\infty |E_t| \, dt &\le C_1 \left(\int_0^\infty e^{-C_2 t/K} \, dt \right) |Q| \\
&= \frac{C_1 \, K}{C_2} \, |Q| \, .
\end{aligned}
$$

Hence it follows that, for any cube Q,

$$
\frac{1}{|Q|} \int_Q |f(y) - f_Q| \, dy \le \frac{C_1 \, K}{C_2}.
$$

This proves that

$$
\|f\|_* \le \frac{C_1 \, K}{C_2}.
$$

The proof of Corollary 4.13 is complete. □

4.6 The Sharp Function and the Space BMO

If $f(x) \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$, we define the *sharp function* $f^\sharp(x)$ by the formula

$$f^\sharp(x) := \sup_{x\in Q} \frac{1}{|Q|} \int_Q |f(y) - f_Q| \, dy,$$

where the supremum is taken over all cubes Q containing x and $|Q|$ is the Lebesgue measure of Q. It should be noticed that we can take only those cubes Q for which x is an interior point.

If we define the sharp function $f^{\sharp\prime}(x)$ by the formula

$$f^{\sharp\prime}(x) := \sup_{x\in B} \frac{1}{|B|} \int_B |f(y) - f_B| \, dy,$$

where the supremum is taken over all balls B containing x, then, by arguing as in the proof of Lemma 4.1 we can obtain the inequalities

$$\frac{V_n}{2^n} f^{\sharp\prime}(x) \le f^\sharp(x) \le \frac{V_n}{2^{n/2}} f^{\sharp\prime}(x), \quad x \in \mathbf{R}^n,$$

where $V_n = \omega_n/n = \pi^{n/2}/\Gamma(n/2+1)$ is the volume of the unit ball in $\mathbf{R}^n$.

We begin by proving the following:

Claim 4.3. For the maximal functions and sharp functions, we have the useful inequalities

$$f^\sharp(x) \le 2Mf(x), \quad x \in \mathbf{R}^n. \tag{4.27}$$

$$(|f|)^\sharp(x) \le 2f^\sharp(x), \quad x \in \mathbf{R}^n. \tag{4.28}$$

Proof. (i) First, we remark that

$$|f(y) - f_Q| \le |f(y)| + |f_Q|,$$

$$|f_Q| \le \frac{1}{|Q|} \int_Q |f(y)| \, dy \le Mf(x).$$

Hence we have, for all cubes Q containing x,

$$\frac{1}{|Q|} \int_Q |f(y) - f_Q| \, dy \le \frac{1}{|Q|} \int_Q |f(y)| \, dy + |f_Q| \le 2Mf(x).$$

This proves the desired inequality (4.27).

(ii) Similarly, we remark that

$$\begin{aligned} ||f(y)| - (|f|)_Q| &\le ||f(y)| - (|f_Q|)| + |(|f_Q|) - (|f|)_Q| \\ &\le |f(y) - f_Q| + |(|f_Q|) - (|f|)_Q|, \end{aligned} \tag{4.29}$$

where

$$f_Q = \frac{1}{|Q|} \int_Q f(y)\, dy,$$

$$(|f|)_Q = \frac{1}{|Q|} \int_Q |f(y)|\, dy.$$

However, it follows that

$$\begin{aligned} |(|f|)_Q - |f_Q|| &= \frac{1}{|Q|} \left| \int_Q (|f(y)| - |f_Q|)\, dy \right| \\ &\le \frac{1}{|Q|} \int_Q ||f(y)| - |f_Q||\, dy \\ &\le \frac{1}{|Q|} \int_Q |f(y) - f_Q|\, dy. \end{aligned} \tag{4.30}$$

Therefore, by combining inequalities (4.29) and (4.30) we obtain that, for all cubes Q containing x,

$$\begin{aligned} \frac{1}{|Q|} \int_Q ||f(y)| - (|f|)_Q|\, dy &\le \frac{1}{|Q|} \int_Q |f(y) - f_Q|\, dy + ||f_Q| - (|f|)_Q| \\ &\le \frac{2}{|Q|} \int_Q |f(y) - f_Q|\, dy \\ &\le 2 f^\sharp(x). \end{aligned}$$

This proves the desired inequality (4.28).

The proof of Claim 4.3 is complete. □

Secondly, it should be noticed that a function f is in BMO exactly when $f^\sharp$ is a bounded function in $\mathbf{R}^n$. More precisely, it is easy to verify that

$$\mathrm{BMO} = \left\{ f \in L^1_{\mathrm{loc}}(\mathbf{R}^n) : f^\sharp \in L^\infty(\mathbf{R}^n) \right\},$$

and further that

$$\|f\|_* = \left\|f^\sharp\right\|_{L^\infty(\mathbf{R}^n)} = \sup_Q \frac{1}{|Q|} \int_Q |f(y) - f_Q|\, dy.$$

In view of inequality (4.27), we obtain from Theorem 4.8 that, for $1 < p < \infty$,

$$\left\|f^\sharp\right\|_{L^p(\mathbf{R}^n)} \le 2\, \|Mf\|_{L^p(\mathbf{R}^n)} \le 2C_p\, \|f\|_{L^p(\mathbf{R}^n)}.$$

On the other hand, by inequality (4.17) it follows that

$$|f(x)| \le Mf(x) \quad \text{for almost every point } x \text{ of } \mathbf{R}^n,$$

so that

$$\|f\|_{L^p(\mathbf{R}^n)} \leq \|Mf\|_{L^p(\mathbf{R}^n)}.$$

Hence, we have the inequality

$$\left\|f^\sharp\right\|_{L^p(\mathbf{R}^n)} \leq 2C_p \|Mf\|_{L^p(\mathbf{R}^n)}.$$

The next theorem asserts that the converse inequality holds true (see [68, Chapter IV, Section 2.2, Theorem 2]):

Theorem 4.14. *Let $1 < p < \infty$. If $f \in L^p(\mathbf{R}^n)$, then we have the inequality*

$$\|Mf\|_{L^p(\mathbf{R}^n)} \leq A_p \left\|f^\sharp\right\|_{L^p(\mathbf{R}^n)}, \tag{4.31}$$

with a positive constant A_p. For example, we may take

$$A_p := 2\left(2 \cdot 3^n\right)^{1/p} 4^{n+1} 2^{(n+1)p}.$$

Proof. The proof is divided into four steps.

Step I: It suffices to show that

$$\|M(|f|)\|_{L^p(\mathbf{R}^n)} \leq c_p \left\|(|f|)^\sharp\right\|_{L^p(\mathbf{R}^n)}. \tag{4.31$'$}$$

Indeed, since we have, by Claim 4.3,

$$\begin{aligned} &Mf(x) = M(|f|)(x), \quad x \in \mathbf{R}^n, \\ &(|f|)^\sharp(x) \leq 2f^\sharp(x), \quad x \in \mathbf{R}^n, \end{aligned}$$

it follows from inequality (4.31′) that

$$\|Mf\|_{L^p(\mathbf{R}^n)} = \|M(|f|)\|_{L^p(\mathbf{R}^n)} \leq c_p \left\|(|f|)^\sharp\right\|_{L^p(\mathbf{R}^n)} \leq 2c_p \left\|f^\sharp\right\|_{L^p(\mathbf{R}^n)}.$$

This proves the desired inequality (4.31) with $A_p := 2c_p$.

Step II: Therefore, we may assume that

$$f(x) \geq 0 \quad \text{on } \mathbf{R}^n.$$

Our proof is based on the Calderón–Zygmund decomposition (Theorem 4.7). First, we see that this decomposition can be carried out for our function f. Let $t > 0$ and assume that Q is a cube such that

$$f_Q = \frac{1}{|Q|}\int_Q f(y)\,dy > t. \tag{4.32}$$

Then we have, for every $x \in Q$,

$$t < \frac{1}{|Q|}\int_Q f(y)\,dy \leq Mf(x),$$

and so

$$t^p \le \frac{1}{|Q|} \int_Q (Mf(x))^p\,dx \le \frac{1}{|Q|} \int_{\mathbf{R}^n} (Mf(x))^p\,dx.$$

This proves that

$$|Q| \le \frac{1}{t^p} \int_{\mathbf{R}^n} (Mf(x))^p\,dx. \tag{4.33}$$

Hence, if $Q_1 \subset Q_2 \subset \cdots \subset Q_k \subset \cdots$ is an increasing family of dyadic cubes such that

$$f_{Q_k} = \frac{1}{|Q_k|} \int_{Q_k} f(y)\,dy > t,$$

then it follows from inequality (4.33) that the family is necessarily finite. Therefore, we find that each dyadic cube Q satisfying condition (4.32) is contained in a maximal one. If $\{Q_{t,j}\}$ is a family of these maximal dyadic cubes, it follows from an application of the Calderón–Zygmund decomposition (Theorem 4.7) that

$$t < f_{Q_{t,j}} = \frac{1}{|Q_{t,j}|} \int_{Q_{t,j}} f(y)\,dy \le 2^n t.$$
$$f(x) \le t \quad \text{for almost every } x \in \mathbf{R}^n \setminus \textstyle\bigcup_j Q_{t,j}.$$

It should be noticed that if $t < s$, then each $Q_{s,j}$ is contained in some $Q_{t,k}$, since we have the inequality

$$f_{Q_{s,j}} = \frac{1}{|Q_{s,j}|} \int_{Q_{s,j}} f(y)\,dy > s > t.$$

For any given $t > 0$, we take

$$Q_0 := Q_{t/(2^{n+1}),j_0},$$
$$A := 2^{(n+1)p+2}.$$

Then there are two possibilities:

(A) $Q_0 \subset \{x \in \mathbf{R}^n : f^\sharp(x) > t/A\}$.
(B) $Q_0 \not\subset \{x \in \mathbf{R}^n : f^\sharp(x) > t/A\}$.

In the first case (A), we have the inequality

$$\sum_{\{j : Q_{t,j} \subset Q_0\}} |Q_{t,j}| \le \left| \left\{ x \in \mathbf{R}^n : f^\sharp(x) > \frac{t}{A} \right\} \right|. \tag{4.34}$$

In the second case (B), we can find a point $x_0 \in Q_0$ such that

$$\frac{1}{|Q_0|} \int_{Q_0} |f(y) - f_{Q_0}|\, dy \leq f^\sharp(x_0) \leq \frac{t}{A}.$$

Here we remark the following inequalities:

$$\frac{t}{2^{n+1}} < f_{Q_0} = \frac{1}{|Q_0|} \int_{Q_0} f(y)\, dy \leq 2^n \cdot \frac{t}{2^{n+1}} = \frac{t}{2}.$$

$$f_{Q_{t,j}} = \frac{1}{|Q_{t,j}|} \int_{Q_{t,j}} f(y)\, dy > t.$$

Then, since we have the inequality

$$\begin{aligned}
\frac{t}{2}|Q_{t,j}| &= \left(t - \frac{t}{2}\right)|Q_{t,j}| \\
&\leq \int_{Q_{t,j}} f(y)\, dy - f_{Q_0}|Q_{t,j}| = \int_{Q_{t,j}} (f(y) - f_{Q_0})\, dy \\
&\leq \int_{Q_{t,j}} |f(y) - f_{Q_0}|\, dy,
\end{aligned}$$

we obtain that

$$\begin{aligned}
\frac{t}{2} \sum_{\{j : Q_{t,j} \subset Q_0\}} |Q_{t,j}| &\leq \sum_{\{j : Q_{t,j} \subset Q_0\}} \int_{Q_{t,j}} |f(y) - f_{Q_0}|\, dy \\
&\leq \int_{Q_0} |f(y) - f_{Q_0}|\, dy \leq \frac{t}{A}|Q_0|.
\end{aligned}$$

This proves that

$$\sum_{\{j : Q_{t,j} \subset Q_0\}} |Q_{t,j}| \leq \frac{2}{A}|Q_0|. \tag{4.35}$$

Therefore, by combining inequalities (4.34) and (4.35) we have proved that

$$\sum_j |Q_{t,j}| \leq \left|\left\{x \in \mathbf{R}^n : f^\sharp(x) > \frac{t}{A}\right\}\right| + \frac{2}{A} \sum_k \left|Q_{t/(2^{n+1}),k}\right|. \tag{4.36}$$

Step III: If we introduce a function

$$\alpha(t) := \sum_j |Q_{t,j}|,$$

then we can rewrite inequality (4.36) as follows:

$$\alpha(t) \leq \left|\left\{x \in \mathbf{R}^n : f^\sharp(x) > \frac{t}{A}\right\}\right| + \frac{2}{A}\alpha\left(\frac{t}{2^{n+1}}\right). \tag{4.37}$$

If N is an arbitrary positive integer, we let

$$I_N := \int_0^N p\, t^{p-1}\, \alpha(t)\, dt.$$

In order to estimate the integral I_N, we need the following claim:

Claim 4.4. If we introduce the *distribution function* of Mf by the formula

$$\beta(t) := |\{x \in \mathbf{R}^n : Mf(x) > t\}|\,,$$

then we have the inequalities

$$\alpha(t) \leq \beta(t), \tag{4.38a}$$

$$\beta(t) \leq 3^n\, \alpha\left(\frac{t}{4^n}\right). \tag{4.38b}$$

Proof. (i) Since we have, for any $x \in Q_{t,j}$,

$$t < \frac{1}{|Q_{t,j}|} \int_{Q_{t,j}} f(y)\, dy \leq Mf(x),$$

it follows that

$$\bigcup_j Q_{t,j} \subset \{x \in \mathbf{R}^n : Mf(x) > t\}\,.$$

This proves that

$$\alpha(t) = \sum_j |Q_{t,j}| \leq |\{x \in \mathbf{R}^n : Mf(x) > t\}| = \beta(t).$$

(ii) Just as in the proof of Theorem 4.7, we obtain that

$$\begin{aligned}\beta(t) = |\{x \in \mathbf{R}^n : Mf(x) > t\}| &\leq \sum_j \left|3Q_{t/4^n,j}\right| = 3^n \sum_j \left|Q_{t/4^n,j}\right| \\ &\leq 3^n\, \alpha\left(\frac{t}{4^n}\right).\end{aligned}$$

The proof of Claim 4.4 is complete. □

First, by inequality (4.38a) and Theorem 3.29 it follows that

$$\begin{aligned}I_N &= \int_0^N p\, t^{p-1}\, \alpha(t)\, dt \leq \int_0^N p\, t^{p-1}\, \beta(t)\, dt \\ &\leq \int_0^\infty p\, t^{p-1}\, \beta(t)\, dt = \int_{\mathbf{R}^n} Mf(x)^p\, dx = \|Mf\|^p_{L^p(\mathbf{R}^n)}\,.\end{aligned}$$

This proves that the finite limit

$$\lim_{N\to\infty} I_N = \int_0^\infty p\, t^{p-1}\,\alpha(t)\,dt \tag{4.39}$$

exists.

Secondly, we obtain from inequality (4.37) that

$$\begin{aligned} I_N &\le \int_0^N p\,t^{p-1}\left|\left\{x\in \mathbf{R}^n : f^\sharp(x) > \frac{t}{A}\right\}\right| dt \\ &\quad + \frac{2}{A}\int_0^N p\,t^{p-1}\alpha\left(\frac{t}{2^{n+1}}\right) dt \\ &= \int_0^N p\,t^{p-1}\left|\left\{x\in \mathbf{R}^n : f^\sharp(x) > \frac{t}{A}\right\}\right| dt \\ &\quad + \frac{2}{A}\int_0^{N/(2^{n+1})} p\,(2^{n+1})^{p-1}\, s^{p-1}\alpha(s)\, 2^{n+1}\, ds \\ &\le \int_0^N p\,t^{p-1}\left|\left\{x\in \mathbf{R}^n : f^\sharp(x) > \frac{t}{A}\right\}\right| dt + \frac{1}{2}\int_0^N p\, s^{p-1}\alpha(s)\, ds \\ &= \int_0^N p\,t^{p-1}\left|\left\{x\in \mathbf{R}^n : f^\sharp(x) > \frac{t}{A}\right\}\right| dt + \frac{1}{2} I_N, \end{aligned} \tag{4.40}$$

since we have chosen

$$A = 2^{(n+1)p+2}.$$

Hence, by inequality (4.40) it follows that

$$I_N \le 2\int_0^N p\,t^{p-1}\left|\left\{x\in \mathbf{R}^n : f^\sharp(x) > \frac{t}{A}\right\}\right| dt.$$

Therefore, in view of assertion (4.39) we can let $N \to \infty$ to obtain that

$$\int_0^\infty p\,t^{p-1}\,\alpha(t)\,dt \le 2\int_0^\infty p\,t^{p-1}\left|\left\{x\in \mathbf{R}^n : f^\sharp(x) > \frac{t}{A}\right\}\right| dt. \tag{4.41}$$

Step IV: By inequalities (4.38b) and (4.41), it follows from an application of Theorem 3.29 that

$$\begin{aligned} &\int_{\mathbf{R}^n} Mf(x)^p\, dx = \int_0^\infty p\,t^{p-1}\,\beta(t)\,dt \\ &\le 3^n \int_0^\infty p\,t^{p-1}\,\alpha\left(\frac{t}{4^n}\right) dt = 3^n\int_0^\infty p\,(4^n)^{p-1}\, s^{p-1}\,\alpha(s)\, 4^n\, ds \\ &= 3^n\cdot 4^{np}\left(\int_0^\infty p\, s^{p-1}\,\alpha(s)\,ds\right) \end{aligned}$$

$$\leq 3^n \cdot 4^{np} \left(2 \int_0^\infty p\, t^{p-1} \left| \left\{ x \in \mathbf{R}^n : f^\sharp(x) > \frac{t}{A} \right\} \right| dt \right)$$
$$= 2 \cdot 3^n \cdot 4^{np} \int_0^\infty p\, A^{p-1}\, \sigma^{p-1} \left| \left\{ x \in \mathbf{R}^n : f^\sharp(x) > \sigma \right\} \right| A\, d\sigma$$
$$= 2 \cdot 3^n \cdot 4^{np} \cdot A^p \int_0^\infty p\, \sigma^{p-1} \left| \left\{ x \in \mathbf{R}^n : f^\sharp(x) > \sigma \right\} \right| d\sigma$$
$$= 2 \cdot 3^n \cdot 4^{np} \cdot A^p \int_{\mathbf{R}^n} f^\sharp(x)^p\, dx.$$

Summing up, we have proved the desired inequality

$$\|Mf\|_{L^p(\mathbf{R}^n)} \leq c_p \left\| f^\sharp \right\|_{L^p(\mathbf{R}^n)}, \tag{4.31$'$}$$

with

$$c_p := \left(2 \cdot 3^n\right)^{1/p} 4^{n+1}\, 2^{(n+1)p}.$$

Now the proof of Theorem 4.14 is complete. □

Corollary 4.15. *Let* $1 < p < \infty$*. If* $f \in L^p(\mathbf{R}^n)$*, then we have the inequality*

$$\|f\|_{L^p(\mathbf{R}^n)} \leq A_p \left\| f^\sharp \right\|_{L^p(\mathbf{R}^n)},$$

with a positive constant A_p*.*

Corollary 4.15 is an immediate consequence of inequality (4.17) and Theorem 4.14.

Remark 4.3. Theorem 4.14 and Corollary 4.15 remain valid if the norm is taken on the half-space $\mathbf{R}^n_+$. In this case, the definition of sharp functions need to be modified by taking only cubes Q contained in $\mathbf{R}^n_+$ (see Figure 4.4).

4.7 Spherical Harmonics

In this section we introduce spherical harmonics in order to prove the fundamental results of singular integrals in Chapter 10 and Chapter 11. The presentation here is a slightly expanded version of Neri [51] (see also [6, Chapter 5]).

If m is a non-negative integer, then we denote by Π_m the set of all polynomials in the variables $x = (x_1, x_2, \ldots, x_n)$, $n \geq 2$, which are homogeneous of degree m. A typical example of homogeneous polynomials of degree 2 is given by the formula

$$|x|^2 = x_1^2 + x_2^2 + \cdots + x_n^2.$$

For simplicity, we assume throughout this section that our polynomials have real coefficients, the complex case being of course identical except for the occasional presence of complex conjugates.

We begin by proving the following fundamental result:

Lemma 4.16. *The space Π_m is a finite dimensional real vector space. Moreover, if we denote by $g(m)$ the dimension of the vector space Π_m, then we have the formula*

$$g(m) = \binom{m+n-1}{n-1} = \frac{(n+m-1)!}{m!(n-1)!}.$$

Proof. (1) The first statement is obvious.

(2) Since the monomials x^α with $|\alpha| = m$ form a basis in the vector space Π_m, it follows that the dimension $g(m)$ is equal to the number of such monomials. However, it should be noticed that these monomials occur as coefficients in the formula

$$\prod_{j=1}^{n} (1 - x_j\, t)^{-1} = \prod_{j=1}^{n} \left(1 + x_j\, t + \cdots + x_j^k\, t^k + \cdots\right).$$

This implies that their number is equal to the coefficient of t^m in $(1 - t)^{-n}$. Therefore, we obtain that

$$\begin{aligned} g(m) &= \frac{1}{m!} \frac{d^m}{dt^m} (1-t)^{-n} \Big|_{t=0} \\ &= \frac{1}{m!}\, n(n+1) \cdots (n+m-1)(1-t)^{-(n+m)} \Big|_{t=0} \\ &= \frac{(n+m-1)!}{m!(n-1)!} = \binom{n+m-1}{n-1}. \end{aligned}$$

The proof of Lemma 4.16 is complete. □

We let

$$P(x) = \sum_{|\alpha|=m} a_\alpha x^\alpha \in \Pi_m,$$

and introduce the homogeneous differential polynomial

$$P\left(\frac{\partial}{\partial x}\right) = \sum_{|\alpha|=m} a_\alpha \left(\frac{\partial}{\partial x}\right)^\alpha,$$

which is obtained from $P(x)$ by replacing each monomial x^α by the corresponding differential monomial $(\partial/\partial x)^\alpha$. Now we can prove the following:

Lemma 4.17. *Every space Π_m is a real inner product space endowed with the inner product*

$$\langle P, Q \rangle := P\left(\frac{\partial}{\partial x}\right) Q(x), \quad P, Q \in \Pi_m,$$

Proof. First, it should be noticed that we have, for $|\alpha| = |\beta|$,

$$\left(\frac{\partial}{\partial x}\right)^{\alpha} x^{\beta} = \begin{cases} \alpha! & \text{if } \alpha = \beta, \\ 0 & \text{otherwise.} \end{cases}$$

Hence it is easy to verify that the formula $\langle P, Q \rangle$, P, $Q \in \Pi_m$, defines a bilinear form on the product space $\Pi_m \times \Pi_m$. Moreover, if $P(x) = \sum_\alpha a_\alpha x^\alpha$ and $Q(x) = \sum_\beta b_\beta x^\beta$ with $|\alpha| = |\beta| = m$, then we obtain that

$$\langle P, Q \rangle = \sum_{|\alpha|=m} \alpha!\, a_\alpha b_\alpha = \sum_{|\alpha|=m} \alpha!\, b_\alpha a_\alpha = \langle Q, P \rangle\,,$$

$$\langle P, P \rangle = \sum_{|\alpha|=m} \alpha!\, a_\alpha^2 \geq 0,$$

$$\langle P, P \rangle = 0 \text{ if and only if } P(x) \equiv 0.$$

This proves that the quantity $\langle P, Q \rangle$ satisfies the axioms (I1) through (I4) of inner product in Section 2.10.

The proof of lemma 4.16 is complete. □

Let $P \in \Pi_\ell$ and $Q \in \Pi_m$ with $\ell \leq m$. Then it is easy to see that the polynomial

$$\langle P, Q \rangle = P\left(\frac{\partial}{\partial x}\right) Q(x)$$

belongs to the space $\Pi_{m-\ell}$. Hence it follows that the associated differential polynomial $P\,(\partial/\partial x)$ defines a linear operator

$$P\left(\frac{\partial}{\partial x}\right) : \Pi_m \longrightarrow \Pi_{m-\ell}$$

by the formula $\langle P, Q \rangle$. Moreover, we have the following:

Lemma 4.18. *Let $P \in \Pi_\ell$ such that $P(x) \not\equiv 0$, and $m \geq \ell$. Then the linear map $P\left(\frac{\partial}{\partial x}\right) : \Pi_m \to \Pi_{m-\ell}$ is surjective.*

Proof. Our proof is based on a reduction to absurdity. Assume, to the contrary, that the range $P\,(\partial/\partial x)\,\Pi_m$ is a proper subspace V of $\Pi_{m-\ell}$. Then, by using Lemma 4.17 we can find an element $R(x) \not\equiv 0$ in the inner product space $\Pi_{m-\ell}$ which is orthogonal to the space V. Therefore, we have, for all $Q \in \Pi_m$,

$$\left\langle R, P\left(\frac{\partial}{\partial x}\right) Q(x)\right\rangle = R\left(\frac{\partial}{\partial x}\right) P\left(\frac{\partial}{\partial x}\right) Q(x) \equiv 0.$$

In particular, by choosing $Q(x) := R(x)P(x)$ we obtain that $\langle Q, Q\rangle = 0$, so that $Q(x) = R(x)P(x) \equiv 0$. This implies that $R(x) \equiv 0$, since $P(x) \not\equiv 0$. This contradiction proves that $P\left(\partial/\partial x\right) : \Pi_m \to \Pi_{m-\ell}$ is surjective.

The proof of Lemma 4.18 is complete. □

Theorem 4.19 (the decomposition theorem)**.** *Let $P \in \Pi_\ell$ such that $P(x) \not\equiv 0$, and $m \geq \ell$. Then, every element $T \in \Pi_m$ can be decomposed uniquely in the form*

$$T(x) = \sum_k P^k(x) R_k(x) \tag{4.42}$$

where

$$R_k \in \Pi_{m-k\ell},$$
$$P\left(\frac{\partial}{\partial x}\right) R_k(x) \equiv 0,$$

and the summation is taken over all non-negative integers k such that $k\ell \leq m$. Moreover, the $R_k(x) \not\equiv 0$ are not divisible by $P(x)$.

Proof. (1) First, we show that every element $T \in \Pi_m$ can be decomposed uniquely in the form

$$T = PS_1 + R_0,$$

where $S_1 \in \Pi_{m-\ell}$ and $P\left(\partial/\partial x\right) R_0(x) \equiv 0$.

To do this, we introduce the subspace

$$M := \{PS : S \in \Pi_{m-\ell}\}$$

of all elements of Π_m which are divisible by $P(x)$, and the subspace

$$N := \left\{R \in \Pi_m : P\left(\frac{\partial}{\partial x}\right) R(x) \equiv 0\right\}$$

of all elements of Π_m which are annihilated by $P\left(\partial/\partial x\right)$. Then we have the assertion

$$M \cap N = \{0\}.$$

Indeed, if $PS \in M$ belongs to N, then it follows that

$$P\left(\frac{\partial}{\partial x}\right) P(x)S(x) \equiv 0,$$

so that

$$S\left(\frac{\partial}{\partial x}\right) P\left(\frac{\partial}{\partial x}\right) P(x)S(x) = \langle PS, PS\rangle = 0.$$

This implies that $PS = 0$. Hence, we have the formula

$$\dim(M \oplus N) = \dim(M) + \dim(N) = g(m-\ell) + \dim(N).$$

However, by Lemma 4.18 it follows that

$$g(m-\ell) + \dim(N) = \dim\left(\Pi_m\right),$$

so that,

$$\dim(M \oplus N) = \dim\left(\Pi_m\right).$$

Therefore, we have proved that

$$\Pi_m = M \oplus N,$$

and further that each element $T \in \Pi_m$ can be decomposed uniquely in the form

$$T = PS_1 + R_0,$$

with $S_1 \in \Pi_{m-\ell}$ and $R_0 \in N$. Moreover, it should be emphasized that $R_0(x)$ is not divisible by $P(x)$ unless $R_0(x) \equiv 0$, since we have $M \cap N = \{0\}$.

(2) Similarly, by repeating the same procedure for $S_1(x)$ (in place of $T(x)$) we obtain that

$$S_1 = PS_2 + R_1,$$

where $S_2 \in \Pi_{m-2\ell}$ and $R_1(x)$ is annihilated by $P\left(\partial/\partial x\right)$ and not divisible by $P(x)$ (unless $R_1(x) \equiv 0$). We remark that

$$T = PS_1 + R_0 = P\left(PS_2 + R_1\right) + R_0 = P^2S_2 + PR_1 + R_0.$$

(3) This recursive process ends when we reach an element $S_k \in \Pi_{m-k\ell}$ of degree less than ℓ. In this case, it follows that $S_k(x)$ is not divisible by $P(x)$ and that $P\left(\partial/\partial x\right) S_k(x) \equiv 0$. Hence, we may take $S_k = R_k$. This proves the desired decomposition (4.42).

The proof of Theorem 4.19 is complete. □

It should be emphasized that the homogeneous polynomial

$$|x|^2 = x_1^2 + x_2^2 + \cdots + x_n^2$$

of degree 2 corresponds to the usual Laplacian

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \ldots + \frac{\partial^2}{\partial x_n^2}.$$

All polynomials $P \in \Pi_m$ are called *solid harmonics* of degree m if they satisfy the Laplace equation: $\Delta P = 0$ in $\mathbf{R}^n$. The restrictions of these solid harmonics to the unit sphere Σ_{n-1} are called *spherical harmonics* of degree m.

We shall use the following notation:

$$\mathcal{S}_m := \text{all solid harmonics of degree } m,$$
$$\mathcal{H}_m := \text{all spherical harmonics of degree } m.$$

Remark 4.4. (a) By Lemma 4.18, it follows that $\mathcal{S}_m$ is the nullspace of the Laplacian as a linear map of Π_m onto Π_{m-2}. If $m = 0$ and $m = 1$, then we have $\mathcal{S}_m = \Pi_m$. However, if $m \geq 2$, then we obtain that $\mathcal{S}_m$ is a proper subspace Π_m of dimension

$$d(m) := g(m) - g(m-2).$$

This formula remains valid for all non-negative integers m if we let $g(-1) = g(-2) = 0$.

(b) If $S_m \in \mathcal{S}_m$, then, we have, for all $x \neq 0$ and $x' = x/|x|$,

$$S_m(x) = S_m\left(|x| \frac{x}{|x|}\right) = |x|^m\, S_m\left(\frac{x}{|x|}\right) = |x|^m\, Q_m(x'),$$

where $Q_m(x') = S_m(x')$ is the corresponding spherical harmonic in $\mathcal{H}_m$. Since this correspondence is an isomorphism, it follows from assertion (a) that $\mathcal{H}_m$ is a vector space of dimension $d(m) = g(m) - g(m-2)$.

(c) On the Euclidean space $\mathbf{R}^2$, we have, for any integer $m \geq 1$,

$$g(m) = \binom{m+1}{1} = m+1, \quad g(m-2) = m-1.$$

Hence it follows that $d(m) = g(m) = m+1$ for $m = 0$ and 1, and that $d(m) = g(m) - g(m-2) = 2$ for $m \geq 2$. In terms of the complex variable $z = x + iy = re^{i\theta}$, we obtain that $z^m = r^m \cos m\theta + ir^m \sin m\theta$. Therefore, we can conclude that, for every m, $\cos m\theta$ and $\sin m\theta$ are the only linearly independent spherical harmonics on the plane.

On the Euclidean space $\mathbf{R}^3$, it is easy to see that $d(m) = 2m + 1$, so the dimension of $\mathcal{S}_m$, and hence that of $\mathcal{H}_m$, increases as the odd integers.

(d) On the general Euclidean space $\mathbf{R}^n$, we have, as $m \to \infty$,

$$g(m) = \binom{m+n-1}{n-1} = \frac{(m+n-1)(m+n-2)\cdots(m+1)}{(n-1)!} \sim \frac{m^{n-1}}{(n-1)!}.$$

Moreover, by using the mean value theorem we obtain that

$$d(m) = g(m) - g(m-2) \sim \left[\frac{m^{n-1}}{(n-1)!} - \frac{(m-2)^{n-1}}{(n-1)!}\right] \sim 2\frac{m^{n-2}}{(n-2)!}.$$

Summing up, we have proved that, for any dimension $n \geq 2$,

$$d(m) \sim c(n)\, m^{n-2}$$

as $m \to \infty$, where $c(n)$ is a positive constant depending only on the dimension n.

Corollary 4.20. *Any continuous function on the unit sphere Σ_{n-1} can be approximated uniformly by a finite linear combination of spherical harmonics.*

Proof. Let $f(x')$ be an arbitrary continuous function on Σ_{n-1}. We may assume that $f(x')$ is real valued. By the Stone–Weierstrass theorem (Theorem 2.70), it follows that $f(x')$ can be approximated uniformly by the restriction to Σ_{n-1} of some polynomial $T(x)$. However, since $T(x)$ is a finite sum of homogeneous pieces, we have only to consider each piece $T_m \in \Pi_m$.

We remark that the assertion is obvious for $m = 0$ and $m = 1$.

If $m \geq 2$, by applying the decomposition theorem (Theorem 4.19) to $P(x) := |x|^2$ and $P(\partial/\partial x) := \Delta$ we obtain that

$$T_m(x) = \sum_{2k \leq m} |x|^{2k} R_k(x),$$

where $R_k(x)$ are solid harmonics of degree $m - 2k$. Hence, we have, on the unit sphere Σ_{n-1},

$$T_m(x') = \sum_{2k \leq m} R_k(x'),$$

where the $R_k(x')$ belong to $\mathcal{H}_{m-2k}$.

The proof of Corollary 4.20 is complete. □

Lemma 4.21. *Let $Q_j \in \mathcal{H}_j$ and $S_j(x) = |x|^j Q_j(x')$ its corresponding solid harmonics. Then we have, for $k \neq m$,*

$$\int_{\Sigma_{n-1}} Q_k(x')\, Q_m(x')\, d\sigma(x') = 0, \tag{4.43}$$

$$\int_{|x|\le 1} S_k(x)\, S_m(x)\, dx = 0. \tag{4.44}$$

Here $d\sigma$ is the surface measure on the unit sphere Σ_{n-1}.

Proof. Since we have, in terms of polar coordinates,

$$\begin{aligned}
&\int_{|x|\le 1} S_k(x)\, S_m(x)\, dx \\
&= \left(\int_0^1 r^{k+m} r^{n-1}\, dr\right) \int_{\Sigma_{n-1}} Q_k(x')\, Q_m(x')\, d\sigma(x') \\
&= \frac{1}{k+m+n} \int_{\Sigma_{n-1}} Q_k(x')\, Q_m(x')\, d\sigma(x'),
\end{aligned}$$

it suffices to prove formula (4.43) for spherical harmonics.

By using Green's formula (see Theorem 5.3 in Chapter 5), we obtain that

$$\begin{aligned}
&\int_{|x|\le 1} \left(S_k(x)\cdot \Delta S_m(x) - S_m(x)\cdot \Delta S_k(x)\right) dx \\
&= \int_{\Sigma_{n-1}} \left(S_k(x')\, \frac{\partial S_m}{\partial \boldsymbol{\nu}}(x') - S_m(x')\, \frac{\partial S_k}{\partial \boldsymbol{\nu}}\right) d\sigma(x'),
\end{aligned}$$

where $\boldsymbol{\nu}$ is the outward normal to the unit sphere Σ_{n-1}. However, since we have the formula

$$\frac{\partial}{\partial \boldsymbol{\nu}} = \frac{\partial}{\partial |x|},$$

it follows that

$$\frac{\partial}{\partial \boldsymbol{\nu}}\left(S_j(x)\right) = j|x|^{j-1} Q_j(x').$$

Moreover, by the definition of solid harmonics it follows that $\Delta S_j = 0$ in $\mathbf{R}^n$. Therefore, we obtain that

$$\begin{aligned}
0 &= \int_{\Sigma_{n-1}} \left(Q_k(x')\cdot m Q_m(x') - Q_m(x')\cdot k Q_k(x')\right) d\sigma(x') \\
&= (m-k) \int_{\Sigma_{n-1}} Q_k(x')\, Q_m(x')\, d\sigma(x').
\end{aligned}$$

This proves that we have, for $k \neq m$,

$$\int_{\Sigma_{n-1}} Q_k(x')\, Q_m(x')\, d\sigma(x') = 0.$$

The proof of Lemma 4.21 is complete. □

We consider the vector spaces $\mathcal{H}_m$ as linear subspaces of the real Hilbert space $L^2(\Sigma_{n-1})$ with inner product

$$(f, g) = \int_{\Sigma_{n-1}} f(x')\, g(x')\, d\sigma(x').$$

With respect to this inner product, we can construct an orthonormal basis $\{Y_{\ell m}\}$, $1 \leq \ell \leq d(m)$, in each space $\mathcal{H}_m$. Hence, we have, for each m,

$$(Y_{im}, Y_{jm}) = \int_{\Sigma_{n-1}} Y_{im}(x')\, Y_{jm}(x')\, d\sigma(x') = \delta_{ij}, \quad 1 \leq i, j \leq d(m).$$

Moreover, we have the following theorem (cf. Theorem 2.64):

Theorem 4.22. *(i) The Hilbert space $L^2(\Sigma_{n-1})$ can be decomposed into the infinite direct sum in the sense of Hilbert space theory:*

$$L^2(\Sigma_{n-1}) = \sum_{m=0}^{\infty} \mathcal{H}_m.$$

More precisely, any function $f \in L^2(\Sigma_{n-1})$ has the development

$$f(x') = \sum_{m=0}^{\infty} Y_m(x'), \quad Y_m(x') \in \mathcal{H}_m,$$

where the convergence is in the $L^2(\Sigma_{n-1})$ norm, and we have the formula

$$\int_{\Sigma_{n-1}} |f(x')|^2\, d\sigma(x') = \sum_{m=0}^{\infty} \int_{\Sigma_{n-1}} |Y_m(x')|^2\, d\sigma(x').$$

(ii) Furthermore, the family $\{Y_{\ell m}\}_{\substack{m=0,1,\ldots,\\ \ell=1,2,\ldots,d(m)}}$ forms a complete orthonormal system in the Hilbert space $L^2(\Sigma_{n-1})$. More precisely, any function $f \in L^2(\Sigma_{n-1})$ has the Fourier series expansion with respect to $\{Y_{\ell m}\}$

$$f(x') = \sum_{m=0}^{\infty} \sum_{\ell=1}^{d(m)} a_{\ell m}\, Y_{\ell m}(x'),$$

where the Fourier coefficients $a_{\ell m}$ *are given by the formulas*

$$a_{\ell m} = \int_{\Sigma_{n-1}} f(y)\, Y_{\ell m}(y)\, d\sigma(y),$$

and satisfy the Parseval identity

$$\sum_{m=0}^{\infty} \sum_{\ell=1}^{d(m)} |a_{\ell m}|^2 = \int_{\Sigma_{n-1}} |f(x')|^2\, d\sigma(x').$$

Proof. Since part (i) is an immediate consequence of part (ii), we have only to prove part (ii).

First, by Lemma 4.21 it follows that spherical harmonics of distinct degrees are orthogonal, so that the family $\{Y_{\ell m}\}$ is orthonormal. On the other hand, since the continuous functions are dense in $L^2(\Sigma_{n-1})$, it follows from an application of Corollary 4.20 that the set of all finite linear combinations of the $Y_{\ell m}$ is dense in $L^2(\Sigma_{n-1})$. This implies that the orthonormal system $\{Y_{\ell m}\}$ is complete in the Hilbert space $L^2(\Sigma_{n-1})$.

The proof of Theorem 4.22 is complete. □

Now we prove some important bounds on the spherical harmonics $\{Y_{\ell m}\}$ and their derivatives. These bounds can be deduced from the fact that we have, for each integer $m \geq 0$,

$$\sum_{\ell=1}^{d(m)} Y_{\ell m}^2(x') = \frac{d(m)}{\omega_n}, \quad x' \in \Sigma_{n-1}, \tag{4.45}$$

where

$$d(m) = g(m) - g(m-2),$$

and

$$\omega_n := |\Sigma_{n-1}| = \frac{2\pi^{n/2}}{\Gamma(n/2)}$$

is the surface area of the unit sphere Σ_{n-1}. In order to prove formula (4.45), we introduce the concept of zonal harmonics:

Lemma 4.23. *For each* m *and each point* $x' \in \Sigma_{n-1}$*, we can find a spherical harmonic* $Z_{x'} \in \mathcal{H}_m$*, called a* zonal harmonic *with pole* x'*, such that we have, for all* $Q \in \mathcal{H}_m$*,*

$$Q(x') = (Q, Z_{x'}). \tag{4.46}$$

Moreover, we have, for all $y \in \Sigma_{n-1}$,

$$Z_{x'}(y) = \sum_{\ell=1}^{d(m)} Y_{\ell m}(x')\, Y_{\ell m}(y). \tag{4.47}$$

Proof. (1) First, we remark that the map $Q \to Q(x')$ is a linear functional over the finite dimensional Hilbert space $\mathcal{H}_m$. Hence, the desired formula (4.46) follows from an application of the Riesz representation theorem (Theorem 2.58).

(2) Secondly, since the $\{Y_{\ell m}\}$, $1 \leq \ell \leq d(m)$, form an orthonormal basis in each space $\mathcal{H}_m$, we obtain that any element $Q \in \mathcal{H}_m$ can be written uniquely in the form

$$Q(x') = \sum_{\ell=1}^{d(m)} a_\ell\, Y_{\ell m}(x'), \quad x' \in \Sigma_{n-1}, \tag{4.48}$$

where the Fourier coefficients a_ℓ are given by the formulas

$$a_\ell := a_{\ell m} = (Q, Y_{\ell m}) = \int_{\Sigma_{n-1}} Q(y)\, Y_{\ell m}(y)\, d\sigma(y).$$

Hence we have, by formula (4.46) and formula (4.48),

$$(Q, Z_{x'}) = \sum_{\ell=1}^{d(m)} (Q, Y_{\ell m}) Y_{\ell m}(x') = \left(Q, \sum_{\ell=1}^{d(m)} Y_{\ell m}(x')\, Y_{\ell m} \right). \tag{4.49}$$

This proves the desired formula (4.47), since formula (4.49) holds true for all Q in $\mathcal{H}_m$.

The proof of Lemma 4.23 is complete. □

Lemma 4.24. *If $u \colon \Sigma_{n-1} \to \Sigma_{n-1}$ is a rotation, then we have, for any zonal harmonic $Z_{x'}$,*

$$Z_{ux'}(uy) = Z_{x'}(y), \quad y \in \Sigma_{n-1}. \tag{4.50}$$

Proof. If the function $Q(x')$ belongs to $\mathcal{H}_m$, then so does the function $Q(ux')$, since the Laplacian Δ is invariant under rotations. By applying formula (4.46) to the harmonic $Q_u(x') := Q(ux')$, we obtain that

$$\begin{aligned} Q(ux') = Q_u(x') &= \int_{\Sigma_{n-1}} Q_u(y)\, Z_{x'}(y)\, d\sigma(y) \qquad (4.51) \\ &= \int_{\Sigma_{n-1}} Q(uy)\, Z_{x'}(y)\, d\sigma(y). \end{aligned}$$

On the other hand, by applying again formula (4.46) to the harmonic $Q(x')$ and by changing variables we obtain that

$$Q(ux') \tag{4.52}$$
$$= \int_{\Sigma_{n-1}} Q(y)\, Z_{ux'}(y)\, d\sigma(y) = \int_{\Sigma_{n-1}} Q(uy)\, Z_{ux'}(uy)\, d\sigma(y),$$

since the surface measure $d\sigma$ is also invariant under rotations.

Therefore, it follows from formula (4.51) and formula (4.52) that we have, for all $Q \in \mathcal{H}_m$,

$$\int_{\Sigma_{n-1}} Z_{x'}(y)\, Q(uy)\, d\sigma(y) = \int_{\Sigma_{n-1}} Z_{ux'}(uy)\, Q(uy)\, d\sigma(y).$$

This proves the desired formula (4.50).

The proof of Lemma 4.24 is complete. □

Remark 4.5. If the pole $x' \in \Sigma_{n-1}$ of a zonal harmonic $Z_{x'}$ lies on the axis of a rotation u, then we have $ux' = x'$ and so, by Lemma 4.24,

$$Z_{x'}(y) = Z_{x'}(uy) \quad \text{for } y \in \Sigma_{n-1}.$$

In other words, a zonal harmonic $Z_{x'}$ is constant along each "parallel" relative to the pole x'.

The next lemma proves the desired formula (4.45):

Lemma 4.25. *For any integer $m \geq 0$, we have the formula*

$$\sum_{\ell=1}^{d(m)} Y_{\ell m}^2(x') = \frac{d(m)}{\omega_n} \quad \textit{for } x' \in \Sigma_{n-1}. \tag{4.45}$$

Proof. By formula (4.46), we have, for any $x' \in \Sigma_{n-1}$,

$$Z_{x'}(x') = \sum_{\ell=1}^{d(m)} Y_{\ell m}^2(x'),$$

and we have, for any rotation u,

$$Z_{ux'}(ux') = \sum_{\ell=1}^{d(m)} Y_{\ell m}^2(ux').$$

Hence, by applying Lemma 4.24 to our situation we obtain that

$$\sum_{\ell=1}^{d(m)} Y_{\ell m}^2(x') = \sum_{\ell=1}^{d(m)} Y_{\ell m}^2(ux') \quad \text{for } x' \in \Sigma_{n-1}. \tag{4.53}$$

Since each point y of Σ_{n-1} is the image $y = ux'$ under a suitable rotation u, we find from formula (4.53) that

$$\sum_{\ell=1}^{d(m)} Y_{\ell m}^2(x') = A_m \quad \text{for } x' \in \Sigma_{n-1}, \tag{4.54}$$

where A_m is a constant depending only on the degree m and the dimension n.

Hence, by integrating formula (4.54) over Σ_{n-1} we obtain that

$$\int_{\Sigma_{n-1}} \sum_{\ell=1}^{d(m)} Y_{\ell m}^2(x')\, d\sigma(x') = A_m \int_{\Sigma_{n-1}} d\sigma(x') = A_m \omega_n. \tag{4.55}$$

On the other hand, since the $\{Y_{\ell m}\}$, $1 \le \ell \le d(m)$, form an orthonormal basis in the Hilbert space $L^2(\Sigma_{n-1})$, it follows that

$$\begin{aligned}\int_{\Sigma_{n-1}} \sum_{\ell=1}^{d(m)} Y_{\ell m}^2(x')\, d\sigma(x') &= \sum_{\ell=1}^{d(m)} \int_{\Sigma_{n-1}} Y_{\ell m}(x')\, Y_{\ell m}(x')\, d\sigma(x') \\ &= d(m).\end{aligned} \tag{4.56}$$

Therefore, by combining formula (4.55) and formula (4.56) we have proved that

$$A_m = \frac{d(m)}{\omega_n}.$$

The proof of Lemma 4.25 is complete. □

Now we can prove the desired bounds on the spherical harmonics $Y_{\ell m}$:

Theorem 4.26. *(i) For all $x' \in \Sigma_{n-1}$, we have the estimate*

$$|Y_{\ell m}(x')| \le C_1\, m^{(n-2)/2}, \tag{4.57}$$

where C_1 is a positive constant depending only on the dimension n.

(ii) More generally, we have, for any multi-index α,

$$\left| \left(\frac{\partial}{\partial x} \right)^{\alpha} (|x|^m Y_{\ell m}(x')) \right| \le C_2\, m^{(n-2)/2+|\alpha|} |x|^{m-|\alpha|}, \tag{4.58}$$

where C_2 is a positive constant depending only on $|\alpha|$ and n.

Proof. (i) By Lemma 4.25 and assertion (d) of Remark 4.4, it follows that we have, for some positive constant C depending only on n,

$$Y_{\ell m}^2(x') \le \sum_{\ell=1}^{d(m)} Y_{\ell m}^2(x') = \frac{d(m)}{\omega_n} \le C\, m^{n-2} \quad \text{for } x' \in \Sigma_{n-1}.$$

Hence, the desire estimate (4.57) follows by taking the square root of this inequality.

(ii) First, we remark that estimate (4.58) for $|\alpha| = 0$ is reduced to estimate (4.57).

Secondly, we consider the case where $|\alpha| = 1$. To do this, we let

$$P(x) := |x|^m \, Y_{\ell m}(x'),$$

and estimate the first partial derivatives $\partial P/\partial x_j$.

Since we have the assertion

$$P \in \mathcal{S}_m,$$

and so

$$\Delta \left(\frac{\partial P}{\partial x_j} \right) = \frac{\partial}{\partial x_j} \, (\Delta P) = 0 \quad \text{in } \mathbf{R}^n,$$

it follows that

$$\frac{\partial P}{\partial x_j} \in \mathcal{S}_{m-1}.$$

Hence, by applying Lemma 4.25 to the functions $\partial P/\partial x_j$ and by using Schwarz's inequality (Theorem 3.14 with $p = q := 2$) we obtain that

$$\begin{aligned} \left| \frac{\partial P}{\partial x_j}(x') \right|^2 &= \left| \sum_{\ell=1}^{d(m-1)} b_\ell \, Y_{\ell m-1}(x') \right|^2 \\ &\leq \left(\sum_{\ell=1}^{d(m-1)} b_\ell^2 \right) \left(\sum_{\ell=1}^{d(m-1)} Y_{\ell m-1}^2(x') \right) \\ &= A \, \frac{d(m-1)}{\omega_n}, \end{aligned}$$

where the Fourier coefficients b_ℓ are given by the formulas

$$\begin{aligned} b_\ell &:= \left(\frac{\partial P}{\partial x_j}, Y_{\ell m-1} \right) \\ &= \int_{\Sigma_{n-1}} \frac{\partial P}{\partial x_j}(y) \, Y_{\ell m-1}(y) \, d\sigma(y), \quad 1 \leq \ell \leq d(m-1), \end{aligned}$$

and A is a positive constant given by the formula

$$A := \sum_{\ell=1}^{d(m-1)} b_\ell^2 = \int_{\Sigma_{n-1}} \left| \frac{\partial P}{\partial x_j}(x') \right|^2 d\sigma(x').$$

Therefore, we have the estimate

$$\left|\frac{\partial P}{\partial x_j}(x')\right|^2 \leq \frac{d(m-1)}{\omega_n} \int_{\Sigma_{n-1}} \left|\frac{\partial P}{\partial x_j}\right|^2 d\sigma \qquad (4.59)$$
$$\leq \frac{d(m-1)}{\omega_n} \int_{\Sigma_{n-1}} |\operatorname{grad} P|^2 \, d\sigma, \quad x' \in \Sigma_{n-1}.$$

However, since $\operatorname{grad} P$ is homogeneous of degree $m-1$, by letting $r := |x|$ and by using polar coordinates we obtain that

$$\int_{|x|\leq 1} |\operatorname{grad} P|^2 \, dx = \left(\int_0^1 r^{2m+n-3} \, dr\right) \int_{\Sigma_{n-1}} |\operatorname{grad} P|^2 \, d\sigma$$
$$= \frac{1}{2m+n-2} \int_{\Sigma_{n-1}} |\operatorname{grad} P|^2 \, d\sigma,$$

so that

$$\int_{\Sigma_{n-1}} |\operatorname{grad} P|^2 \, d\sigma = (2m+n-2) \int_{|x|\leq 1} |\operatorname{grad} P|^2 \, dx. \qquad (4.60)$$

Moreover, since $P(x)$ is harmonic in $\mathbf{R}^n$, it follows that

$$\operatorname{div}(P \operatorname{grad} P) = |\operatorname{grad} P|^2 + P\,\Delta P = |\operatorname{grad} P|^2.$$

Hence, we have, by the divergence theorem (see Theorem 5.2 in Chapter 5),

$$\int_{|x|\leq 1} |\operatorname{grad} P|^2 \, dx = \int_{|x|\leq 1} \operatorname{div}(P \operatorname{grad} P) \, dx \qquad (4.61)$$
$$= \int_{\Sigma_{n-1}} P(\operatorname{grad} P) \cdot \boldsymbol{\nu} \, d\sigma.$$

However, since $P(x)$ is homogeneous of degree m, we have, by Euler's formula for homogeneous functions,

$$(\operatorname{grad} P) \cdot \boldsymbol{\nu} = m\, P(x').$$

Hence it follows that

$$\int_{\Sigma_{n-1}} P(\operatorname{grad} P) \cdot \boldsymbol{\nu} \, d\sigma \qquad (4.62)$$
$$= m \int_{\Sigma_{n-1}} P(x')^2 \, d\sigma(x') = m \int_{\Sigma_{n-1}} Y_{\ell m}(x')^2 \, d\sigma(x') = m.$$

By combining formulas (4.60), (4.61) and (4.62), we obtain that

$$\int_{\Sigma_{n-1}} |\operatorname{grad} P|^2 \, d\sigma = m(2m+n-2).$$

Therefore, by carrying this formula into estimate (4.59) we obtain from assertion (d) with $m := m-1$ of Remark 4.4 that

$$\begin{aligned}\left|\frac{\partial P}{\partial x_j}(x')\right|^2 &\leq \frac{d(m-1)}{\omega_n}\, m(2m+n-2) \\ &\leq C(m-1)^{n-2}\, m(2m+n-2) \\ &\leq C'\, m^n \quad \text{for all } x' \in \Sigma_{n-1},\end{aligned}$$

where C and C' are positive constants depending on n. Moreover, since $\partial P/\partial x_j$ is homogeneous of degree $m-1$, we have, for some positive constant C'' depending on n,

$$\left|\frac{\partial P}{\partial x_j}(x)\right| \leq C''\, m^{n/2}\, |x|^{m-1}, \quad x \in \mathbf{R}^n.$$

This proves the desired estimate (4.58) for $|\alpha| = 1$.

Repeating this argument, we obtain the desired estimate (4.58) for the general case.

The proof of Theorem 4.26 is complete. □

By part (ii) of Theorem 4.22, we know that any function $f \in L^2(\Sigma_{n-1})$ can be expanded by the Fourier series

$$f(x') = \sum_{m=0}^{\infty} \sum_{\ell=1}^{d(m)} a_{\ell m}\, Y_{\ell m}(x'), \tag{4.63}$$

where the Fourier coefficients $a_{\ell m}$ are given by the formulas

$$a_{\ell m} = \int_{\Sigma_{n-1}} f(y)\, Y_{\ell m}(y)\, d\sigma(y). \tag{4.64}$$

and satisfy the Parseval identity

$$\sum_{m=0}^{\infty} \sum_{\ell=1}^{d(m)} |a_{\ell m}|^2 = \int_{\Sigma_{n-1}} |f(x')|^2\, d\sigma(x').$$

Remark 4.6. Any continuous function $f(x')$ on the unit sphere Σ_{n-1} has a continuous extension $\widetilde{f}(x) := f(x/|x|)$ to the space $\mathbf{R}^n \setminus \{0\}$ which is homogeneous of degree zero. Conversely, any function $g(x) \in C(\mathbf{R}^n \setminus \{0\})$ which is homogeneous of degree zero is of the form $g(x) = \widetilde{f}(x)$ where $f(x)$ is its restriction to the unit sphere Σ_{n-1}. In the following, we shall not distinguish between $f(x')$ and $\widetilde{f}(x)$, and if we write $f(x') \in C^\infty(\Sigma_{n-1})$, then we mean that $\widetilde{f}(x) \in C^\infty(\mathbf{R}^n \setminus \{0\})$.

If $f(x') \in C^\infty(\Sigma_{n-1})$, then it is well known that the Fourier series (4.63) converges absolutely and uniformly to $f(x')$ in the unit sphere Σ_{n-1}. In fact, we shall prove (Theorem 4.30) that a necessary and sufficient condition for a sequence

$$\{a_{\ell m}\}_{\substack{m=0,1,\ldots,\\ \ell=1,2,\ldots,d(m)}}$$

to be the harmonic Fourier coefficients of some function

$$f(x') \in C^\infty(\Sigma_{n-1})$$

is that the sequence $\{a_{\ell m}\}$ be rapidly decreasing. This result is a consequence of the formula (4.71) analogous to formula (4.64), which we shall prove in Theorem 4.29 below.

First, we introduce a second-order differential operator $\mathcal{L}$, which preserves the degree of homogeneity of a function, given by the formula

$$\mathcal{L}f := |x|^2 \,\Delta f.$$

We remark that $\mathcal{L}f = \Delta f$ on the unit sphere Σ_{n-1}. Then we have the following lemma:

Lemma 4.27. *For any function $Y_{\ell m}(x')$ and any integer $r \geq 1$, we have the formula*

$$\mathcal{L}^r Y_{\ell m}(x') = (-m)^r (m+n-2)^r Y_{\ell m}(x') \quad \text{for } x' \in \Sigma_{n-1}. \tag{4.65}$$

Proof. First, we prove formula (4.65) for $r = 1$.

If $P(x) := |x|^m Y_{\ell m}(x')$ is the corresponding solid harmonic, then it follows that

$$\mathcal{L}Y_{\ell m}(x') = |x|^2 \Delta Y_{\ell m} = |x|^2 \Delta \left(|x|^{-m}\, P(x)\right). \tag{4.66}$$

However, we remark the following elementary formulas:

$$\begin{aligned}
&\Delta(fg) = f\,\Delta g + 2(\operatorname{grad} f)\cdot(\operatorname{grad} g) + g\,\Delta f,\\
&\operatorname{grad}\left(|x|^k\right) = k|x|^{k-2}\,x,\\
&\Delta\left(|x|^k\right) = k(k+n-2)\,|x|^{k-2},\\
&x\cdot(\operatorname{grad} P) = m\,P(x).
\end{aligned}$$

Hence, by the harmonicity of $P(x)$ it follows that

$$\begin{aligned}
&\Delta\left(|x|^{-m}\,P(x)\right)\\
&= 2\left(\operatorname{grad}|x|^{-m}\right)\cdot(\operatorname{grad} P) + (-m)(-m+n-2)\,|x|^{-m-2}\,P(x)\\
&= 2(-m)|x|^{-m-2}\,x\cdot(\operatorname{grad} P) + (-m)(-m+n-2)\,|x|^{-m-2}\,P(x)
\end{aligned}$$

$$= 2(-m^2)|x|^{-m-2}\, P(x) + (-m)(-m+n-2)|x|^{-m-2}\, P(x)$$
$$= (-m)(m+n-2)\, |x|^{-m-2}\, P(x).$$

Therefore, by substituting this formula into formula (4.66) we obtain that

$$\begin{aligned}\mathcal{L}Y_{\ell m}(x') &= (-m)(m+n-2)\, |x|^{-m}\, \Delta\left(|x|^{-m}\, P(x)\right)\\ &= (-m)(m+n-2)Y_{\ell m}(x').\end{aligned}$$

This proves the desired formula (4.65) for $r = 1$.

The desired formula (4.65) for any integer $r \geq 2$ follows by iteration.

The proof of lemma 4.27 is complete. □

Lemma 4.28. *If f, $g \in C^{2r}(\mathbf{R}^n \setminus \{0\})$ are homogeneous of degree zero, then we have the formula*

$$\int_{\Sigma_{n-1}} f \cdot \mathcal{L}^r g \, d\sigma = \int_{\Sigma_{n-1}} g \cdot \mathcal{L}^r f \, d\sigma. \tag{4.67}$$

Proof. Using polar coordinates, we obtain that, for some constant $C \neq 0$,

$$\begin{aligned}&\int_{1 \leq |x| \leq 2} (f(x) \cdot \Delta g(x) - g(x) \cdot \Delta f(x))\, dx \qquad (4.68)\\ &= C \int_{\Sigma_{n-1}} (f(x') \cdot \Delta g(x') - g(x') \cdot \Delta f(x'))\, d\sigma(x').\end{aligned}$$

However, we have, by Green's formula (see Theorem 5.3 in Chapter 5),

$$\begin{aligned}&\int_{1 \leq |x| \leq 2} (f(x) \cdot \Delta g(x) - g(x) \cdot \Delta f(x))\, dx \qquad (4.69)\\ &= \int_{|x|=2} \left(f(x')\, \frac{\partial g}{\partial \boldsymbol{\nu}}(x') - g(x')\, \frac{\partial f}{\partial \boldsymbol{\nu}}(x') \right) d\sigma(x')\\ &\quad - \int_{|x|=1} \left(f(x')\, \frac{\partial g}{\partial \boldsymbol{\nu}}(x') - g(x')\, \frac{\partial f}{\partial \boldsymbol{\nu}}(x') \right) d\sigma(x')\\ &= 0.\end{aligned}$$

Indeed, it suffices to note that

$$\frac{\partial}{\partial \boldsymbol{\nu}} = \frac{\partial}{\partial |x|}$$

is the radical derivative and that $f(x)$ and $g(x)$ are constant along radii.

Therefore, by combining formula (4.68) and formula (4.69) we obtain that

$$\int_{\Sigma_{n-1}} (f \cdot \Delta g - g \cdot \Delta f)\, d\sigma = 0. \tag{4.70}$$

Since we have $\mathcal{L}f = \Delta f$ and $\mathcal{L}g = \Delta g$ on Σ_{n-1}, it follows from formula (4.70) that

$$\int_{\Sigma_{n-1}} f \cdot \mathcal{L}g\, d\sigma = \int_{\Sigma_{n-1}} g \cdot \mathcal{L}f\, d\sigma.$$

Moreover, by replacing g by $\mathcal{L}g$ in this formula we obtain that

$$\int_{\Sigma_{n-1}} f \cdot \mathcal{L}^2 g\, d\sigma = \int_{\Sigma_{n-1}} \mathcal{L}g \cdot \mathcal{L}f\, d\sigma = \int_{\Sigma_{n-1}} g \cdot \mathcal{L}^2 f\, d\sigma.$$

Hence, the desired formula (4.67) follows by repeating this process.

The proof of lemma 4.28 is complete. □

Theorem 4.29. *If $f \in C^\infty(\Sigma_{n-1})$ and if $\{a_{\ell m}\}$ are its Fourier coefficients with respect to the spherical harmonics $\{Y_{\ell m}\}$, then we have, for any integer $r \geq 0$,*

$$a_{\ell m} = (-m)^{-r}(m+n-2)^{-r} \int_{\Sigma_{n-1}} Y_{\ell m} \cdot \mathcal{L}^r f\, d\sigma. \tag{4.71}$$

Proof. By combining Lemma 4.27 and Lemma 4.28, we obtain that

$$\begin{aligned} a_{\ell m} &= \int_{\Sigma_{n-1}} f \cdot Y_{\ell m}\, d\sigma = (-m)^{-r}(m+n-2)^{-r} \int_{\Sigma_{n-1}} f \cdot \mathcal{L}^r Y_{\ell m}\, d\sigma \\ &= (-m)^{-r}(m+n-2)^{-r} \int_{\Sigma_{n-1}} Y_{\ell m} \cdot \mathcal{L}^r f\, d\sigma. \end{aligned}$$

This proves the desired formula (4.71). □

Theorem 4.30. *(i) If $f \in C^\infty(\Sigma_{n-1})$ and if $\{a_{\ell m}\}$ are its Fourier coefficients with respect to the spherical harmonics $\{Y_{\ell m}\}$, then we have, for any integer $r \geq 0$,*

$$\sum_{m=0}^{\infty} \sum_{\ell=1}^{d(m)} m^r \, |a_{\ell m}| < \infty. \tag{4.72}$$

(ii) Conversely, given any family $\{a_{\ell m}\}_{\substack{m=0,1,\ldots, \\ \ell=1,2,\ldots,d(m)}}$ of constants which

satisfies condition (4.72) *for any integer* $r \geq 0$*, we can construct a function* $f \in C^\infty(\Sigma_{n-1})$ *such that*

$$f(x') = \sum_{m=0}^{\infty} \sum_{\ell=1}^{d(m)} a_{\ell m} \, Y_{\ell m}(x').$$

Proof. (i) We have only to prove estimate (4.72) for all r sufficiently large. By formula (4.71), it follows that

$$|a_{\ell m}| = m^{-r}(m+n-2)^{-r} \left| \int_{\Sigma_{n-1}} Y_{\ell m}(y) \, \mathcal{L}^r f(y) \, d\sigma(y) \right|.$$

Hence, by using Schwarz's inequality (Theorem 3.14 with $p = q := 2$) and the normality of $Y_{\ell m}$ we obtain that

$$\begin{aligned}
&|a_{\ell m}| \qquad (4.73)\\
&\leq m^{-2r} \left(\int_{\Sigma_{n-1}} |Y_{\ell m}(y)|^2 \, d\sigma(y) \right)^{1/2} \left(\int_{\Sigma_{n-1}} |\mathcal{L}^r f(y)|^2 \, d\sigma(y) \right)^{1/2} \\
&= m^{-2r} \left(\int_{\Sigma_{n-1}} |\mathcal{L}^r f(y)|^2 \, d\sigma(y) \right)^{1/2} \\
&= C(r) \, m^{-2r},
\end{aligned}$$

where $C(r)$ is a positive constant defined by the formula

$$C(r) := \left(\int_{\Sigma_{n-1}} |\mathcal{L}^r f(y)|^2 \, d\sigma(y) \right)^{1/2}.$$

Since we have, by assertion (d) of Remark 4.4,

$$d(m) \leq C \, m^{n-2}$$

for some positive constant C, it follows from inequality (4.73) that we have, for all integers $r > n-1$,

$$\begin{aligned}
\sum_{m=0}^{\infty} \sum_{\ell=1}^{d(m)} m^r \, |a_{\ell m}| &\leq C(r) \sum_{m=1}^{\infty} d(m) \, m^{-r} \leq C \, C(r) \sum_{m=1}^{\infty} m^{n-2-r} \\
&< \infty.
\end{aligned}$$

This proves part (i).

(ii) Conversely, we assume that the constants $\{a_{\ell m}\}$ satisfy condition

(4.72) for all integers $r \geq 0$. Then it follows from an application of Theorem 4.26 that we have the estimate

$$|Y_{\ell m}(x')| \leq C\, m^{(n-2)/2} \quad \text{for all } x' \in \Sigma_{n-1},$$

with a positive constant C independent of m. This proves that

$$\sum_{m=0}^{\infty} \sum_{\ell=1}^{d(m)} |a_{\ell m}\, Y_{\ell m}(x')| \leq C \sum_{m=0}^{\infty} \sum_{\ell=1}^{d(m)} |a_{\ell m}|\, m^{(n-2)/2} \quad \text{for all } x' \in \Sigma_{n-1}.$$

Therefore, we obtain from condition (4.72) that the series

$$\sum_{\ell,m} a_{\ell m}\, Y_{\ell m}(x')$$

converges absolutely and uniformly to some continuous function $f(x')$ in the unit sphere Σ_{n-1}.

By repeating this argument and by using estimate (4.58) of Theorem 4.26, we obtain that the series $\sum_{\ell,m} a_{\ell m}\, (\partial/\partial x)^{\alpha} Y_{\ell m}(x')$ is absolutely and uniformly convergent to the continuous function $(\partial/\partial x)^{\alpha} f(x')$ in the unit sphere Σ_{n-1}. Summing up, we have proved that $f \in C^{\infty}(\Sigma_{n-1})$.

The proof of Theorem 4.30 is complete. □

By combining Theorem 4.22, Theorem 4.26 and Theorem 4.30, we have proved the following fundamental results for spherical harmonics:

Theorem 4.31. *(i) The space $\mathcal{H}_m$ of n-dimensional spherical harmonics of degree m has dimension*

$$d(m) = g(m) - g(m-2),$$

where

$$g(\ell) = \binom{\ell + n - 1}{n-1} = \frac{(\ell + n - 1)!}{(n-1)!\,\ell!}, \quad \ell \geq 0,$$
$$g(-1) = g(-2) = 0.$$

Moreover, we have the estimate

$$d(m) \leq c(n)\, m^{n-2}, \tag{4.74}$$

with a positive constant $c(n)$ depending on n.

(ii) The Hilbert space $L^2(\Sigma_{n-1})$ can be decomposed into the infinite direct sum in the sense of Hilbert space theory:

$$L^2(\Sigma_{n-1}) = \sum_{m=0}^{\infty} \mathcal{H}_m.$$

More precisely, any function $f \in L^2(\Sigma_{n-1})$ *has the development*

$$f(x') = \sum_{m=0}^{\infty} Y_m(x'), \quad Y_m \in \mathcal{H}_m,$$

where the convergence is in the $L^2(\Sigma_{n-1})$ *norm, and we have the formula*

$$\int_{\Sigma_{n-1}} |f(x')|^2 \, d\sigma(x') = \sum_{m=0}^{\infty} \int_{\Sigma_{n-1}} |Y_m(x')|^2 \, d\sigma(x').$$

(iii) If $Y_m(x')$ *is any* n*-dimensional spherical harmonic of degree* m*, then its partial derivatives* $(\partial^\alpha Y_m / \partial x^\alpha)(x')$ *satisfy the estimates*

$$\sup_{x' \in \Sigma_{n-1}} \left| \frac{\partial^\alpha Y_m}{\partial x^\alpha}(x') \right| \leq c \, m^{|\alpha| + (n-2)/2}, \tag{4.75}$$

with a positive constant $c = c(n, |\alpha|)$ *depending on the multi-indices* α *and* n.

(iv) The spherical harmonics $\{Y_{km}\}_{\substack{m=0,1,\ldots, \\ k=1,2,\ldots,d(m)}}$ *form a complete orthonormal system in the Hilbert space* $L^2(\Sigma_{n-1})$*. If* $f \in C^\infty(\Sigma_{n-1})$ *has the Fourier series expansion with respect to* $\{Y_{km}\}$

$$f(x') = \sum_{m=0}^{\infty} \sum_{k=1}^{d(m)} a_{km} \, Y_{km}(x'),$$

$$a_{km} = \int_{\Sigma_{n-1}} f(y) \, Y_{km}(y) \, d\sigma(y),$$

then the Fourier coefficients a_{km} *satisfy, for any integer* $r > 1$*, the estimate*

$$|a_{km}| \leq c(r) \, m^{-2r}, \tag{4.76}$$

with a positive constant $c(r)$ *depending on* r.

4.8 Notes and Comments

For more thorough treatments of real analytic tools, the reader might be referred to Duoandikoetxea [23], Garcia-Cuerva–Rubio de Francia [31], Garnett [32], Malý–Ziemer [45], Neri [51], Stein [68], [70] and Torchinsky [91].

Part II

Elements of Function Spaces

5

Harmonic Functions and Poisson Integrals

The purpose of this chapter is to study harmonic functions in the half-space $\mathbf{R}_+^{n+1}$ in terms of Poisson integrals of functions in various L^p spaces (Theorems 5.8, 5.9 and 5.10). In particular, we establish fundamental relationships between means of derivatives of Poisson integrals $u(x, y)$ taken with respect to the normal variable y and those taken with respect to the tangential variables x_i (Theorems 5.14 and 5.19). More precisely, Theorem 5.14 (resp. Theorem 5.19) establishes fundamental relationships between means of the first (resp. second) derivatives of $u(x, y)$ taken with respect to y and those taken with respect to x_i.

5.1 Lipschitz Domains and Green's Formulas

An open set Ω in $\mathbf{R}^n$ is called a *Lipschitz hypograph* if its boundary $\partial\Omega$ can be represented as the graph of a Lipschitz continuous function. Namely, there exists a Lipschitz continuous function $\zeta \colon \mathbf{R}^{n-1} \to \mathbf{R}$ such that (see Figure 5.1)

$$\Omega = \left\{ x = (x', x_n) \in \mathbf{R}^n : x_n < \zeta(x'), \quad x' \in \mathbf{R}^{n-1} \right\}. \tag{5.1}$$

Sometimes, a different smoothness condition will be needed, so we broaden the above terminology as follows (see [98, Section 2]): For any non-negative integer k and any $0 < \theta \leq 1$, we say that the domain Ω defined by formula (5.1) is a $C^{k,\theta}$ *hypograph* if the function ζ is of class $C^{k,\theta}$, that is, if ζ is of class C^k and its k-th order partial derivatives are Hölder continuous with exponent θ.

The next definition requires that, roughly speaking, the boundary of Ω can be represented locally as the graph of a Lipschitz continuous function, by using different systems of Cartesian coordinates for different parts of the boundary:

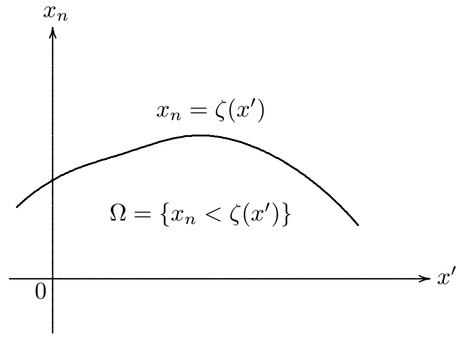

Fig. 5.1. The Lipschitz hypograph $\Omega = \{x_n < \zeta(x')\}$

Definition 5.1. Let Ω be a bounded domain in Euclidean space $\mathbf{R}^n$ with boundary $\partial\Omega$. We say that Ω is a *Lipschitz domain* if there exist finite families $\{U_j\}_{j=1}^J$ and $\{\Omega_j\}_{j=1}^J$ having the following three properties (i), (ii) and (iii) (see Figure 5.2):

(i) The family $\{U_j\}_{j=1}^J$ is a finite open covering of $\partial\Omega$.
(ii) Each Ω_j can be transformed to a Lipschitz hypograph by a rigid motion, that is, by a rotation plus a translation.
(iii) The set Ω satisfies the conditions

$$U_j \cap \Omega = U_j \cap \Omega_j, \quad 1 \leq j \leq J.$$

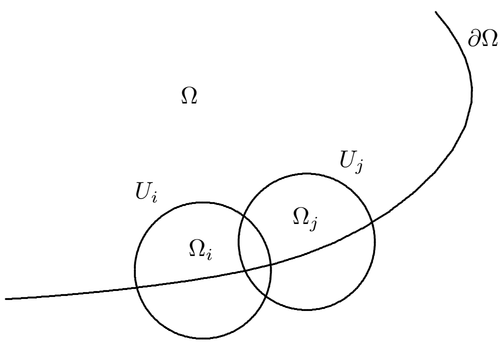

Fig. 5.2. The families $\{U_j\}$ and $\{\Omega_j\}$ of the Lipschitz domain Ω

In the obvious way, we define a $C^{k,\theta}$ *domain* by substituting "$C^{k,\theta}$" for "Lipschitz" throughout Definition 5.1. It should be emphasized that a Lipschitz domain is the same thing as a $C^{0,1}$ domain.

If Ω is a Lipschitz hypograph defined by formula (5.1), then we remark that its boundary

$$\partial\Omega = \left\{x = (x', \zeta(x')) : x' \in \mathbf{R}^{n-1}\right\}$$

is an $(n-1)$-dimensional, $C^{0,1}$ submanifold of $\mathbf{R}^n$. Hence we find that $\partial\Omega$ has a surface measure $d\sigma$ and an outward unit normal $\boldsymbol{\nu}$ which exists $d\sigma$-almost everywhere in $\mathbf{R}^{n-1}$, if we apply the following Rademacher's theorem (see [45, Chapter 1, Corollary 1.73], [69, Theorem]):

Theorem 5.1 (Rademacher)**.** *We have the assertion*

$$C^{0,1}(\mathbf{R}^n) = W^{1,\infty}(\mathbf{R}^n).$$

In other words, any Lipschitz continuous function on $\mathbf{R}^n$ admits L^∞ first partial derivatives almost everywhere in $\mathbf{R}^n$.

Indeed, it follows from an application of Rademacher's theorem that the function $\zeta(x')$ is Fréchet differentiable almost everywhere in $\mathbf{R}^{n-1}$ with

$$\|\nabla\zeta\|_{L^\infty(\mathbf{R}^{n-1})} \leq C, \tag{5.2}$$

where C is any Lipschitz constant for the function $\zeta(x')$. Then we have the following formulas for $d\sigma$ and $\boldsymbol{\nu}$ (see Figure 5.3):

$$\begin{aligned} d\sigma &= \sqrt{1 + |\nabla\zeta(x')|^2}\, dx', \\ \boldsymbol{\nu} &= \frac{(-\nabla\zeta(x'), 1)}{\sqrt{1 + |\nabla\zeta(x')|^2}}. \end{aligned}$$

Here it should be noticed that we have, by inequality (5.2),

$$1 \leq \sqrt{1 + |\nabla\zeta(x')|^2} \leq \sqrt{1 + C^2},$$

so that the surface measure $d\sigma$ is equivalent locally to the Lebesgue measure dx'.

We consider the case where Ω is a bounded Lipschitz domain. By using the notation of Definition 5.1, we choose a partition of unity $\{\phi_j\}_{j=1}^J$ subordinate to the open covering $\{U_j\}_{j=1}^J$ of $\partial\Omega$ (see Figure 5.4), that is,

$$\begin{aligned} &\phi_j \in C_0^\infty(U_j), \\ &0 \leq \phi_j(x) \leq 1 \quad \text{in } U_j, \\ &\sum_{j=1}^J \phi_j(x) = 1 \quad \text{on } \partial\Omega. \end{aligned}$$

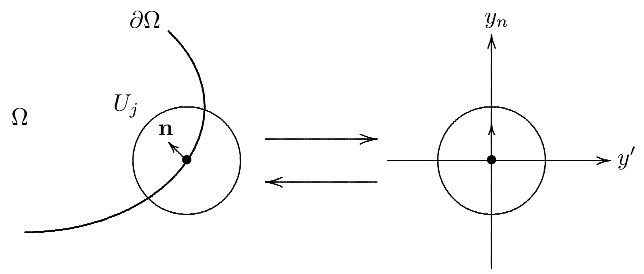

Fig. 5.3. The unit outward normal $\mathbf{n} = -\boldsymbol{\nu}$ to the boundary $\partial\Omega = \{x_n = \zeta(x')\}$

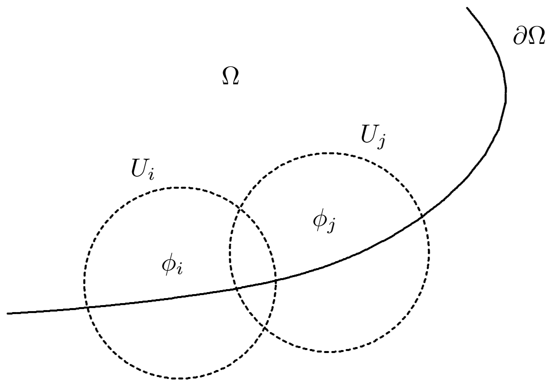

Fig. 5.4. The partition of unity $\{\phi_j\}$ subordinate to the open covering $\{U_j\}$ of $\partial\Omega$

Then, upon using local coordinate systems flattening out the boundary $\partial\Omega$, together with a partition of unity, we can prove the following divergence theorem for Lipschitz domains:

Theorem 5.2 (the divergence theorem)**.** *Let* Ω *be a bounded, Lipschitz domain of* $\mathbf{R}^n$ *with boundary* $\partial\Omega$*. If* $F = (f_1, f_2, \ldots, f_n)$*, is a* C^1 *vector field, then we have the formula*

$$\int_\Omega \sum_{i=1}^n \frac{\partial f_i}{\partial x_i}\, dx = \int_{\partial\Omega} \sum_{i=1}^n f_i \nu_i \, d\sigma, \quad f_i \in C^1(\overline{\Omega}), \tag{5.3}$$

or equivalently,

$$\int_\Omega \operatorname{div} F\, dx = \int_{\partial\Omega} (F, \boldsymbol{\nu})\, d\sigma,$$

where $\boldsymbol{\nu} = (\nu_1, \nu_2, \ldots, \nu_n)$ *is the unit outward normal to* $\partial\Omega$ *and* $d\sigma$ *is the surface measure of* $\partial\Omega$.

5.2 Harmonic Functions

We start with the following elementary result:

Let $I = (a, b)$ be an open interval of $\mathbf{R}$. If $u \in C^2(I)$ and $d^2u/dx^2 = 0$ in I, then $u(x)$ is a linear function. In particular, we have, for all sufficiently small $r > 0$,

$$u(x) = \frac{1}{2}[u(x+r) + u(x-r)] \quad \text{(the mean value property)},$$

$$u(x) = \frac{1}{2r}\int_{-r}^{r} u(x+z)\,dz = \frac{1}{2}\int_{-1}^{1} u(x+ry)\,dy.$$

In this section, we extend these results to the n-dimensional case by replacing the operator d^2/dx^2 by the usual Laplacian

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \ldots + \frac{\partial^2}{\partial x_n^2}.$$

To do this, we need the following Green formulas for Lipschitz domains:

Theorem 5.3 (Green's formula)**.** *Let* Ω *be a bounded, Lipschitz domain of* $\mathbf{R}^n$ *with boundary* $\partial\Omega$. *For* u, $v \in C^2(\overline{\Omega})$, *we have the formulas*

$$\int_\Omega (v\Delta u + \nabla v \cdot \nabla u)\,dx = \int_{\partial\Omega} v\frac{\partial u}{\partial \boldsymbol{\nu}}\,d\sigma. \tag{5.4a}$$

$$\int_\Omega (v\Delta u - u\Delta v)\,dx = \int_{\partial\Omega} \left(v\frac{\partial u}{\partial \boldsymbol{\nu}} - \frac{\partial v}{\partial \boldsymbol{\nu}}u\right)\,d\sigma. \tag{5.4b}$$

Proof. (i) Formula (5.4a) follows by applying the divergence theorem (Theorem 5.2) with

$$f = v\,\nabla u = \left(v\frac{\partial u}{\partial x_1}, v\frac{\partial u}{\partial x_2}, \ldots, v\frac{\partial u}{\partial x_n}\right).$$

(ii) Formula (5.4b) follows by interchanging u and v in formula (5.4a) and then by subtracting this formula from formula (5.4a).

The proof of Theorem 5.3 is complete. □

A function $u(x)$ defined in Ω is said to be *harmonic* if it is twice continuously differentiable in Ω and satisfies the Laplace equation

$$\Delta u = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \ldots + \frac{\partial^2 u}{\partial x_n^2} = 0 \quad \text{in } \Omega.$$

Corollary 5.4. *If a function $u(x) \in C^2(\Omega)$ is harmonic in Ω, then we have, for any Lipschitz subdomain Ω' of Ω,*

$$\int_{\partial\Omega'} \frac{\partial u}{\partial \boldsymbol{\nu}} \, d\sigma = 0.$$

Proof. It suffices to take $v := 1$ in formula (5.4a). □

Theorem 5.5 (the mean value theorem)**.** *Assume that $u(x) \in C^2(\Omega)$ is* harmonic *in Ω. If $x \in \Omega$ and $r > 0$ is small enough so that*

$$\overline{B(x,r)} \subset \Omega,$$

then we have the formulas

$$u(x) = \frac{1}{r^{n-1}\omega_n} \int_{S(x,r)} u(y) \, d\sigma(y) \tag{5.5a}$$

$$= \frac{1}{\omega_n} \int_{S(0,1)} u(x + rz) \, d\sigma(z). \tag{5.5b}$$

Here

$$\omega_n := |\Sigma_{n-1}| = \frac{2\pi^{n/2}}{\Gamma(n/2)}$$

is the surface area of the unit sphere $\Sigma_{n-1} = S(0,1)$ in $\mathbf{R}^n$.

Proof. For any $0 < \varepsilon < r$, we consider the annular domain (see Figure 5.5)

$$\Gamma_{\varepsilon r} := \{y \in \Omega : \varepsilon < |y - x| < r\} = B(x,r) \setminus \overline{B(x,\varepsilon)}.$$

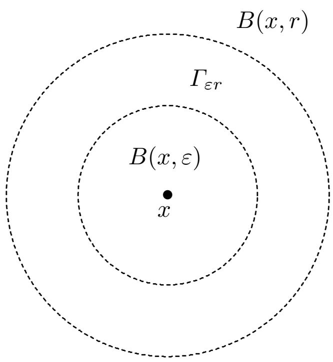

Fig. 5.5. The open balls $B(x,r)$, $B(x,\varepsilon)$ and the annular domain $\Gamma_{\varepsilon r}$

(i) The proof of the first formula (5.5a): We let

$$v(y) = \begin{cases} \frac{1}{|x-y|^{n-2}} & \text{if } n \geq 3, \\ -\log|x-y| & \text{if } n = 2. \end{cases}$$

Then, since we have the formula

$$\frac{\partial}{\partial \boldsymbol{\nu}} = \begin{cases} \frac{1}{r}\sum_{j=1}^{n}(y_j - x_j)\frac{\partial}{\partial y_j} & \text{on } S(x,r) = \{x \in \Omega : |y-x| = r\}, \\ \frac{1}{\varepsilon}\sum_{j=1}^{n}(x_j - y_j)\frac{\partial}{\partial y_j} & \text{on } S(x,\varepsilon) = \{x \in \Omega : |y-x| = \varepsilon\}, \end{cases}$$

it follows that

$$\frac{\partial v}{\partial \boldsymbol{\nu}} = \begin{cases} (2-n)\frac{1}{r^{n-1}} & \text{on } S(x,r), \\ -(2-n)\frac{1}{\varepsilon^{n-1}} & \text{on } S(x,\varepsilon). \end{cases}$$

Furthermore, it is easy to see that

$$\Delta v(y) = 0 \quad \text{for all } y \in \Omega \setminus \{x\}.$$

Thus, by applying Green's formula (5.4b) with $\Omega := \Gamma_{\varepsilon r}$ we obtain that

$$\begin{aligned} 0 &= \int_{S(x,r)} \left(v\frac{\partial u}{\partial \boldsymbol{\nu}} - \frac{\partial v}{\partial \boldsymbol{\nu}}u \right) d\sigma + \int_{S(x,\varepsilon)} \left(v\frac{\partial u}{\partial \boldsymbol{\nu}} - \frac{\partial v}{\partial \boldsymbol{\nu}}u \right) d\sigma \qquad (5.6) \\ &= \int_{S(x,r)} \frac{1}{r^{n-2}}\frac{\partial u}{\partial \boldsymbol{\nu}}\, d\sigma - \int_{S(x,r)} (2-n)\frac{1}{r^{n-1}} u\, d\sigma \\ &\quad + \int_{S(x,\varepsilon)} \frac{1}{\varepsilon^{n-2}}\frac{\partial u}{\partial \boldsymbol{\nu}}\, d\sigma + \int_{S(x,\varepsilon)} (2-n)\frac{1}{\varepsilon^{n-1}} u\, d\sigma. \end{aligned}$$

However, by Corollary 5.4 it follows that

$$\int_{S(x,r)} \frac{\partial u}{\partial \boldsymbol{\nu}}\, d\sigma = 0,$$
$$\int_{S(x,\varepsilon)} \frac{\partial u}{\partial \boldsymbol{\nu}}\, d\sigma = 0.$$

Hence we have, by formula (5.6),

$$\frac{1}{r^{n-1}} \int_{S(x,r)} u(y)\, d\sigma(y) = \frac{1}{\varepsilon^{n-1}} \int_{S(x,\varepsilon)} u(z)\, d\sigma(z). \qquad (5.7)$$

Now we need the following:

Claim 5.1. We have, as $\varepsilon \downarrow 0$,

$$\frac{1}{\varepsilon^{n-1}} \int_{S(x,\varepsilon)} u(z)\, d\sigma(z) \longrightarrow \omega_n\, u(x).$$

Proof. Indeed, by the continuity of $u(x)$ it follows that

$$\begin{aligned}
&\frac{1}{\omega_n \varepsilon^{n-1}} \left| \int_{S(x,\varepsilon)} u(z)\, d\sigma(z) - u(x) \right| \\
&= \frac{1}{\omega_n \varepsilon^{n-1}} \left| \int_{S(x,\varepsilon)} (u(z) - u(x))\, d\sigma(z) \right| \\
&\leq \frac{1}{\omega_n \varepsilon^{n-1}} \int_{S(x,\varepsilon)} |u(z) - u(x)|\, d\sigma(z) \\
&\leq \sup_{z \in B(x,\varepsilon)} |u(z) - u(x)| \longrightarrow 0 \quad \text{as } \varepsilon \downarrow 0.
\end{aligned}$$

This proves Claim 5.1. □

Therefore, by letting $\varepsilon \downarrow 0$ in formula (5.7) we obtain from Claim 5.1 that

$$u(x) = \frac{1}{r^{n-1}\omega_n} \int_{S(x,r)} u(y)\, d\sigma(y).$$

(ii) The proof of the second formula (5.5b): We make the change of variables

$$y = x + rz, \quad z \in S(0,1).$$

Then, since we have the formula

$$d\sigma(y) = r^{n-1}\, d\sigma(z),$$

it follows from formula (5.5a) that

$$\begin{aligned}
u(x) &= \frac{1}{r^{n-1}\omega_n} \int_{S(x,r)} u(y)\, d\sigma(y) = \frac{1}{r^{n-1}\omega_n} \int_{S(0,1)} u(x+rz) r^{n-1}\, d\sigma(z) \\
&= \frac{1}{\omega_n} \int_{S(0,1)} u(x+rz)\, d\sigma(z).
\end{aligned}$$

The proof of Theorem 5.5 is complete. □

Corollary 5.6. *Assume that $u(x)$ is harmonic in Ω. Then we have the formulas*

$$u(x) = \frac{n}{r^n \omega_n} \int_{B(x,r)} u(y)\, dy \tag{5.8a}$$

$$= \frac{n}{\omega_n} \int_{B(0,1)} u(x+rz)\, dz. \tag{5.8b}$$

Remark 5.1. The volume V_n of the unit ball $B(0,1)$ is given by the formula

$$V_n = \int_0^1 \int_{S(0,1)} r^{n-1}\, d\sigma(y)\, dr = \omega_n \int_0^1 r^{n-1}\, dr = \frac{\omega_n}{n} = \frac{\pi^{n/2}}{\Gamma(n/2+1)}.$$

Hence we can rewrite formula (5.8b) as follows:

$$u(x) = \frac{1}{V_n} \int_{B(0,1)} u(x+rz)\, dz. \tag{5.9}$$

Proof. (i) The proof of the first formula (5.8a): By formula (5.5b), it follows that, for $0 < \rho < 1$,

$$u(x) = \frac{1}{\omega_n} \int_{S(0,1)} u(x+\rho r y)\, d\sigma(y).$$

Therefore, integrating both sides from 0 to 1 with respect to $\rho^{n-1} d\rho$ we obtain that

$$\begin{aligned}
\frac{1}{n} u(x) &= \frac{1}{\omega_n} \int_0^1 \int_{S(0,1)} u(x+\rho r y) \rho^{n-1}\, d\rho\, d\sigma(y) \\
&= \frac{1}{\omega_n} \int_0^r \int_{S(0,1)} u(x+ty) \left(\frac{t}{r}\right)^{n-1} \frac{dt}{r}\, d\sigma(y) \\
&= \frac{1}{\omega_n} \frac{1}{r^n} \int_0^r \int_{S(0,1)} u(x+ty)\, t^{n-1}\, d\sigma(y)\, dt \\
&= \frac{1}{r^n \omega_n} \int_{B(x,r)} u(y)\, dy.
\end{aligned}$$

This proves formula (5.8a).

(ii) The proof of the second equality (5.8b): We make the change of variables

$$y = x + rz, \quad z \in B(0,1).$$

Then, since we have the formula

$$dy = r^n\, dz,$$

it follows from formula (5.8a) that

$$\begin{aligned}
u(x) &= \frac{n}{r^n \omega_n} \int_{B(x,r)} u(y)\, dy = \frac{n}{r^n \omega_n} \int_{B(0,1)} u(x+rz) r^n\, dz \\
&= \frac{n}{\omega_n} \int_{B(0,1)} u(x+rz)\, dz.
\end{aligned}$$

The proof of Corollary 5.6 is complete. □

5.3 Poisson Kernels

To motivate the definition of Poisson kernels, we begin with some heuristic remarks on the Dirichlet boundary value problem

$$\begin{cases} \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \cdots + \frac{\partial^2}{\partial x_n^2} + \frac{\partial^2}{\partial y^2}\right) u(x,y) = 0 & \text{in } \mathbf{R}_+^{n+1}, \\ u(x,0) = f(x) & \text{on } \mathbf{R}^n \end{cases} \tag{5.10}$$

in the half-space $\mathbf{R}_+^{n+1} = \{(x,y) : x \in \mathbf{R}^n, y > 0\}$. If we consider the partial Fourier transform $\widetilde{u}(\xi, y)$ of a solution $u(x,y)$ of problem (5.10), then we have the following initial value problem

$$\begin{cases} \left(\frac{d^2}{dy^2} - |\xi|^2\right) \widetilde{u}(\xi, y) = 0, \quad y > 0, \\ \widetilde{u}(\xi, 0) = \widehat{f}(\xi). \end{cases}$$

Hence it follows that the partial Fourier transform $\widetilde{u}(\xi, y)$ is given by the formula

$$\widetilde{u}(\xi, y) = \widehat{f}(\xi)\, e^{-|\xi| y}, \quad y > 0.$$

Therefore, by the partial Fourier inversion formula we obtain that the solution $u(x,y)$ of problem (5.10) may be "formally" expressed as follows:

$$\begin{aligned} u(x,y) &= \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\xi}\, e^{-|\xi| y}\, \widehat{f}(\xi)\, d\xi \\ &= \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{-|\xi| y} \left(\int_{\mathbf{R}^n} e^{i(x-z)\xi}\, f(z)\, dz \right) d\xi \\ &= \int_{\mathbf{R}^n} \left(\frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{i(x-z)\xi}\, e^{-|\xi| y}\, d\xi \right) f(z)\, dz \\ &:= \int_{\mathbf{R}^n} P(x-z, y)\, f(y)\, dy, \end{aligned}$$

where

$$P(x,y) := \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\xi} e^{-|\xi| y}\, d\xi, \quad y > 0. \tag{5.11}$$

The function $P(x,y)$ is called the *Poisson kernel* for the half-space $\mathbf{R}_+^{n+1}$. However, we can calculate explicitly the Poisson kernel $P(x,y)$. Indeed, we have the following:

Claim 5.2. The Poisson kernel $P(x,y)$ is given by the formula

$$P(x,y) = \frac{1}{c_n} \frac{y}{(|x|^2 + y^2)^{(n+1)/2}}, \quad x \in \mathbf{R}^n,\ y > 0, \tag{5.12a}$$

where c_n is a positive constant given by the formula

$$c_n = \frac{\pi^{(n+1)/2}}{\Gamma((n+1)/2)} = \frac{\omega_{n+1}}{2}. \tag{5.12b}$$

Here

$$\omega_{n+1} := |\Sigma_n| = (n+1)V_{n+1} = \frac{2\pi^{(n+1)/2}}{\Gamma((n+1)/2)}$$

is the surface area of the unit sphere Σ_n in $\mathbf{R}^{n+1}$.

Proof. First, we recall the following two well-known formulas (5.13) and (5.14):

$$e^{-\gamma} = \frac{1}{\sqrt{\pi}} \int_0^\infty \frac{e^{-u}}{\sqrt{u}}\, e^{-\frac{\gamma^2}{4u}}\, du, \quad \gamma > 0. \tag{5.13}$$

$$\frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\xi} e^{-\alpha|\xi|^2}\, dx = \frac{1}{(4\alpha\pi)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4\alpha}} \quad \text{for } \alpha > 0. \tag{5.14}$$

Therefore, by using Fubini's theorem (Theorem 3.10) we obtain from formula (5.13) with $\gamma := |\xi| y$ and formula (5.14) with $\alpha := y^2/(4u)$ that

$$\begin{aligned}
&\frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\xi} e^{-|\xi| y}\, d\xi \\
&= \int_{\mathbf{R}^n} e^{ix\xi} \left(\frac{1}{\sqrt{\pi}} \int_0^\infty \frac{e^{-u}}{\sqrt{u}} e^{-|\xi|^2 y^2/(4u)}\, du \right) d\xi \\
&= \frac{1}{\sqrt{\pi}} \int_0^\infty \frac{e^{-u}}{\sqrt{u}} \left(\frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} e^{ix\xi} e^{-|\xi|^2 y^2/(4u)}\, d\xi \right) du \\
&= \frac{1}{\sqrt{\pi}} \int_0^\infty \frac{e^{-u}}{\sqrt{u}} \left(\frac{u}{\pi y^2} \right)^{n/2} e^{-|x|^2 u/y^2}\, du \\
&= \frac{1}{\pi^{(n+1)/2}} \frac{1}{y^n} \int_0^\infty u^{(n-1)/2} e^{-(1+|x|^2/y^2)u}\, du \\
&= \frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}} \frac{y}{(|x|^2+y^2)^{(n+1)/2}} \\
&= P(x,y).
\end{aligned}$$

The proof of Claim 5.2 is complete. □

Now we prove fundamental properties of the Poisson kernel $P(x, y)$:

Lemma 5.7. *The Poisson kernel $P(x, y)$ enjoys the following properties:*

$$\int_{\mathbf{R}^n} P(x,y)\, dx = 1 \quad \textit{for all } x \in \mathbf{R}^n. \tag{5.15a}$$

$$P_y(x,y) = \frac{1}{c_n} \frac{|x|^2 - ny^2}{(|x|^2+y^2)^{(n+3)/2}} \quad \text{for all } x \in \mathbf{R}^n. \tag{5.15b}$$

$$|P_y(x,y)| \leq \frac{n+1}{y} P(x,y) \quad \text{for all } x \in \mathbf{R}^n.$$

$$P_{x_i}(x,y) = \frac{1}{c_n} \frac{-(n+1)x_i y}{(|x|^2+y^2)^{(n+3)/2}} \quad \text{for all } x \in \mathbf{R}^n, \tag{5.15c}$$

$$|P_{x_i}(x,y)| \leq \frac{n+1}{2y} P(x,y) \quad \text{for all } x \in \mathbf{R}^n \text{ and } 1 \leq i \leq n.$$

Here $P_y(x,y)$ and $P_{x_i}(x,y)$ indicate the partial derivative of $P(x,y)$ in the direction of the element subscripted, that is,

$$P_y(x,y) = \frac{\partial P}{\partial y}, \quad P_{x_i}(x,y) = \frac{\partial P}{\partial x_i}.$$

Proof. It is easy to verify assertions (5.15b) and (5.15c). Moreover, assertions (5.15a) follow from a direct calculation. Indeed, we have, by formula (5.12a),

$$P(x,y) = \frac{1}{c_n} \frac{y}{(|x|^2+y^2)^{(n+1)/2}} = \frac{1}{y^n} P\left(\frac{x}{y}, 1\right) \quad \text{for all } y > 0,$$

and also

$$\begin{aligned}
&\int_{\mathbf{R}^n} P(z,1)\, dz \\
&= \frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}} \int_{\mathbf{R}^n} (1+|z|^2)^{-(n+1)/2}\, dz \\
&= \frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}} \int_0^\infty \int_\Sigma (1+r^2)^{-(n+1)/2} r^{n-1}\, dr\, d\sigma \\
&= \frac{\Gamma((n+1)/2)}{\pi^{(n+1)/2}} \left(\frac{2\pi^{n/2}}{\Gamma(n/2)}\right) \int_0^\infty \int_\Sigma (1+r^2)^{-(n+1)/2} r^{n-1}\, dr\, d\sigma \\
&= \frac{2}{\sqrt{\pi}} \frac{\Gamma((n+1)/2)}{\Gamma(n/2)} \int_1^\infty s^{-(n+1)/2} (s-1)^{(n-1)/2} \frac{1}{2\sqrt{s-1}}\, ds \\
&= \frac{1}{\sqrt{\pi}} \frac{\Gamma((n+1)/2)}{\Gamma(n/2)} \int_1^\infty \left(1 - \frac{1}{s}\right)^{(n-2)/2} s^{-3/2}\, ds \\
&= \frac{1}{\sqrt{\pi}} \frac{\Gamma((n+1)/2)}{\Gamma(n/2)} \int_0^1 (1-\sigma)^{(n-2)/2} \sigma^{-1/2}\, d\sigma \\
&= \frac{1}{\sqrt{\pi}} \frac{\Gamma((n+1)/2)}{\Gamma(n/2)} B(1/2, n/2) = \frac{1}{\sqrt{\pi}} \frac{\Gamma((n+1)/2)}{\Gamma(n/2)} \frac{\Gamma(1/2)\Gamma(n/2)}{\Gamma((n+1)/2))} \\
&= 1.
\end{aligned}$$

Therefore, we obtain that

$$\int_{\mathbf{R}^n} P(x,y)\,dx = \int_{\mathbf{R}^n} \frac{1}{y^n} P\left(\frac{x}{y},1\right)\,dx = \int_{\mathbf{R}^n} P(z,1)\,dz = 1.$$

The proof of Lemma 5.7 is complete. □

5.4 Poisson Integrals

If $f(x) \in L^p(\mathbf{R}^n)$ with $1 \le p \le \infty$, then the *Poisson integral* of $f(x)$ is defined as the convolution of $f(x)$ with the Poisson kernel $P(x,y)$:

$$\begin{aligned} u(x,y) &= P(x,y) * f(x) \qquad (5.16)\\ &= \frac{1}{c_n}\int_{\mathbf{R}^n} \frac{y}{(|x-z|^2+y^2)^{(n+1)/2}} f(z)\,dz \quad \text{for all } y>0. \end{aligned}$$

Then we have the following assertions for the Poisson integral $u(x,y)$:

Theorem 5.8. *(i) Let $f \in L^p(\mathbf{R}^n)$ with $1 \le p \le \infty$. Then the function*

$$u(x,y) = P(x,y) * f(x)$$

is well defined in the half-space $\mathbf{R}^{n+1}_+$, and satisfies the inequality

$$\|u(\cdot,y)\|_p \le \|f\|_p \quad \text{for all } y>0.$$

Furthermore, the function $u(x,y)$ is harmonic in $\mathbf{R}^{n+1}_+$.

(ii) If $1 \le p < \infty$, we have, as $y \downarrow 0$,

$$u(\cdot,y) \longrightarrow f \quad \text{in } L^p(\mathbf{R}^n). \tag{5.17}$$

If, in addition, $f(x)$ is bounded and continuous on $\mathbf{R}^n$, then the function $u(x,y)$ is continuous on $\overline{\mathbf{R}^{n+1}_+} = \{(x,y) : x \in \mathbf{R}^n, y \ge 0\}$, and satisfies the Dirichlet condition

$$u(x,0) = f(x) \quad \text{for every } x \in \mathbf{R}^n. \tag{5.18}$$

Proof. The proof is divided into three steps.

Step 1: Since we have, for all $y > 0$,

$$P(\cdot,y) \in \bigcap_{p=1}^{\infty} L^p(\mathbf{R}^n),$$

it follows that the function $u(x,y) = P(x,y) * f(x)$ is well defined on $\mathbf{R}^{n+1}_+$.

Furthermore, we can differentiate formula (5.16) under the integral sign to obtain the following formulas (5.19) and (5.20):

$$\frac{\partial^2 u}{\partial y^2}(x,y) \tag{5.19}$$
$$= \frac{1}{c_n} \int_{\mathbf{R}^n} \left[\frac{n(n+1)y^3 - 3(n+1)y|x-z|^2}{(|x-z|^2 + y^2)^{(n+1)/2+1}} \right] f(z)\,dz,$$

$$\frac{\partial^2 u}{\partial x_i^2}(x,y) \tag{5.20}$$
$$= \frac{1}{c_n} \int_{\mathbf{R}^n} \left[\frac{(n+1)(n+3)y(x_i - z_i)^2 - (n+1)(t^2 + |x-z|^2)}{(|x-z|^2 + y^2)^{(n+1)/2+2}} \right] f(z)\,dz.$$

It should be noticed that the integrals converge absolutely on compact subsets of $\mathbf{R}^{n+1}_+$. Indeed, it suffices to note that the terms in the bracket $[\cdot]$ in formulas (5.19) and (5.20) are bounded by a function in $L^{p'}(\mathbf{R}^n)$ uniformly in (x,y) of a compact subset of $\mathbf{R}^{n+1}_+$, where $p' = p/(p-1)$ is the exponent conjugate to p.

Summing up, we find that

$$u(x,y) \in C^2(\mathbf{R}^{n+1}_+),$$

and that

$$\left(\sum_{j=1}^{n} \frac{\partial^2}{\partial x_j^2} + \frac{\partial^2}{\partial y^2} \right) u(x,y) = 0 \quad \text{in } \mathbf{R}^{n+1}_+,$$

that is, the function $u(x,y)$ is harmonic in the half-space $\mathbf{R}^{n+1}_+$.

Step 2: We recall that

$$\int_{\mathbf{R}^n} P(x,1)\,dx = 1,$$

and further that

$$\frac{1}{y^n} P\left(\frac{x}{y}, 1\right) = P(x,y) \quad \text{for all } y > 0.$$

Therefore, the desired assertion (5.17) follows from an application of Theorem 3.25 with

$$\varphi(x) := P(x,1), \quad \varepsilon := y.$$

Step 3: To prove assertion (5.18), we assume that $f(x)$ is bounded and continuous on $\mathbf{R}^n$. Then we have, for all $y > 0$,

$$P(x,y) * f(x) - f(x) = \int_{\mathbf{R}^n} P(x-z,y)[f(z) - f(x)]\,dz \tag{5.21}$$

$$= \int_{\mathbf{R}^n} P(z,y)[f(x-z)-f(x)]\,dz$$
$$= \int_{\mathbf{R}^n} P(\xi,1)[f(x-y\xi)-f(x)]\,d\xi.$$

However, we remark that

$$|P(\xi,1)\,[f(x-y\xi)-f(x)]\,| \leq 2\max|f|\cdot P(\xi,1),$$

and that

$$\int_{\mathbf{R}^n} P(\xi,1)\,d\xi = 1.$$

Therefore, by applying the Lebesgue dominated convergence theorem (Theorem 3.9) we obtain from formula (5.21) that

$$\lim_{y\downarrow 0} P(x,y) * f(x) = f(x) \quad \text{for every } x \in \mathbf{R}^n.$$

The proof of Theorem 5.8 is complete. □

Assertions (5.17) and (5.18) may be made precise. To do this, we recall some important results in the theory of differentiation of integrals of functions defined on $\mathbf{R}^n$ (Theorem 4.8 and Corollary 4.6).

First, Corollary 4.6 may be restated as follows:

Claim 5.3. If $f(t)$ is a locally integrable function on $\mathbf{R}^n$, then we have, for almost every $x \in \mathbf{R}^n$,

$$\lim_{r\downarrow 0} \frac{1}{r^n} \int_{|t|<r} (f(x-t)-f(x))\,dt = 0. \tag{5.22}$$

In particular, this is true for functions $f \in L^p(\mathbf{R}^n)$ with $1 \leq p \leq \infty$.

Indeed, it suffices to note that

$$\left| \frac{1}{r^n} \int_{|t|<r} (f(x-t)-f(x))\,dt \right|$$
$$= \left| \frac{1}{r^n} \int_{B(x,r)} (f(y)-f(x))\,dy \right| \leq \frac{1}{r^n} \int_{B(x,r)} |f(y)-f(x)|\,dy$$
$$= 2^n\,\frac{1}{(2r)^n} \int_{B(x,r)} |f(y)-f(x)|\,dy$$
$$\leq 2^n\,\frac{1}{|Q(x,r)|} \int_{Q(x,r)} |f(y)-f(x)|\,dy.$$

Therefore, the desired assertion (5.22) follows from assertion (4.15).

We let

$D(f) =$ the set of all points $x \in \mathbf{R}^n$ for which condition (5.22) holds.

The set $D(f)$ is called the set of points where the integral of f is *differentiable*. There is a smaller but closely related set which is also naturally associated with each locally integrable function.

We let

$$L(f) = \text{the set of all Lebesgue points } x \in \mathbf{R}^n \text{ for } f, \text{ i.e.,}$$
$$\lim_{r \downarrow 0} \frac{1}{r^n} \int_{|t|<r} |f(x-t) - f(x)| \, dt = 0. \tag{5.23}$$

The set $L(f)$ is called the *Lebesgue set* of f (see Definition 4.1).

Secondly, Theorem 4.8 may be restated as follows:

Claim 5.4. If $f(x)$ is a locally integrable function on $\mathbf{R}^n$, then the complement of the Lebesgue set $L(f)$ has measure *zero*.

Now we can generalize assertions (5.17) and (5.18) as follows:

Theorem 5.9. *Let $f \in L^p(\mathbf{R}^n)$ with $1 \leq p \leq \infty$. Then we have the formula*

$$\lim_{\varepsilon \downarrow 0} P(x, \varepsilon) * f(x) = \lim_{\varepsilon \downarrow 0} \int_{\mathbf{R}^n} P(x-t, \varepsilon) f(t) \, dt = f(x), \tag{5.24}$$

whenever x belongs to the Lebesgue set $L(f)$ of f. In particular, the convergence (5.24) *holds true for almost all $x \in \mathbf{R}^n$.*

Proof. If x is a point of the Lebesgue set $L(f)$, then it follows from condition (5.23) that, for each $\delta > 0$, there exists a constant $\eta = \eta(x, \delta) > 0$ such that

$$\frac{1}{r^n} \int_{|t|<r} |f(x-t) - f(x)| \, dt < \delta \quad \text{for all } 0 < r \leq \eta. \tag{5.25}$$

We let

$$\varphi(t) := P(t, 1) = \frac{1}{c_n} \frac{1}{(|t|^2+1)^{(n+1)/2}}, \quad t \in \mathbf{R}^n,$$
$$\varphi_\varepsilon(t) := P(t, \varepsilon) = \frac{1}{c_n} \frac{\varepsilon}{(|t|^2+\varepsilon^2)^{(n+1)/2}}, \quad t \in \mathbf{R}^n, \quad \varepsilon > 0.$$

Then we have the inequality

$$|P(x, \varepsilon) * f(x) - f(x)| = \left| \int_{\mathbf{R}^n} P(t, \varepsilon) f(x-t) \, dt - f(x) \right| \tag{5.26}$$

$$
\begin{aligned}
&= \left| \int_{\mathbf{R}^n} (f(x-t) - f(x))\, \varphi_\varepsilon(t)\, dt \right| \\
&\le \int_{|t|\ge\eta} |f(x-t) - f(x)| \varphi_\varepsilon(t)\, dt \\
&\quad + \int_{|t|<\eta} |f(x-t) - f(x)| \varphi_\varepsilon(t)\, dt \\
&:= I_1^{(\varepsilon)}(x,\eta) + I_2^{(\varepsilon)}(x,\eta).
\end{aligned}
$$

We estimate the two integrals $I_1^{(\varepsilon)}(x,\eta)$ and $I_2^{(\varepsilon)}(x,\eta)$ on the right of inequality (5.26).

(i) The estimate of $I_1^{(\varepsilon)}(x,\eta)$: We let

$$\chi_\eta(x) = \text{ the characteristic function of the set } \{t \in \mathbf{R}^n : |t| \ge \eta\}.$$

Then we have, by Hölder's inequality (Theorem 3.14),

$$
\begin{aligned}
I_1^{(\varepsilon)}(x,\eta) &= \int_{|t|\ge\eta} |f(x-t) - f(x)|\, \varphi_\varepsilon(t)\, dt \\
&= \int_{\mathbf{R}^n} |f(x-t) - f(x)|\, \chi_\eta(t)\varphi_\varepsilon(t)\, dt \\
&\le \int_{\mathbf{R}^n} |f(x-t)|\, \chi_\eta(t)\varphi_\varepsilon(t)\, dt + \int_{\mathbf{R}^n} |f(x)|\, \chi_\eta(t)\varphi_\varepsilon(t)\, dt \\
&\le \|f\|_p \, \|\chi_\eta\varphi_\varepsilon\|_{p'} + |f(x)| \, \|\chi_\eta\varphi_\varepsilon\|_1 .
\end{aligned}
$$

However, it follows that, as $\varepsilon \downarrow 0$,

$$\|\chi_\eta\varphi_\varepsilon\|_1 = \int_{|t|\ge\eta} \frac{1}{\varepsilon^n} \varphi\left(\frac{t}{\varepsilon}\right) dt = \int_{|z|\ge\frac{\eta}{\varepsilon}} \varphi(z)\, dz \longrightarrow 0,$$

and further that

$$
\begin{aligned}
\|\chi_\eta\varphi_\varepsilon\|_{p'} &= \left(\int_{|t|\ge\eta} \varphi_\varepsilon(t)^{p'}\, dt \right)^{1/p'} \\
&= \left(\int_{\mathbf{R}^n} (\chi_\eta(t)\varphi_\varepsilon(t))\, (\chi_\eta(t)\varphi_\varepsilon(t))^{p/p'}\, dt \right)^{1/p'} \\
&\le \|\chi_\eta\varphi_\varepsilon\|_\infty^{1/p} \, \|\chi_\eta\varphi_\varepsilon\|_1^{1/p'} \longrightarrow 0,
\end{aligned}
$$

since we have the formula

$$\|\chi_\eta\varphi_\varepsilon\|_\infty = \sup_{|t|\ge\eta} |\varphi_\varepsilon(t)| = \frac{1}{c_n} \frac{\varepsilon}{(\eta^2+\varepsilon^2)^{(n+1)/2}}.$$

Hence we find that, as $\varepsilon \downarrow 0$,

$$I_1^{(\varepsilon)}(x,\eta) \le \|f\|_p \, \|\chi_\eta\varphi_\varepsilon\|_{p'} + |f(x)| \, \|\chi_\eta\varphi_\varepsilon\|_1 \longrightarrow 0. \tag{5.27}$$

(ii) The estimate of $I_2^{(\varepsilon)}(x,\eta)$: We let

$$g(r) := \int_{\Sigma_{n-1}} |f(x - r\sigma) - f(x)| \, d\sigma,$$

where Σ_{n-1} is the unit sphere in $\mathbf{R}^n$. Then we find that condition (5.25) is equivalent to the following condition: For each $\delta > 0$, there exists a constant $\eta = \eta(x,\delta) > 0$ such that

$$\begin{aligned} G(r) := \int_0^r s^{n-1} g(s)\, ds &= \int_{|t|<r} |f(x-t) - f(x)| \, dt \\ &< \delta r^n \quad \text{for all } 0 < r \leq \eta. \end{aligned} \tag{5.25'}$$

Moreover, we introduce a function $\psi(r)$ (associated with the function $\varphi(x)$) by the formula

$$\psi(r) := \frac{1}{c_n} \frac{1}{(r^2+1)^{(n+1)/2}}, \quad r > 0.$$

By integration by parts, it follows that

$$\begin{aligned} I_2^{(\varepsilon)}(x,\eta) &\leq \int_{|t|<\eta} |f(x-t) - f(x)| \, \varphi_\varepsilon(t)\, dt \\ &= \int_0^\eta \int_{\Sigma_{n-1}} |f(x - r\sigma) - f(x)| \, \varepsilon^{-n} \, \psi\left(\frac{r}{\varepsilon}\right) r^{n-1}\, dr\, d\sigma \\ &= \int_0^\eta r^{n-1} g(r)\, \varepsilon^{-n}\, \psi\left(\frac{r}{\varepsilon}\right) dr \\ &= \left[G(r)\, \varepsilon^{-n}\, \psi\left(\frac{r}{\varepsilon}\right)\right]_0^\eta - \int_0^\eta G(r) \left(\varepsilon^{-n}\, \psi'\left(\frac{r}{\varepsilon}\right)\right) \varepsilon^{-1}\, dr \\ &= G(\eta)\, \varepsilon^{-n}\, \psi\left(\frac{\eta}{\varepsilon}\right) - \int_0^{\eta/\varepsilon} G(\varepsilon s)\, \varepsilon^{-n}\, \psi'(s)\, ds. \end{aligned}$$

Hence we have, by condition (5.25′),

$$I_2^{(\varepsilon)}(x,\eta) \leq \delta\, \eta^n\, \varepsilon^{-n}\, \psi\left(\frac{\eta}{\varepsilon}\right) - \delta \int_0^{\eta/\varepsilon} (\varepsilon s)^n \varepsilon^{-n}\, \psi'(s)\, ds \leq A\delta, \tag{5.28}$$

where A is a positive constant defined by the formula

$$A := \frac{1}{c_n}\left(\sup_{r>0}\left\{\frac{r^n}{(r^2+1)^{(n+1)/2}}\right\} + n \int_0^\infty \frac{s^{n-1}}{(s^2+1)^{(n+1)/2}}\, ds\right).$$

Therefore, by combining inequalities (5.26), (5.27) and (5.28) we obtain that

$$\limsup_{\varepsilon\downarrow 0} |P(x,\varepsilon) * f(x) - f(x)| \leq \limsup_{\varepsilon \downarrow 0} \left\{I_1^{(\varepsilon)}(x,\eta) + I_2^{(\varepsilon)}(x,\eta)\right\} \leq A\delta.$$

This proves the desired assertion (5.24), since δ is arbitrary.

The proof of Theorem 5.9 is complete. □

5.5 Manipulations of Harmonic Functions

Theorem 5.8 has a converse:

Theorem 5.10. *Assume that a function $u(x,y)$ is harmonic in the half-space $\mathbf{R}_+^{n+1}$ and that there exist constants $c > 0$ and $1 \leq p < \infty$ such that*

$$\|u(\cdot, y)\|_p \leq c \quad \text{for all } y > 0. \tag{5.29}$$

Then we have the following two assertions (a) and (b):

(a) If $1 < p < \infty$, there exists a function $f(x) \in L^p(\mathbf{R}^n)$ such that

$$u(x,y) = P(x,y) * f(x).$$

(b) If $p = 1$, there exists a finite Borel measure μ on $\mathbf{R}^n$ such that

$$u(x,y) = \int_{\mathbf{R}^n} P(x-t, y)\, d\mu(t).$$

Furthermore, if $\{u(\cdot, y)\}$ is a Cauchy sequence in $L^1(\mathbf{R}^n)$ as $y \downarrow 0$, then there exists a function $f(x) \in L^1(\mathbf{R}^n)$ such that $u(x,y) = P(x,y) * f(x)$.

The proof of Theorem 5.10 is divided into two steps.

Step 1: In order to prove Theorem 5.10, we need some lemmas:

Lemma 5.11. *Assume that a function $u(x,y)$ is harmonic in the half-space $\mathbf{R}_+^{n+1}$ and satisfies condition (5.29). Then we have the inequality*

$$\|u(\cdot, y)\|_\infty \leq \left(\frac{2^{n+1}}{V_{n+1}}\right)^{1/p} c\, y^{-n/p} \quad \text{for all } y > 0,$$

where V_{n+1} is the volume of the unit ball in $\mathbf{R}^{n+1}$

$$V_{n+1} := \frac{2\pi^{(n+1)/2}}{(n+1)\Gamma((n+1)/2)}.$$

In particular, the function $u(x,y)$ is bounded in each proper sub-half space of $\mathbf{R}_+^{n+1}$.

Proof. By applying the mean value theorem for harmonic functions (Theorem 5.5) with

$$\Omega := \mathbf{R}^{n+1}, \quad r := \frac{y}{2},$$

we obtain that

$$\begin{aligned}
u(x,y) &= \frac{n+1}{(y/2)^{n+1}} \int_0^{y/2} u(x,y) r^n \, dr \\
&= \frac{(n+1)2^{n+1}}{y^{n+1}} \int_0^{y/2} \left(\frac{1}{\omega_{n+1}} \int_{\Sigma_n} u((x,y) + r\sigma) \, d\sigma \right) r^n \, dr \\
&= \frac{2^{n+1}}{V_{n+1}} \frac{1}{y^{n+1}} \int_{|t| \le y/2} u((x,y) + t) \, dt.
\end{aligned}$$

Here Σ_n is the unit sphere in $\mathbf{R}^{n+1}$ and

$$\omega_{n+1} := |\Sigma_n| = (n+1) V_{n+1} = \frac{2\pi^{(n+1)/2}}{\Gamma((n+1)/2)}$$

is its surface area. Therefore, we have, by Hölder's inequality (Theorem 3.14) and inequality (5.29),

$$\begin{aligned}
|u(x,y)| &\le \frac{2^{n+1}}{V_{n+1}} \frac{1}{y^{n+1}} \int_{|(x,y)-(\xi,\eta)| \le y/2} |u(\xi,\eta)| \, d\xi \, d\eta \\
&\le \frac{2^{n+1}}{V_{n+1}} \frac{1}{y^{n+1}} \left(\int_{|(x,y)-(\xi,\eta)| \le y/2} |u(\xi,\eta)|^p \, d\xi \, d\eta \right)^{1/p} \\
&\quad \times \left(V_{n+1} \left(\frac{y}{2} \right)^{n+1} \right)^{1-1/p} \\
&\le \left(\frac{2^{n+1}}{V_{n+1}} \right)^{1/p} y^{-(n+1)/p} \left(\int_{y/2}^{3y/2} \left(\int_{\mathbf{R}^n} |u(\xi,\eta)|^p \, d\xi \right) d\eta \right)^{1/p} \\
&\le \left(\frac{2^{n+1}}{V_{n+1}} \right)^{1/p} y^{-(n+1)/p} \left(\int_{y/2}^{3y/2} c^p \, d\eta \right)^{1/p} \\
&\le \left(\frac{2^{n+1}}{V_{n+1}} \right)^{1/p} c y^{-n/p}.
\end{aligned}$$

The proof of Lemma 5.11 is complete. □

The next lemma is a special case of the *Schwarz reflection principle* for hyperplanes (see [6, Chapter 4, Theorem 4.12]):

Lemma 5.12. *If $u(x,y)$ is harmonic in the half-space $\mathbf{R}^{n+1}_+$, everywhere continuous and bounded on the closure $\overline{\mathbf{R}^{n+1}_+} = \mathbf{R}^{n+1}_+ \cup \mathbf{R}^n$ and is equal to zero on the boundary $\mathbf{R}^n$, then it follows that $u(x,y) \equiv 0$ on $\mathbf{R}^{n+1}_+$.*

Proof. The proof is divided into two steps.

(1) We let

$$u^*(x,y) := \begin{cases} u(x,y) & \text{if } y \geq 0, \\ -u(x,-y) & \text{if } y < 0. \end{cases}$$

Then it follows that

$$\begin{cases} \Delta u^* = 0 & \text{in } \mathbf{R}^{n+1}_+, \\ \Delta u^* = 0 & \text{in } \mathbf{R}^{n+1}_-, \end{cases}$$

and further that

$$u^* = 0 \quad \text{on } \mathbf{R}^n.$$

Hence we have the assertion

$$u^* \in C\left(\mathbf{R}^{n+1}\right) \cap C^2\left(\mathbf{R}^{n+1}_+\right) \cap C^2\left(\mathbf{R}^{n+1}_-\right).$$

Now we show that u^* is harmonic near $\mathbf{R}^n$ in $\mathbf{R}^{n+1}$. To do so, let $(x_0, 0)$ be an arbitary point of $\mathbf{R}^n$ and let

$$B_r := B\left((x_0,0), r\right) \subset \mathbf{R}^{n+1}$$

be an arbitrary small open ball of radius r about $(x_0, 0)$ (see Figure 5.6).

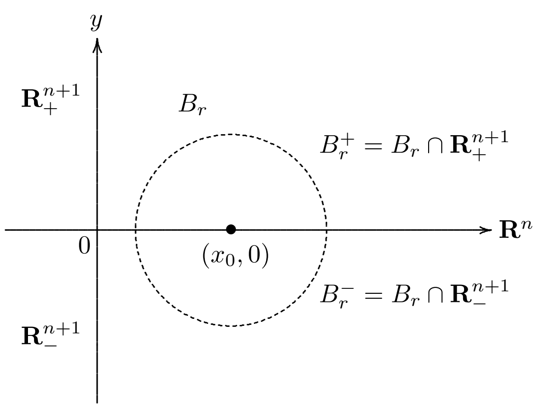

Fig. 5.6. The open ball B_r of radius r about $(x_0, 0)$

Then we have the formula

$$\begin{aligned} &\int_{\partial B_r} u^*(x,y)\, d\sigma(x,y) \\ &= \int_{\partial B_r^+} u(x,y)\, d\sigma(x,y) - \int_{\partial B_r^-} u(x,-y)\, d\sigma(x,y) \end{aligned}$$

$$= \int_{\partial B_r^+} u(x,y)\, d\sigma(x,y) - \int_{\partial B_r^+} u(x,y)\, d\sigma(x,y)$$
$$= 0.$$

Hence there exists a constant $\varepsilon > 0$ such that

$$\int_{\partial B_r} u^*(x,y)\, d\sigma(x,y) = 0 \quad \text{for all } 0 < r < \varepsilon.$$

Therefore, by using Green's identity (5.4a) with $\Omega := B_r$ and $v := 1$ we obtain that

$$\begin{aligned}
0 &= \frac{d}{dr}\left(\frac{1}{r^n\,\omega_{n+1}} \int_{\partial B_r} u^*(x,y)\, d\sigma(x,y)\right) \\
&= \frac{d}{dr}\left(\frac{1}{\omega_{n+1}} \int_{S(0,1)} u^*\left((x_0,0) + rz\right)\, d\sigma(z)\right) \\
&= \frac{1}{\omega_{n+1}} \int_{S(0,1)} z \cdot \nabla u^*\left((x_0,0) + rz\right)\, d\sigma(z) \\
&= \frac{1}{\omega_{n+1}} \int_{S(0,r)} \frac{w}{r} \cdot \nabla u^*\left((x_0,0) + w\right)\, r^{-n} d\sigma(w) \\
&= \frac{1}{r^n\,\omega_{n+1}} \int_{S(0,r)} \frac{w}{r} \cdot \nabla u^*\left((x_0,0) + w\right)\, d\sigma(w) \\
&= \frac{1}{r^n\,\omega_{n+1}} \int_{S(0,r)} \frac{\partial u^*}{\partial \boldsymbol{\nu}}\left((x_0,0) + w\right) d\sigma(w) \\
&= \frac{1}{r^n\,\omega_{n+1}} \int_{B_r} \Delta u^*(x,y)\, dx\, dy \quad \text{for all } 0 < r < \varepsilon,
\end{aligned}$$

where

$$S(0,r) = \left\{ w \in \mathbf{R}^{n+1} : |w| = r \right\}.$$

Since the integral of Δu^* over any ball near the point $(x_0, 0)$ vanishes, it follows that

$$\Delta u^*(x,y) = 0 \quad \text{in a neighborhood of each point } (x_0,0) \text{ of } \mathbf{R}^n.$$

Summing up, we have proved that $u^*(x,y)$ is harmonic in the whole space $\mathbf{R}^{n+1}$.

(2) By Lemma 5.11, it follows that the function $u^*(x,y)$ is harmonic and bounded on the whole space $\mathbf{R}^{n+1}$. Hence, by applying Liouville's theorem for harmonic functions ([6, Chapter 2, Theorem 2.1]) we obtain that

$$u^*(x,y) \equiv 0 \quad \text{on } \mathbf{R}^{n+1},$$

so that

$$u(x,y) \equiv 0 \quad \text{on } \mathbf{R}_+^{n+1}.$$

The proof of Lemma 5.12 is complete. □

Lemma 5.13. *Assume that a function $u(x,y)$ is harmonic in the half-space $\mathbf{R}_+^{n+1}$ and bounded in a proper sub-half space $\{(x,y) \in \mathbf{R}_+^{n+1} : y > y_0\}$ of $\mathbf{R}_+^{n+1}$, for each $y_0 > 0$. Then we have the formula*

$$u(x, y_1 + y_2) = u(x, y_1) * P(x, y_2) \quad \textit{for all } y_1,\ y_2 > 0.$$

Proof. For each $y_1 > 0$, we let

$$w_1(x,y) := u(x, y_1 + y) \quad \text{for all } y \geq 0,$$

and

$$w_2(x,y) := u(x,y_1) * P(x,y) = \int_{\mathbf{R}^n} P(x-t,y)\, u(t,y_1)\, dt \quad \text{for all } y > 0.$$

Then it suffices to show that

$$w_1(x,y) \equiv w_2(x,y) \quad \text{for all } (x,y) \in \mathbf{R}_+^{n+1}.$$

We find that the functions $w_1(x,y)$ and $w_2(x,y)$ are both harmonic in $\mathbf{R}_+^{n+1}$, everywhere continuous and bounded on $\overline{\mathbf{R}_+^{n+1}}$, and further take the same boundary values $u(x,y_1)$ on $\mathbf{R}^n$. Hence, by applying Lemma 5.12 to the function

$$h(x,y) = w_1(x,y) - w_2(x,y),$$

we obtain that

$$h(x,y) \equiv 0 \quad \text{on } \mathbf{R}_+^{n+1},$$

that is,

$$w_1(x,y) \equiv w_2(x,y) \quad \text{for all } (x,y) \in \mathbf{R}_+^{n+1}.$$

The proof of Lemma 5.13 is complete. □

Step 2: Now, the proof of Theorem 5.10 may be carried out as follows.

Step 2-a: The case $1 < p < \infty$. Assume that

$$\|u(\cdot,y)\|_p \leq c \quad \text{for all } y > 0.$$

Then, by the *weak compactness* of the space $L^p(\mathbf{R}^n)$ (see Theorem 2.30) we can find a sequence $\{y_k\}$, $y_k \downarrow 0$, and a function $f \in L^p(\mathbf{R}^n)$ such that

$$u(\cdot, y_k) \longrightarrow f \quad \text{weakly in } L^p(\mathbf{R}^n) \text{ as } k \to \infty,$$

that is,

$$\int_{\mathbf{R}^n} u(t, y_k) g(t)\, dt \longrightarrow \int_{\mathbf{R}^n} f(t) g(t)\, dt \quad \text{for all } g \in L^{p'}(\mathbf{R}^n) \text{ as } k \to \infty.$$

However, we remark that

$$P(\cdot, y) = \frac{1}{c_n} \frac{y}{(|\cdot|^2 + y^2)^{(n+1)/2}} \in L^{p'}(\mathbf{R}^n) \quad \text{for each } y > 0.$$

Hence we have the assertion

$$\begin{aligned} &\int_{\mathbf{R}^n} P(x-t, y) u(t, y_k)\, dt \\ &\longrightarrow \int_{\mathbf{R}^n} P(x-t, y) f(t)\, dt \quad \text{for all } y > 0. \end{aligned} \tag{5.30}$$

Thus it remains to show the following:

Claim 5.5. $u(x, y) = \int_{\mathbf{R}^n} P(x-t, y)\, f(t)\, dt = P(x, y) * f(x)$ for all $y > 0$.

Proof. By Lemma 5.13, it follows that

$$u(x, y + y_k) = \int_{\mathbf{R}^n} P(x-t, y)\, u(t, y_k)\, dt. \tag{5.31}$$

Hence, by letting $k \to \infty$ in formula (5.31) we obtain from assertion (5.30) that

$$\begin{aligned} u(x, y) &= \lim_{k\to\infty} u(x, y + y_k) = \lim_{k\to\infty} \int_{\mathbf{R}^n} P(x-t, y)\, u(t, y_k)\, dt \\ &= \int_{\mathbf{R}^n} P(x-t, y) f(t)\, dt. \end{aligned}$$

The proof of Claim 5.5 is complete. □

Step 2-b: The case $p = 1$. Assume that

$$\|u(\cdot, y)\|_1 \le c \quad \text{for all } y > 0.$$

Then, by the *compactness* of measures (Theorem 2.35) we can find a sequence $\{y_k\}$, $y_k \downarrow 0$, and a finite Borel measure μ on $\mathbf{R}^n$ such that

$$\int_{\mathbf{R}^n} u(t, y_k)\, g(t)\, dt \longrightarrow \int_{\mathbf{R}^n} g(t)\, d\mu(t) \quad \text{for all } g \in C_0(\mathbf{R}^n),$$

where

$$\begin{aligned} &C_0(\mathbf{R}^n) \\ &= \text{the space of continuous functions on } \mathbf{R}^n \text{ vanishing at infinity.} \end{aligned}$$

However, we find that

$$P(\cdot, y) = \frac{1}{c_n} \frac{y}{(|\cdot|^2 + y^2)^{(n+1)/2}} \in C_0(\mathbf{R}^n) \quad \text{for all } y > 0.$$

Hence we have the assertion

$$\begin{aligned} &\int_{\mathbf{R}^n} P(x-t, y)\, u(t, y_k)\, dt \\ &\longrightarrow \int_{\mathbf{R}^n} P(x-t, y)\, d\mu(t) \quad \text{for all } y > 0. \end{aligned} \tag{5.32}$$

By arguing just as in the proof of Claim 5.5, we obtain that

$$u(x, y) = \int_{\mathbf{R}^n} P(x-t, y)\, d\mu(t) \quad \text{for all } y > 0.$$

Furthermore, if $\{u(\cdot, y)\}$ is a Cauchy sequence in $L^1(\mathbf{R}^n)$ as $y \downarrow 0$, then there exists a function $f \in L^1(\mathbf{R}^n)$ such that $u(\cdot, y) \longrightarrow f$ in $L^1(\mathbf{R}^n)$ as $y \downarrow 0$. Hence it follows that

$$\int_{\mathbf{R}^n} u(t, y_k)\, g(t)\, dt \longrightarrow \int_{\mathbf{R}^n} f(t)\, g(t)\, dt \quad \text{for all } g \in L^\infty(\mathbf{R}^n).$$

However, we remark that

$$P(\cdot, y) = \frac{1}{c_n} \frac{y}{(|\cdot|^2 + y^2)^{(n+1)/2}} \in L^\infty(\mathbf{R}^n) \quad \text{for all } y > 0.$$

Thus we have the assertion

$$\int_{\mathbf{R}^n} P(x-t, y)\, u(t, y_k)\, dt \longrightarrow \int_{\mathbf{R}^n} P(x-t, y) f(t)\, dt \quad \text{for all } y > 0,$$

and so, by assertion (5.32),

$$u(x, y) = \int_{\mathbf{R}^n} P(x-t, y)\, f(t)\, dt \quad \text{for all } y > 0.$$

The proof of Theorem 5.10 is now complete. □

The next theorem establishes some fundamental relationships between means of the first derivatives of Poisson integrals

$$u(x, y) = \int_{\mathbf{R}^n} P(x-t, y)\, f(t)\, dt$$

taken with respect to the variable y and those taken with respect to the variables x_i:

Theorem 5.14. *Assume that a function $u(x,y)$ is harmonic in the half-space $\mathbf{R}^{n+1}_+$ and bounded in a proper sub-half space $\{(x,y) \in \mathbf{R}^{n+1}_+ : y > y_0\}$ of $\mathbf{R}^{n+1}_+$, for each $y_0 > 0$. Then we have, for $\alpha > 0$, $1 \leq p,\ q \leq \infty$,*

$$\sup_{1 \leq i \leq n} \|y^\alpha u_{x_i}(\cdot,\cdot)\|_{pq} \leq M_\alpha \|y^\alpha u_y(\cdot,\cdot)|\|_{pq} . \tag{a}$$

$$\|y^\alpha u_y(\cdot,\cdot)\|_{pq} \leq M_\alpha \sup_{1 \leq i \leq n} \|y^\alpha u_{x_i}(\cdot,\cdot)\|_{pq} . \tag{b}$$

Here $u_y(x,y)$ and $u_{x_i}(x,y)$ indicate the partial derivative of $u(x,y)$ in the direction of the element subscripted, that is,

$$u_y(x,y) = \frac{\partial u}{\partial y}, \quad u_{x_i}(x,y) = \frac{\partial u}{\partial x_i},$$

and the mixed norm $\|\cdot\|_{pq}$ is defined by the formula

$$\|y^\alpha u(\cdot,\cdot)\|_{pq} := \begin{cases} \left(\int_0^\infty (y^\alpha \|u(\cdot,y)\|_p)^q \frac{dy}{y} \right)^{1/q} & \text{if } 1 \leq q < \infty, \\ \sup_{y>0} \left\{ y^\alpha \|u(\cdot,y)\|_p \right\} & \text{if } q = \infty. \end{cases}$$

The proof of Theorem 5.14 is divided into two steps.

Step 1: In order to prove Theorem 5.14, we need several lemmas.

Lemma 5.15. *Let $1 \leq p \leq \infty$. Then we have the inequality*

$$\|P(\cdot,y)\|_p \leq \left(\frac{1}{c_n} \right)^{1/p'} y^{-n/p'} \quad \text{for all } y > 0, \tag{5.33}$$

where p' is the exponent conjugate to p:

$$p' = \frac{p}{p-1}.$$

Proof. We remark that, by formulas (5.12a) and (5.15a),

$$\|P(\cdot,y)\|_1 = 1,$$
$$\|P(\cdot,y)\|_\infty = \frac{1}{c_n}\, y^{-n} \quad \text{for all } y > 0.$$

Hence the desired inequality (5.33) follows from the logarithmic convexity of the L^p norm as a function of $1/p$, since we have the formula

$$\frac{1}{p} = 1 - \frac{1}{p'} = \left(1 - \frac{1}{p'} \right) \frac{1}{1} + \frac{1}{p'} \frac{1}{\infty}.$$

The proof of Lemma 5.15 is complete. □

Lemma 5.16. *If $f(x) \in L^p(\mathbf{R}^n)$ with $1 \leq p \leq \infty$ and if $u(x,y) = P(x,y) * f(x)$ is its Poisson integral, then we have the inequality*

$$\|u(\cdot,y)\|_\infty \leq \left(\frac{1}{c_n}\right)^{1/p} \|f\|_p \, y^{-n/p} \quad \textit{for all } y > 0.$$

Proof. By Young's inequality (Theorem 3.23), it follows that

$$\begin{aligned} \|u(\cdot,y)\|_\infty &\leq \|P(\cdot,y)\|_{p'} \|f\|_p \\ &\leq \left(\frac{1}{c_n}\right)^{1/p} y^{-n/p} \|f\|_p \quad \text{for all } y > 0. \end{aligned}$$

The proof of Lemma 5.16 is complete. □

Lemma 5.17. *Assume that a function $u(x,y)$ is harmonic in the half-space $\mathbf{R}_+^{n+1}$ and that, for each $y_0 > 0$, there exists a constant $A = A(y_0) > 0$ such that*

$$\|u(\cdot,y)\|_p \leq A \quad \textit{for all } y \geq y_0, \tag{5.34}$$

with $1 \leq p \leq \infty$. Then we have the following inequalities:

$$\|u_y(\cdot,y)\|_p \leq 2(n+1)\, y^{-1} \|u(\cdot,y/2)\|_p \quad \textit{for all } y > 0. \tag{a}$$

$$\|u_{x_i}(\cdot,y)\|_p \leq (n+1)\, y^{-1} \|u(\cdot,y/2)\|_p \quad \textit{for all } y > 0. \tag{b}$$

Proof. By condition (5.34), we can apply Theorem 5.10 to obtain that there exists a function $h(x) \in L^p(\mathbf{R}^n)$ such that

$$u(x, y + y_0) = P(x,y) * h(x) \quad \text{for all } y > 0.$$

Hence it follows from an application of Lemma 5.15 that

$$\|u(\cdot, y+y_0)\|_\infty \leq \|P(\cdot,y)\|_{p'} \|h\|_p \leq c_n^{-1/p} \|h\|_p \, y^{-n/p} \quad \text{for all } y > 0.$$

This proves that the function $u(x, y + y_0)$ is bounded in each proper sub-half space of $\mathbf{R}_+^{n+1}$. Hence we have, by Lemma 5.13,

$$u(x, y_1 + y_2 + y_0) = P(x, y_2) * u(x, y_1 + y_0) \quad \text{for all } y_1,\ y_2 > 0,$$

or equivalently,

$$u(x, y+z) = P(x,y) * u(x,z) \quad \text{for all } y,\ z > 0.$$

If we differentiate the both sides with respect to y (resp. x_i), we find that

$$\begin{aligned} u_y(x, y+z) &= P_y(x,y) * u(x,z) \quad \text{for all } y,\ z > 0, \\ u_{x_i}(x, y+z) &= P_{x_i}(x,y) * u(x,z) \quad \text{for all } y,\ z > 0, \end{aligned}$$

since the integrals defining the convolutions converge absolutely.

Therefore, part (a) follows from part (b) of Lemma 5.7. Indeed, we have the inequality

$$\begin{aligned}\|u_y(\cdot,y)\|_p &= \|P_y(\cdot,y/2) * u(\cdot,y/2)\|_p \leq \|P_y(\cdot,y/2)\|_1 \|u(\cdot,y/2)\|_p \\ &\leq 2(n+1)y^{-1} \|u(\cdot,y/2)\|_p \quad \text{for all } y>0.\end{aligned}$$

Similarly, part (b) follows from part (c) of Lemma 5.7.

The proof of Lemma 5.17 is complete. □

Lemma 5.18. *Assume that a function $u(x,y)$ is harmonic in the half-space $\mathbf{R}_+^{n+1}$ and bounded in a proper sub-half space $\{(x,y) \in \mathbf{R}_+^{n+1} : y > y_0\}$ of $\mathbf{R}_+^{n+1}$, for each $y_0 > 0$. Then we have, for $\alpha > 0$, $1 \leq p,\ q \leq \infty$,*

$$\left\|y^{\alpha+1} u_y(\cdot,\cdot)\right\|_{pq} \leq 2^{\alpha+1}(n+1) \|y^\alpha u(\cdot,\cdot)\|_{pq}. \tag{a}$$

$$\left\|y^{\alpha+1} u_{x_i}(\cdot,\cdot)\right\|_{pq} \leq 2^{\alpha}(n+1) \|y^\alpha u(\cdot,\cdot)\|_{pq}, \quad 1 \leq i \leq n. \tag{b}$$

If, in addition, $u(x,y) \to 0$ as $y \to \infty$, then we have the inequality

$$\|y^\alpha u(\cdot,\cdot)\|_{pq} \leq \frac{1}{\alpha} \left\|y^{\alpha+1} u_y(\cdot,\cdot)\right\|_{pq}. \tag{c}$$

Proof. We recall that

$$u(x,y+z) = P(x,y) * u(x,z) \quad \text{for all } y,\ z>0.$$

Hence, if we differentiate the both sides with respect to y (resp. x_i), we obtain that

$$\begin{aligned}u_y(x,y+z) &= P_y(x,y) * u(x,z) \quad \text{for all } y,\ z>0, \\ u_{x_i}(x,y+z) &= P_{x_i}(x,y) * u(x,z) \quad \text{for all } y,\ z>0.\end{aligned}$$

In particular, it follows that

$$\begin{aligned}u_y(x,y) &= P_y(x,y/2) * u(x,y/2) \quad \text{for all } y>0, \\ u_{x_i}(x,y) &= P_{x_i}(x,y/2) * u(x,y/2) \quad \text{for all } y>0.\end{aligned}$$

Therefore, we have, by part (b) of Lemma 5.17,

$$\begin{aligned}\|u_y(\cdot,y)\|_p &= \|P_y(\cdot,y/2) * u(\cdot,y/2)\|_p \\ &\leq \|P_y(\cdot,y/2)\|_1 \|u(\cdot,y/2)\|_p \\ &\leq 2(n+1)y^{-1} \|u(\cdot,y/2)\|_p \quad \text{for all } y>0.\end{aligned} \tag{5.35}$$

Similarly, we have the inequality

$$\|u_{x_i}(\cdot,y)\|_p \leq (n+1)y^{-1} \|u(\cdot,y/2)\|_p \quad \text{for all } y>0. \tag{5.36}$$

(a-1) The case $1 \leq q < \infty$: By inequality (5.35), it follows that

$$\begin{aligned}
\left\| y^{\alpha+1} u_y(\cdot,\cdot) \right\|_{pq} &= \left(\int_0^\infty \left(y^{\alpha+1} \left\| u_y(\cdot,y) \right\|_p \right)^q \frac{dy}{y} \right)^{1/q} \\
&\leq \left(\int_0^\infty \left(y^{\alpha+1} 2(n+1)\, y^{-1} \left\| u(\cdot,y/2) \right\|_p \right)^q \frac{dy}{y} \right)^{1/q} \\
&= 2^{\alpha+1}(n+1) \left(\int_0^\infty \left(z^\alpha \left\| u(\cdot,z) \right\|_p \right)^q \frac{dz}{z} \right)^{1/q} \\
&= 2^{\alpha+1}(n+1)\ \left\| y^\alpha u(\cdot,\cdot) \right\|_{pq}.
\end{aligned}$$

(a-2) The case $q = \infty$: We have, by inequality (5.36),

$$\begin{aligned}
\left\| y^{\alpha+1} u_y(\cdot,\cdot) \right\|_{p\infty} &= \sup_{y>0} \left\{ y^{\alpha+1} \left\| u_y(\cdot,y) \right\|_p \right\} \\
&\leq 2(n+1) \sup_{y>0} \left\{ y^\alpha \left\| u(\cdot,y/2) \right\|_p \right\} \\
&= 2^{\alpha+1}(n+1) \sup_{z>0} \left\{ z^\alpha \left\| u(\cdot,z) \right\|_p \right\} \\
&= 2^{\alpha+1}(n+1)\ \left\| y^\alpha u(\cdot,\cdot) \right\|_{p\infty}.
\end{aligned}$$

This completes the proof of part (a).

(b) Similarly, we have the inequality

$$\left\| y^{\alpha+1} u_{x_i}(\cdot,\cdot) \right\|_{pq} \leq 2^\alpha (n+1) \left\| y^\alpha u(\cdot,\cdot) \right\|_{pq}, \quad 1 \leq q \leq \infty.$$

(c) If $u(x,y) \to 0$ as $y \to \infty$, it follows that

$$u(x,y) = -\int_y^\infty u_s(x,s)\, ds \quad \text{for all } y > 0.$$

Hence, by applying Minkowski's inequality for integrals (Theorem 3.18) we obtain that

$$\left\| u(\cdot,y) \right\|_p \leq \int_y^\infty \left\| u_s(\cdot,s) \right\|_p \, ds \quad \text{for all } y > 0.$$

(c-1) The case $1 \leq q < \infty$: By Hardy's inequality (Theorem 3.20), it follows that

$$\begin{aligned}
\left\| y^\alpha u(\cdot,\cdot) \right\|_{pq} &= \left(\int_0^\infty \left(y^\alpha \left\| u(\cdot,y) \right\|_p \right)^{\alpha+1} \frac{dy}{y} \right)^{1/q} \\
&\leq \left(\int_0^\infty \left(y^\alpha \int_y^\infty \left\| u_s(\cdot,s) \right\|_p \, ds \right)^q \frac{dy}{y} \right)^{1/q}
\end{aligned}$$

$$\leq \frac{1}{\alpha} \left(\int_0^\infty \left(y^{\alpha+1} \|u_y(\cdot, y)\|_p \right)^q \frac{dy}{y} \right)^{1/q}$$
$$= \frac{1}{\alpha} \left\| y^{\alpha+1} u_y(\cdot, \cdot) \right\|_{pq}.$$

(c-2) The case $q = \infty$: We let

$$C_\alpha = \left\| y^{\alpha+1} u_y(\cdot, \cdot) \right\|_{p\infty}.$$

Then we have the inequality

$$\|u_y(\cdot, y)\|_p \leq C_\alpha y^{-\alpha-1} \quad \text{for all } y > 0,$$

so that

$$\|u(\cdot, y)\|_p \leq \int_y^\infty \|u_s(\cdot, s)\|_p \, ds \leq C_\alpha \alpha^{-1} y^{-\alpha} \quad \text{for all } y > 0.$$

This proves that

$$\|y^\alpha u(\cdot, \cdot)\|_{pq} = \sup_{y>0} \left\{ y^\alpha \|u(\cdot, y)\|_p \right\} \leq \frac{1}{\alpha} \left\| y^{\alpha+1} u_y(\cdot, \cdot) \right\|_{p\infty}.$$

The proof of Lemma 5.18 is complete. □

Step 2: First, we remark that, for each $y_0 > 0$, there exists a constant $A_{y_0} > 0$ such that

$$\|u(\cdot, y)\|_\infty \leq A_{y_0} \quad \text{for all } y > y_0.$$

Hence, by virtue of Lemma 5.17 it follows that

$$\|u_y(\cdot, y)\|_\infty \leq \frac{2(n+1)}{y} \|u(\cdot, y/2)\|_\infty \qquad (5.37)$$
$$\leq \frac{2(n+1)A_{y_0}}{y} \quad \text{for all } y \geq 2y_0,$$

and that

$$\|u_{x_i}(\cdot, y)\|_\infty \leq \frac{(n+1)}{y} \|u(\cdot, y/2)\|_\infty \qquad (5.38)$$
$$\leq \frac{(n+1)A_{y_0}}{y} \quad \text{for all } y \geq 2y_0.$$

Step 2-a: Since the function $u_y(x, y)$ is harmonic in $\mathbf{R}_+^{n+1}$ and satisfies the inequality (5.37), it follows from an application of part (b) of Lemma 5.18 that

$$\left\| y^{\alpha+1} u_{x_i y}(\cdot, \cdot) \right\|_{pq} \leq 2^\alpha (n+1) \|y^\alpha u_y(\cdot, \cdot)\|_{pq}. \qquad (5.39)$$

Moreover, we remark by inequality (5.38) that $u_{x_i}(x, y) \to 0$ as $y \to \infty$. Hence, by applying part (c) of Lemma 5.18 to the function $u_{x_i}(x, y)$ we obtain that

$$\|y^\alpha u_{x_i}(\cdot,\cdot)\|_{pq} \leq \frac{1}{\alpha} \left\|y^{\alpha+1} u_{x_i y}(\cdot,\cdot)\right\|_{pq}. \tag{5.40}$$

Therefore, it follows from inequalities (5.39) and (5.40) that

$$\|y^\alpha u_{x_i}(\cdot,\cdot)\|_{pq} \leq \frac{2^\alpha}{\alpha}(n+1) \|y^\alpha u_y(\cdot,\cdot)\|_{pq}.$$

This proves the desired part (a) of Theorem 5.14.

Step 2-b: Since the function $u_{x_i}(x, y)$ is harmonic in $\mathbf{R}_+^{n+1}$ and satisfies the inequality (5.38), it follows from an application of part (b) of Lemma 5.18 that

$$\left\|y^{\alpha+1} u_{x_i x_j}(\cdot,\cdot)\right\|_{pq} \leq 2^\alpha (n+1) \|y^\alpha u_{x_i}(\cdot,\cdot)\|_{pq}. \tag{5.41}$$

However, we remark that

$$u_{yy}(x, y) = -\sum_{i=1}^n u_{x_i x_i}(x, y).$$

Hence we have, by inequality (5.41),

$$\begin{aligned}
\left\|y^{\alpha+1} u_{yy}(\cdot,\cdot)\right\|_{pq} &\leq \sum_{i=1}^n \left\|y^{\alpha+1} u_{x_i x_i}(\cdot,\cdot)\right\|_{pq} \\
&\leq 2^\alpha (n+1) \sum_{i=1}^n \|y^\alpha u_{x_i}(\cdot,\cdot)\|_{pq} \\
&\leq 2^\alpha n(n+1) \sup_{1 \leq i \leq n} \|y^\alpha u_{x_i}(\cdot,\cdot)\|_{pq}.
\end{aligned} \tag{5.42}$$

Moreover, we remark by inequality (5.37) that $u_y(x, y) \to 0$ as $y \to \infty$. Hence, by applying part (c) of Lemma 5.18 to the function $u_y(x, y)$ we obtain that

$$\|y^\alpha u_y(\cdot,\cdot)\|_{pq} \leq \frac{1}{\alpha} \left\|y^{\alpha+1} u_{yy}(\cdot,\cdot)\right\|_{pq}. \tag{5.43}$$

Therefore, it follows from inequalities (5.42) and (5.43) that

$$\|y^\alpha u_y(\cdot,\cdot)\|_{pq} \leq \frac{2^\alpha}{\alpha} n(n+1) \sup_{1 \leq i \leq n} \|y^\alpha u_{x_i}(\cdot,\cdot)\|_{pq}.$$

This proves the desired part (b) of Theorem 5.14.

The proof of Theorem 5.14 is now complete. □

The next theorem establishes some fundamental relationships between means of the second derivatives of Poisson integrals

$$u(x, y) = \int_{\mathbf{R}^n} P(x - t, y)\, f(t)\, dt$$

taken with respect to the variable y and those taken with respect to the variables x_i:

Theorem 5.19. *Assume that a function $u(x, y)$ is harmonic in the half-space $\mathbf{R}_+^{n+1}$ and bounded in a proper sub-half space $\{(x, y) \in \mathbf{R}_+^{n+1} : y > y_0\}$ of $\mathbf{R}_+^{n+1}$, for each $y_0 > 0$. Then we have, for $\alpha > 0$, $1 \le p,\ q \le \infty$,*

$$\sup_{1 \le i, j \le n} \left\| y^\alpha u_{x_i x_j}(\cdot, \cdot) \right\|_{pq} \le M_\alpha \left\| y^\alpha u_{yy}(\cdot, \cdot) \right\|_{pq}. \tag{a}$$

$$\left\| y^\alpha u_{yy}(\cdot, \cdot) \right\|_{pq} \le M_\alpha \sup_{1 \le i, j \le n} \left\| y^\alpha u_{x_i x_j}(\cdot, \cdot) \right\|_{pq}. \tag{b}$$

Proof. Part (b) follows from the equation

$$u_{yy}(x, y) = -\sum_{i=1}^{n} u_{x_i x_i}(x, y).$$

On the other hand, part (a) follows by applying part (a) of Theorem 5.14 twice.

The proof of Theorem 5.19 is complete. □

5.6 Notes and Comments

The results of this chapter are adapted from Taibleson [72] and Folland [28] (see also [6, Chapter 1]).

6
Besov Spaces via Poisson Integrals

In this chapter we develop the theory of Besov spaces on the Euclidean space $\mathbf{R}^n$, paying particular attention to Poisson integrals. Besov spaces are function spaces defined in terms of the L^p modulus of continuity, and enter naturally in connection with boundary value problems in the framework of Sobolev spaces of L^p type. We prove a variety of equivalent norms for the Besov spaces on $\mathbf{R}^n$ via Poisson integrals (Theorems 6.3, 6.5 and 6.6).

6.1 Definition of Besov Spaces

Let $\alpha > 0$ and $1 \leq p,\, q \leq \infty$. We let

$$
\begin{aligned}
B^{\alpha}_{p,q}(\mathbf{R}^n) &= \text{the space of functions } f \in L^p(\mathbf{R}^n) \text{ for which} \qquad (6.1)\\
\|f\|_{\alpha;p,q} &= \|f\|_p + \left\| y^{\overline{\alpha}-\alpha} u_y^{(\overline{\alpha})}(\cdot,\cdot) \right\|_{pq} \\
&= \|f\|_p + \left(\int_0^\infty \left(y^{\overline{\alpha}-\alpha} \left\| u_y^{(\overline{\alpha})}(\cdot,y) \right\|_p \right)^q \frac{dy}{y} \right)^{1/q} < \infty.
\end{aligned}
$$

Here $\overline{\alpha}$ is the smallest integer greater than α and $u(x,y) = P(x,y) * f(x)$ is the Poisson integral of $f(x)$

$$
\begin{aligned}
u(x,y) &= P(x,y) * f(x) \qquad (5.16)\\
&= \frac{1}{c_n} \int_{\mathbf{R}^n} \frac{y}{(|x-z|^2 + y^2)^{(n+1)/2}} f(z)\, dz \quad \text{for all } y > 0,
\end{aligned}
$$

where c_n is a positive constant given by the formula

$$
c_n = \frac{\pi^{(n+1)/2}}{\Gamma((n+1)/2)}.
$$

Moreover, we recall that the mixed norm $\|\cdot\|_{pq}$ is defined by the formula

$$\|y^\alpha u(\cdot,\cdot)\|_{pq} = \begin{cases} \left(\int_0^\infty (y^\alpha \|u(\cdot,y)\|_p)^q \frac{dy}{y}\right)^{1/q} & \text{if } 1 \leq q < \infty, \\ \sup_{y>0} \left\{ y^\alpha \|u(\cdot,y)\|_p \right\} & \text{if } q = \infty. \end{cases}$$

First, we prove the completeness of $B^\alpha_{p,q}(\mathbf{R}^n)$:

Theorem 6.1. *The space $B^\alpha_{p,q}(\mathbf{R}^n)$ is a Banach space.*

Proof. We only consider the case:

$$0 < \alpha < 1, \quad \overline{\alpha} = 1.$$

Let $\{f^k\}$ be an arbitrary Cauchy sequence in $B^\alpha_{p,q}(\mathbf{R}^n)$, that is, we have, as k, $\ell \to \infty$,

$$\begin{aligned}\left\|f^k - f^\ell\right\|_{\alpha;p,q} &= \left\|f^k - f^\ell\right\|_p + \left\|y^{1-\alpha}\left(u^k_y(\cdot,\cdot) - u^\ell_y(\cdot,\cdot)\right)\right\|_{pq} \\ &\longrightarrow 0,\end{aligned} \tag{6.2}$$

where

$$u^k(x,y) = P(x,y) * f^k(x).$$

Then, since $L^p(\mathbf{R}^n)$ is complete, there exists a function $f \in L^p(\mathbf{R}^n)$ such that

$$\left\|f^k - f\right\|_p \longrightarrow 0 \quad \text{as } k \to \infty.$$

We let

$$u(x,y) = P(x,y) * f(x).$$

Then we obtain the following claim:

Claim 6.1. $\left\|y^{1-\alpha}\left(u^k_y(\cdot,\cdot) - u_y(\cdot,\cdot)\right)\right\|_{pq} \longrightarrow 0$ as $k \to \infty$.

Proof. Since we have the formulas

$$\begin{aligned} u^\ell_y(x,y) &= P_y(x,y) * f^\ell(x), \\ u_y(x,y) &= P_y(x,y) * f(x), \end{aligned}$$

it follows that we have, by Lemma 5.7,

$$\begin{aligned}\left\|u^\ell_y(\cdot,y) - u_y(\cdot,y)\right\|_p &\leq \|P_y(\cdot,y)\|_1 \left\|f^\ell - f\right\|_p \\ &\leq (n+1)y^{-1}\left\|f^\ell - f\right\|_p \quad \text{for all } y > 0.\end{aligned}$$

This implies that we have, for each $y > 0$,

$$\left\|y^{1-\alpha}\left(u^k_y(\cdot,y) - u^\ell_y(\cdot,y)\right)\right\|_p \longrightarrow \left\|y^{1-\alpha}\left(u^k_y(\cdot,y) - u_y(\cdot,y)\right)\right\|_p$$

as $\ell \to \infty$.

Hence, by applying Fatou's lemma (Theorem 3.7) we obtain from assertion (6.2) that

$$\begin{aligned}
&\left\| y^{1-\alpha} \left(u_y^k(\cdot,\cdot) - u_y(\cdot,\cdot) \right) \right\|_{pq}^q \\
&= \int_0^\infty \left(y^{1-\alpha} \left\| u_y^k(\cdot,y) - u_y(\cdot,y) \right\|_p \right)^q \frac{dy}{y} \\
&= \int_0^\infty \lim_{\ell\to\infty} \left(y^{1-\alpha} \left\| u_y^k(\cdot,y) - u_y^\ell(\cdot,y) \right\|_p \right)^q \frac{dy}{y} \\
&\le \liminf_{\ell\to\infty} \int_0^\infty \left(y^{1-\alpha} \left\| u_y^k(\cdot,y) - u_y^\ell(\cdot,y) \right\|_p \right)^q \frac{dy}{y} \\
&\le \sup_{\ell\ge k} \left\| y^{1-\alpha} \left(u_y^k(\cdot,\cdot) - u_y^\ell(\cdot,\cdot) \right) \right\|_{pq}^q \\
&\le \sup_{\ell\ge k} \left\| f^k - f^\ell \right\|_{\alpha;p,q}^q \longrightarrow 0 \quad \text{as } k \to \infty.
\end{aligned}$$

The proof of Claim 6.1 is complete. □

Therefore, we find from Claim 6.1 that

$$\begin{aligned}
&f \in B_{p,q}^\alpha(\mathbf{R}^n), \\
&f^k \longrightarrow f \quad \text{in } B_{p,q}^\alpha(\mathbf{R}^n),
\end{aligned}$$

since we have the assertion

$$\begin{aligned}
&\left\| y^{1-\alpha} u_y(\cdot,\cdot) \right\|_{pq} \\
&\le \left\| y^{1-\alpha} \left(u_y(\cdot,\cdot) - u_y^k(\cdot,\cdot) \right) \right\|_{pq} + \left\| y^{1-\alpha} u_y^k(\cdot,\cdot) \right\|_{pq} < \infty,
\end{aligned}$$

and also

$$\begin{aligned}
\left\| f^k - f \right\|_{\alpha;p,q} &= \left\| f^k - f \right\|_p + \left\| y^{1-\alpha} \left(u_y^k(\cdot,\cdot) - u_y(\cdot,\cdot) \right) \right\|_{pq} \\
&\longrightarrow 0 \quad \text{as } k \to \infty.
\end{aligned}$$

This proves the completeness of the Besov space $B_{p,q}^\alpha(\mathbf{R}^n)$.

The proof of Theorem 6.1 is complete. □

Proposition 6.2. *Let $f \in L^p(\mathbf{R}^n)$ with $1 \le p \le \infty$, and $u(x,y) = P(x,y) * f(x)$ its Poisson integral. Then, for any integer $k > \alpha$, the norm*

$$\|f\|_p + \left\| y^{k-\alpha} u_y^{(k)}(\cdot,\cdot) \right\|_{pq}$$

is equivalent to the norm $\|f\|_{\alpha;p,q}$.

Proof. It suffices to show that the ratio

$$\frac{\left\| y^{k-\alpha} u_y^{(k)}(\cdot,\cdot) \right\|_{pq}}{\left\| y^{\overline{\alpha}-\alpha} u_y^{(\overline{\alpha})}(\cdot,\cdot) \right\|_{pq}}$$

is bounded above and below by positive constants independent of f. However, this follows from a repeated application of part (a) and part (c) of Lemma 5.18.

The proof of Proposition 6.2 is complete. □

6.2 Various Norms of Besov Spaces

The next theorem gives a variety of equivalent norms for the Besov space $B^{\alpha}_{p,q}(\mathbf{R}^n)$:

Theorem 6.3. *Let f be a function in $L^p(\mathbf{R}^n)$ with $1 \le p,\ q \le \infty$, and let $u(x,y) = P(x,y) * f(x)$ be its Poisson integral. We define the eight norms A through H as follows:*

- $$A := \left\| \frac{f(\cdot + h) - f(\cdot)}{|h|^{\alpha}} \right\|_{pq} = \left(\int_{\mathbf{R}^n} \left(\frac{\|f(\cdot + h) - f(\cdot)\|_p}{|h|^{\alpha}} \right)^q \frac{dh}{|h|^n} \right)^{1/q},$$
  $0 < \alpha < 1$.
- $$B := \left\| \frac{f(\cdot + h) - 2f(\cdot) + 2f(\cdot - h)}{|h|^{\alpha}} \right\|_{pq} = \left(\int_{\mathbf{R}^n} \left(\frac{\|f(\cdot + h) - 2f(\cdot) + f(\cdot - h)\|_p}{|h|^{\alpha}} \right)^q \frac{dh}{|h|^n} \right)^{1/q},$$
  $0 < \alpha < 2$.
- $$C := \frac{1}{\omega_n} \left(\int_0^{\infty} \left(t^{-\alpha} \left\| \int_{\Sigma_{n-1}} (f(\cdot + t\sigma) - f(\cdot))\, d\sigma \right\|_p \right)^q \frac{dt}{t} \right)^{1/q},$$
  $0 < \alpha < 2$.
- $$D := \left(\int_0^{\infty} \left(y^{1-\alpha} \|u_y(\cdot, y)\|_p \right)^q \frac{dy}{y} \right)^{1/q} = \left\| y^{1-\alpha} u_y(\cdot,\cdot) \right\|_{pq},$$
  $0 < \alpha < 1$.

- $E := \left(\int_0^\infty (y^{2-\alpha} \|u_{yy}(\cdot, y)\|_p)^q \frac{dy}{y} \right)^{1/q} = \left\| y^{2-\alpha} u_{yy}(\cdot, \cdot) \right\|_{pq},$
  $0 < \alpha < 2.$
- $F := \left(\int_0^\infty \left(t^{-\alpha} \cdot \sup_{0<|h|\le t} \|f(\cdot + h) - f(\cdot)\|_p \right)^q \frac{dt}{t} \right)^{1/q},$
  $0 < \alpha < 1.$
- $G := \left(\int_0^\infty \left(t^{-\alpha} \cdot \sup_{0<|h|\le t} \|f(\cdot + h) - 2f(\cdot) + f(\cdot - h)\|_p \right)^q \frac{dt}{t} \right)^{1/q},$
  $0 < \alpha < 2.$
- $H := \left(\int_0^\infty \left(y^{-\alpha} \|u(\cdot, y) - f(\cdot)\|_p \right)^q \frac{dy}{y} \right)^{1/q},$
  $0 < \alpha < 1.$

Here Σ_{n-1} is the unit sphere in $\mathbf{R}^n$, $d\sigma$ is its surface element and ω_n is its surface area

$$\omega_n := |\Sigma_{n-1}| = \int_{\Sigma_{n-1}} d\sigma = \frac{2\pi^{n/2}}{\Gamma(n/2)}.$$

Then the norms A, B, C, D, E, F, G and H are equivalent for $0 < \alpha < 1$, while the norms B, C, E and G are equivalent for $0 < \alpha < 2$.

Proof. The proof is divided into ten steps.

Step 1: $E \le 2^{2-\alpha}(n+1)D$ and $D \le 1/(1-\alpha)E$ for $0 < \alpha < 1$.

Assume that

$$u(x, y) = P(x, y) * f(x) \quad \text{for } f \in L^p(\mathbf{R}^n).$$

Then we have, by part (i) of Theorem 5.8,

$$\|u(\cdot, y)\|_p = \|P(\cdot, y) * f\|_p \le \|f\|_p \quad \text{for all } y > 0.$$

Hence, by Lemma 5.11 it follows that

$$\|u(\cdot, y)\|_\infty \le \left(\frac{2^{n+1}}{V_{n+1}} \right)^{1/p} \|f\|_p \, y^{-n/p} \quad \text{for all } y > 0,$$

where V_{n+1} is the volume of the unit ball in $\mathbf{R}^{n+1}$

$$V_{n+1} = \frac{\omega_{n+1}}{n+1} = \frac{\pi^{(n+1)/2}}{\Gamma((n+1)/2 + 1)}.$$

By applying Lemma 5.17 with $p := \infty$, we obtain that, for each $y_0 > 0$,

$$\|u_y(\cdot, y)\|_\infty \le 2(n+1)\, A_{y_0}\, y^{-1} \quad \text{for all } y \ge 2y_0, \tag{6.3}$$

with

$$A_{y_0} := \left(\frac{2^{n+1}}{V_{n+1}}\right)^{1/p} \|f\|_p \; y_0^{-n/p}.$$

Therefore, it follows from an application of part (a) of Lemma 5.18 with $\alpha := 1 - \alpha$ that

$$\begin{aligned} E &= \left\|y^{2-\alpha} u_{yy}(\cdot, \cdot)\right\|_{pq} \le 2^{2-\alpha}(n+1) \left\|y^{1-\alpha} u_y(\cdot, \cdot)\right\|_{pq} \\ &= 2^{2-\alpha}(n+1)D, \quad 0 < \alpha < 1. \end{aligned}$$

On the other hand, we find from inequality (6.3) that $u_y(x, y) \to 0$ as $y \to \infty$. Hence, by applying part (c) of Lemma 5.18 with $u(x, y) := u_y(x, y)$ and $\alpha := 1 - \alpha$ we obtain that

$$\begin{aligned} D &= \left\|y^{1-\alpha} u_y(\cdot, \cdot)\right\|_{pq} \le \frac{1}{1-\alpha} \left\|y^{2-\alpha} u_{yy}(\cdot, \cdot)\right\|_{pq} \\ &= \frac{1}{1-\alpha} E, \quad 0 < \alpha < 1. \end{aligned}$$

Step 2: $C \le (2\omega_n^{1/q})^{-1} B$.

Step 2-a: The case $q = \infty$. By Minkowski's inequality for integrals (Theorem 3.16), it follows that

$$\begin{aligned} C &= \frac{1}{\omega_n} \sup_{t>0} \left\{ t^{-\alpha} \left\| \int_{\Sigma_{n-1}} (f(\cdot + t\sigma) - f(\cdot))\, d\sigma \right\|_p \right\} \\ &= \frac{1}{2} \frac{1}{\omega_n} \sup_{t>0} \left\{ t^{-\alpha} \left\| \int_{\Sigma_{n-1}} (f(\cdot + t\sigma) - 2f(\cdot) + f(\cdot - t\sigma))\, d\sigma \right\|_p \right\} \\ &\le \frac{1}{2} \frac{1}{\omega_n} \sup_{t>0} \left\{ \int_{\Sigma_{n-1}} \frac{\|f(\cdot + t\sigma) - 2f(\cdot) + f(\cdot - t\sigma)\|_p}{t^\alpha}\, d\sigma \right\} \\ &\le \frac{1}{2} \frac{1}{\omega_n} \int_{\Sigma_{n-1}} \left(\sup_{h \in \mathbf{R}^n} \frac{\|f(\cdot + h) - 2f(\cdot) + f(\cdot - h)\|_p}{|h|^\alpha} \right) d\sigma \\ &= \frac{1}{2} B, \quad 0 < \alpha < 1. \end{aligned}$$

Step 2-b: The case $1 \le q < \infty$. By applying Hölder's inequality (Theorem 3.14) with $q' = q/(q-1)$, we obtain that

C

$$
\begin{aligned}
&= \frac{1}{\omega_n} \left(\int_0^\infty \left(t^{-\alpha} \left\| \int_{\Sigma_{n-1}} (f(\cdot + t\sigma) - f(\cdot))\, d\sigma \right\|_p \right)^q \frac{dt}{t} \right)^{1/q} \\
&= \frac{1}{2} \frac{1}{\omega_n} \left(\int_0^\infty \left\| \int_{\Sigma_{n-1}} (f(\cdot + t\sigma) - 2f(\cdot) + f(\cdot - t\sigma))\, d\sigma \right\|_p^q \frac{dt}{t^{1+\alpha q}} \right)^{1/q} \\
&\le \frac{1}{2} \frac{1}{\omega_n} \left(\int_0^\infty \left(\int_{\Sigma_{n-1}} \|f(\cdot + t\sigma) - 2f(\cdot) + f(\cdot - t\sigma)\|_p \, d\sigma \right)^q \frac{dt}{t^{1+\alpha q}} \right)^{1/q} \\
&\le \frac{1}{2} \frac{1}{\omega_n} \left(\int_0^\infty \int_{\Sigma_{n-1}} \|f(\cdot + t\sigma) - 2f(\cdot) + f(\cdot - t\sigma)\|_p^q \, d\sigma \, \omega_n^{q/q'} \frac{dt}{t^{1+\alpha q}} \right)^{1/q} \\
&\le \frac{1}{2} \frac{1}{\omega_n^{1/q}} \left(\int_0^\infty \int_{\Sigma_{n-1}} \|f(\cdot + t\sigma) - 2f(\cdot) + f(\cdot - t\sigma)\|_p^q \, d\sigma \frac{dt}{t^{1+\alpha q}} \right)^{1/q} \\
&= \frac{1}{2} \frac{1}{\omega_n^{1/q}} \left(\int_{\mathbf{R}^n} \left(\frac{\|f(\cdot + h) - 2f(\cdot) + f(\cdot - h)\|_p}{|h|^\alpha} \right)^q \frac{dh}{|h|^n} \right)^{1/q} \\
&= \frac{1}{2\omega_n^{1/q}} B, \quad 0 < \alpha < 1.
\end{aligned}
$$

Step 3: $B \le 2A$. We remark that

$$
\begin{aligned}
\|f(\cdot + h) - 2f(\cdot) + f(\cdot - h)\|_p &\le \|f(\cdot + h) - f(\cdot)\|_p + \|f(\cdot) - f(\cdot - h)\|_p \\
&\le 2 \|f(\cdot + h) - f(\cdot)\|_p .
\end{aligned}
$$

Hence we have the inequality

$$
\begin{aligned}
B &= \left(\int_{\mathbf{R}^n} \left(\frac{\|f(\cdot + h) - 2f(\cdot) + f(\cdot - h)\|_p}{|h|^\alpha} \right)^q \frac{dh}{|h|^n} \right)^{1/q} \\
&\le \left(\int_{\mathbf{R}^n} \left(\frac{2\|f(\cdot + h) - f(\cdot)\|_p}{|h|^\alpha} \right)^q \frac{dh}{|h|^n} \right)^{1/q} \\
&= 2A, \quad 0 < \alpha < 1.
\end{aligned}
$$

Step 4: $A \le \omega_n^{1/q} F$. This is obvious. Indeed, we have the inequality

$$
\begin{aligned}
A &= \left(\int_{\mathbf{R}^n} \left(\frac{\|f(\cdot + h) - f(\cdot)\|_p}{|h|^\alpha} \right)^q \frac{dh}{|h|^n} \right)^{1/q} \\
&= \left(\int_0^\infty \int_{\Sigma_{n-1}} \|f(\cdot + t\sigma) - f(\cdot)\|_p^q \frac{dt}{t^{1+\alpha q}} \, d\sigma \right)^{1/q}
\end{aligned}
$$

$$\leq \omega_n^{1/q} \left(\int_0^\infty \left(t^{-\alpha} \cdot \sup_{0<|h|\leq t} \|f(\cdot + h) - f(\cdot)\|_p \right)^q \frac{dt}{t} \right)^{1/q}$$
$$= \omega_n^{1/q} F, \quad 0 < \alpha < 1.$$

Step 5: $F \leq M_\alpha D$. We recall (see Theorem 5.9) that

$$\lim_{y \downarrow 0} u(x, y) = f(x),$$

whenever x belongs to the Lebesgue set $L(f)$ of f. We let

$$\mathcal{L} = \{(x, h) \in \mathbf{R}^n \times \mathbf{R}^n : x + h \in L(f)\}.$$

By Fubini's theorem (Theorem 3.10), we find that the complement of $\mathcal{L}$ has measure zero.

Now we assume that

$$(x, 0), (x, h) \in \mathcal{L} \quad \text{with } 0 < |h| \leq t.$$

Then we have, for $h = s\sigma$, $0 < s \leq t$ and $\sigma \in \Sigma_{n-1}$,

$$\begin{aligned} & f(x+h) - f(x) \\ &= (f(x+h) - u(x+h, t)) + (u(x+h, t) - u(x, t)) + (u(x, t) - f(x)) \\ &= -\int_0^t u_y(x + s\sigma, y)\, dy + \int_0^s u_r(x + r\sigma, t)\, dr + \int_0^t u_y(x, y)\, dy. \end{aligned}$$

Hence it follows from an application of Minkowski's inequality for integrals (Theorem 3.16) that

$$\begin{aligned} \|f(\cdot + h) - f(\cdot)\|_p &\leq \int_0^t \|u_y(\cdot + s\sigma, y)\|_p \, dy + \int_0^s \|u_r(\cdot + r\sigma, t)\|_p \, dr \\ &\quad + \int_0^t \|u_y(\cdot, y)\|_p \, dy \\ &\leq 2 \int_0^t \|u_y(\cdot, y)\|_p \, dy + t \|u_r(\cdot, t)\|_p \\ &\leq 2 \int_0^t \|u_y(\cdot, y)\|_p \, dy + t \sum_{i=1}^n \|u_{x_i}(\cdot, t)\|_p, \end{aligned}$$

since we have the formula

$$u_r(x + r\sigma, t) = \sum_{i=1}^n \frac{\partial u}{\partial x_i}(x + r\sigma, t)\sigma_i.$$

Step 5-a: The case $q = \infty$. Since we have the inequality

$$\|u_y(\cdot, y)\|_p \leq D y^{\alpha - 1} \quad \text{for all } y > 0,$$

it follows from part (a) of Theorem 5.14 that

$$\sup_{1\le i\le n} \|u_{x_i}(\cdot,y)\|_p \le M'_\alpha \|u_y(\cdot,y)\|_p \le M'_\alpha D y^{\alpha-1} \quad \text{for all } y>0.$$

Thus we obtain that

$$\begin{aligned} \|f(\cdot+h)-f(\cdot)\|_p &\le 2\int_0^t \|u_y(\cdot,y)\|_p\,dy + t\sum_{i=1}^n \|u_{x_i}(\cdot,t)\|_p \qquad (6.4)\\ &\le 2D\int_0^t y^{\alpha-1}\,dy + M'_\alpha D\,nt\,t^{\alpha-1}\\ &= \left(\frac{2}{\alpha}+nM'_\alpha\right) D\,t^\alpha \quad \text{for all } 0<|h|\le t. \end{aligned}$$

This proves that

$$t^{-\alpha}\cdot \sup_{0<|h|\le t} \|f(\cdot+h)-f(\cdot)\|_p \le \left(\frac{2}{\alpha}+nM'_\alpha\right) D,$$

so that

$$F=\sup_{t>0}\left(t^{-\alpha}\cdot \sup_{0<|h|\le t}\|f(\cdot+h)-f(\cdot)\|_p\right) \le M_\alpha D,$$

with

$$M_\alpha := \frac{2}{\alpha}+nM'_\alpha, \quad 0<\alpha<1.$$

Step 5-b: The case $1\le q<\infty$. In view of inequality (6.4), we have, for some positive constant M_q,

$$\begin{aligned} F &= \left(\int_0^\infty \left(t^{-\alpha}\cdot \sup_{0<|h|\le t}\|f(\cdot+h)-f(\cdot)\|_p\right)^q \frac{dt}{t}\right)^{1/q} \qquad (6.5)\\ &\le \left(\int_0^\infty t^{-\alpha q}\left(2\int_0^t \|u_y(\cdot,y)\|_p\,dy + t\sum_{i=1}^n \|u_{x_i}(\cdot,t)\|_p\right)^q \frac{dt}{t}\right)^{1/q}\\ &\le M_q\left[\left(\int_0^\infty \left(t^{-\alpha}\int_0^t \|u_y(\cdot,y)\|_p\,dy\right)^q \frac{dt}{t}\right)^{1/q}\right.\\ &\quad \left. + \sum_{i=1}^n \left(\int_0^\infty \left(t^{1-\alpha}\|u_{x_i}(\cdot,t)\|_p\right)^q \frac{dt}{t}\right)^{1/q}\right]. \end{aligned}$$

However, we have, by Hardy's inequality (Theorem 3.18),

$$\left(\int_0^\infty \left(t^{-\alpha}\int_0^t \|u_y(\cdot,y)\|_p\,dy\right)^q \frac{dt}{t}\right)^{1/q} \qquad (6.6)$$

$$\leq \frac{1}{\alpha}\left(\int_0^\infty \left(y^{1-\alpha}\,\|u_y(\cdot,y)\|_p\right)^q \frac{dy}{y}\right)^{1/q}$$
$$= \frac{1}{\alpha}D.$$

Furthermore, we have, by part (a) of Theorem 5.14,

$$\sum_{i=1}^n \left(\int_0^\infty \left(t^{1-\alpha}\,\|u_{x_i}(\cdot,t)\|_p\right)^q \frac{dt}{t}\right)^{1/q} \tag{6.7}$$
$$\leq M'_\alpha \left(\int_0^\infty \left(y^{1-\alpha}\,\|u_y(\cdot,y)\|_p\right)^q \frac{dy}{y}\right)^{1/q}$$
$$= M'_\alpha D.$$

Therefore, by combining inequalities (6.5), (6.6) and (6.7) we find that

$$F \leq M_\alpha D,$$

with

$$M_\alpha := M_q\left(\frac{1}{\alpha} + M'_\alpha\right), \quad 0 < \alpha < 1.$$

Step 6: $B \leq \omega_n^{1/q} G$. This is obvious. Indeed, we have the inequality

$$B$$
$$= \left(\int_{\mathbf{R}^n} \left(\frac{\|f(\cdot+h) - 2f(\cdot) + f(\cdot-h)\|_p}{|h|^\alpha}\right)^q \frac{dh}{|h|^n}\right)^{1/q}$$
$$= \left(\int_0^\infty \int_{\Sigma_{n-1}} \|f(\cdot+t\sigma) - 2f(\cdot) + f(\cdot-t\sigma)\|_p^q \, d\sigma \frac{dt}{t^{1+\alpha q}}\right)^{1/q}$$
$$\leq \omega_n^{1/q}\left(\int_0^\infty \left(t^{-\alpha}\cdot \sup_{0<|h|\leq t} \|f(\cdot+h) - 2f(\cdot) + f(\cdot-h)\|_p\right)^q \frac{dt}{t}\right)^{1/q}$$
$$= \omega_n^{1/q} G, \quad 0 < \alpha < 2.$$

Step 7: $G \leq M_\alpha E$. We have, for $h = s\sigma$, $0 < s \leq t$, $\sigma \in \Sigma_{n-1}$ and $\varepsilon > 0$,

$$\begin{aligned} u(x \pm h, \varepsilon) &= \big\{[t u_y(x \pm h, t) - \varepsilon u_y(x \pm h, \varepsilon)] \\ &\quad - [u(x \pm h, t) - u(x \pm h, \varepsilon)]\big\} \\ &\quad - [t u_y(x \pm h, t) - \varepsilon u_y(x \pm h, \varepsilon)] + u(x \pm h, t) \\ &= \int_\varepsilon^t y u_{yy}(x \pm h, y)\, dy - t u_y(x \pm h, t) + u(x \pm h, t) \end{aligned}$$

$$+ \varepsilon u_y(x \pm h, \varepsilon),$$

and

$$u(x, \varepsilon) = \int_\varepsilon^t y u_{yy}(x, y)\, dy - t u_y(x, t) + u(x, t) + \varepsilon u_y(x, \varepsilon).$$

Hence it follows that

$$\begin{aligned}
&u(x+h, \varepsilon) - 2u(x, \varepsilon) + u(x-h, \varepsilon) \\
&= \int_\varepsilon^t y u_{yy}(x+h, y)\, dy + \int_\varepsilon^t y u_{yy}(x-h, y)\, dy - 2 \int_\varepsilon^t y u_{yy}(x, y)\, dy \\
&\quad - t\left[u_y(x+h, t) - u_y(x, t)\right] - t\left[u_y(x-h, t) - u_y(x, t)\right] \\
&\quad + \left[u(x+h, t) - 2u(x, t) + u(x-h, t)\right] \\
&\quad + \varepsilon\left[u_y(x+h, \varepsilon) - 2u_y(x, \varepsilon) + u_y(x-h, \varepsilon)\right].
\end{aligned}$$

Thus, by applying Minkowski's inequality for integrals (Theorem 3.16) we obtain that

$$\begin{aligned}
&\left\|u(\cdot + h, \varepsilon) - 2u(\cdot, \varepsilon) + u(\cdot - h, \varepsilon)\right\|_p \qquad (6.8) \\
&\le \int_\varepsilon^t y \left\|u_{yy}(\cdot + h, y)\right\|_p\, dy + \int_\varepsilon^t y \left\|u_{yy}(\cdot - h, y)\right\|_p\, dy \\
&\quad + 2 \int_\varepsilon^t y \left\|u_{yy}(\cdot, y)\right\|_p\, dy \\
&\quad + t \left\|u_y(\cdot + h, t) - u_y(\cdot, t)\right\|_p + t \left\|u_y(\cdot - h, t) - u_y(\cdot, t)\right\|_p \\
&\quad + \left\|u(\cdot + h, t) - 2u(\cdot, t) + u(\cdot - h, t)\right\|_p \\
&\quad + \varepsilon \left\|u_y(\cdot + h, \varepsilon) - 2u_y(\cdot, \varepsilon) + u_y(\cdot - h, \varepsilon)\right\|_p \\
&\le 4 \int_\varepsilon^t y \left\|u_{yy}(\cdot, y)\right\|_p\, dy + 2t \left\|u_y(\cdot + h, t) - u_y(\cdot, t)\right\|_p \\
&\quad + \left\|u(\cdot + h, t) - 2u(\cdot, t) + u(\cdot - h, t)\right\|_p + 4\varepsilon \left\|u_y(\cdot, \varepsilon)\right\|_p \\
&\le 4 \int_\varepsilon^t y \left\|u_{yy}(\cdot, y)\right\|_p\, dy + 2t \sum_{i=1}^n \int_0^s \left\|u_{yx_i}(\cdot + r\sigma, t)\right\|_p\, dr \\
&\quad + \sum_{i,j=1}^n \int_0^s \int_0^r \left\|u_{x_i x_j}(\cdot + v\sigma, t)\right\|_p\, dv\, dr \\
&\quad + \sum_{i,j=1}^n \int_0^s \int_0^r \left\|u_{x_i x_j}(\cdot - v\sigma, t)\right\|_p\, dv\, dr + 4\varepsilon \left\|u_y(\cdot, \varepsilon)\right\|_p \\
&\le 4 \int_\varepsilon^t y \left\|u_{yy}(\cdot, y)\right\|_p\, dy + 2t^2 \sum_{i=1}^n \left\|u_{yx_i}(\cdot, t)\right\|_p
\end{aligned}$$

$$+ t^2 \sum_{i,j=1}^{n} \left\| u_{x_i x_j}(\cdot, t) \right\|_p + 4\varepsilon \left\| u_y(\cdot, \varepsilon) \right\|_p .$$

Step 7-1: Here we need the following lemma:

Lemma 6.4. *Assume that $u(x, y)$ is harmonic in the half-space $\mathbf{R}_+^{n+1}$ and bounded in each proper subhalf space of $\mathbf{R}_+^{n+1}$, and that we are given $\alpha > 0$, an integer $k > \alpha$, $D > 0$ and $y_0 > 0$ such that*

$$\left\| y^{k-\alpha} u_y^{(k)}(\cdot, \cdot) \right\|_{pq} \leq D, \tag{6.9}$$

$$\|u(\cdot, y)\|_p \leq D \quad \textit{for all } y \geq y_0. \tag{6.10}$$

Then it follows that the function $u(x, y)$ is the Poisson integral of some function $f \in B_{p,q}^{\alpha}(\mathbf{R}^n)$. Moreover, we have the following two assertions (a) and (b):

$$\|u_y(\cdot, y)\|_p = o(y^{-1}) \quad \textit{as } y \downarrow 0. \tag{a}$$

$$\|f\|_{\alpha;p,q} \leq M_{\alpha,k,y_0} D. \tag{b}$$

Proof. The proof of Lemma 6.4 is divided into two steps.

Step (I): First, we prove assertion (a).

(a-1) The case $0 < \alpha < 1$: We remark that

$$u(x, y) = O(1) \quad \text{as } y \to \infty,$$

uniformly in $x \in \mathbf{R}^n$. However, we have, by the mean value theorem,

$$u(x, y) = u(x, y_0) + u_y(x, z)(y - y_0) \quad \text{for some } z \in (y_0, y).$$

Hence it follows that

$$u_y(x, y) = o(1) \quad \text{as } y \to \infty,$$

uniformly in $x \in \mathbf{R}^n$.

Now, by a repeated application of part (c) of Lemma 5.18 we find from inequality (6.9) that

$$\begin{aligned} \left\| y^{1-\alpha} u_y(\cdot, y) \right\|_{pq} &\leq M_\alpha \left\| y^{2-\alpha} u_{yy}(\cdot, y) \right\|_{pq} \\ &\leq M_{\alpha,k} \left\| y^{k-\alpha} u^{(k)}(\cdot, y) \right\|_{pq} \\ &\leq M_{\alpha,k}\, D \quad \text{for all } y > 0. \end{aligned} \tag{6.11}$$

However, it is easy to see that the function $\|u_y(\cdot, y)\|_p$ is non-increasing of y. Indeed, since $u_y(x, y)$ is harmonic in $\mathbf{R}_+^{n+1}$, it follows that

$$\|u_y(\cdot, y_1 + y_2)\|_p = \|P(\cdot, y_1) * u_y(\cdot, y_2)\|_p$$

$$\leq \|P(\cdot, y_1)\|_1 \|u_y(\cdot, y_2)\|_p$$
$$= \|u_y(\cdot, y_2)\|_p \quad \text{for all } y_1 \text{ and } y_2 > 0.$$

Thus we have, by inequality (6.11),

$$\begin{aligned}\frac{y^{1-\alpha}}{((1-\alpha)q)^{1/q}} \|u_y(\cdot, y)\|_p &= \left(\int_0^y \left(t^{(1-\alpha)} \|u_y(\cdot, y)\|_p\right)^q \frac{dt}{t}\right)^{1/q} \\ &\leq \left(\int_0^y \left(t^{(1-\alpha)} \|u_y(\cdot, t)\|_p\right)^q \frac{dt}{t}\right)^{1/q} \\ &\leq \left(\int_0^\infty \left(t^{(1-\alpha)} \|u_y(\cdot, t)\|_p\right)^q \frac{dt}{t}\right)^{1/q} \\ &= \left\|y^{1-\alpha} u_y(\cdot, \cdot)\right\|_{pq} \\ &\leq M_{\alpha,k} D \quad \text{for all } y > 0.\end{aligned}$$

This proves that

$$y \|u_y(\cdot, y)\|_p \leq M_{\alpha,k} ((1-\alpha)q)^{1/q} D y^\alpha \quad \text{for all } y > 0. \tag{6.12}$$

(a-2) The case $\alpha \geq 1$: We remark that, by inequality (6.9),

$$\begin{aligned}&\left(\int_0^{2y_0} \left(y^{k-1/2} \left\|u_y^{(k)}(\cdot, y)\right\|_p\right)^q \frac{dy}{y}\right)^{1/q} \qquad (6.13) \\ &= \left(\int_0^{2y_0} y^{q(\alpha-1/2)} \left(y^{k-\alpha} \left\|u_y^{(k)}(\cdot, y)\right\|_p\right)^q \frac{dy}{y}\right)^{1/q} \\ &\leq (2y_0)^{(\alpha-1/2)} D.\end{aligned}$$

On the other hand, it follows from inequality (6.10) that, for all $y \geq 2y_0$,

$$\begin{aligned}\left\|u_y^{(k)}(\cdot, y)\right\|_p &= \left\|P_y^{(k)}(\cdot, y/2) * u(\cdot, y/2)\right\|_p \\ &\leq \left\|P_y^{(k)}(\cdot, y/2)\right\|_1 \|u(\cdot, y/2)\|_p \leq \frac{M_k}{y^k} \|u(\cdot, y/2)\|_p \\ &\leq (M_k D) y^{-k}.\end{aligned}$$

Thus we have the inequality

$$\begin{aligned}&\left(\int_{2y_0}^\infty \left(y^{k-1/2} \left\|u_y^{(k)}(\cdot, y)\right\|_p\right)^q \frac{dy}{y}\right)^{1/q} \qquad (6.14) \\ &\leq M_k D \left(\int_{2y_0}^\infty \left(y^{k-1/2} y^{-k}\right)^q \frac{dy}{y}\right)^{1/q} \\ &= \frac{2M_k D}{q} (2y_0)^{-q/2}.\end{aligned}$$

Therefore, by combining inequalities (6.13) and (6.14) we obtain that

$$\begin{aligned}
&\left(\int_0^\infty \left(y^{k-1/2}\left\|u_y^{(k)}(\cdot,y)\right\|_p\right)^q \frac{dy}{y}\right)^{1/q} \\
&= \left(\int_0^{2y_0} \left(y^{k-1/2}\left\|u_y^{(k)}(\cdot,y)\right\|_p\right)^q \frac{dy}{y}\right. \\
&\quad \left. + \int_{2y_0}^\infty \left(y^{k-1/2}\left\|u_y^{(k)}(\cdot,y)\right\|_p\right)^q \frac{dy}{y}\right)^{1/q} \\
&\leq M_{\alpha,k,y_0} D,
\end{aligned}$$

where M_{α,k,y_0} is a positive constant given by the formula

$$M_{\alpha,k,y_0} := (2y_0)^{(\alpha-1/2)} + \frac{2M_k}{q}(2y_0)^{-q/2}.$$

Hence it follows from a repeated application of part (c) of Lemma 5.18 that

$$\begin{aligned}
&\left(\int_0^\infty \left(y^{1/2}\left\|u_y(\cdot,y)\right\|_p\right)^q \frac{dy}{y}\right)^{1/q} \\
&\leq M_k \left(\int_0^\infty \left(y^{k-1/2}\left\|u_y^{(k)}(\cdot,y)\right\|_p\right)^q \frac{dy}{y}\right)^{1/q} \\
&\leq M_k M_{\alpha,k,y_0} D.
\end{aligned}$$

This proves that

$$\begin{aligned}
s^{1/2}\left\|u_y(\cdot,s)\right\|_p &= \left(\frac{q}{2}\right)^{1/q}\left(\int_0^s \left(y^{1/2}\left\|u_y(\cdot,s)\right\|_p\right)^q \frac{dy}{y}\right)^{1/q} \\
&\leq M_k M_{\alpha,k,y_0}\, D \quad \text{for all } s > 0,
\end{aligned}$$

or equivalently,

$$y\left\|u_y(\cdot,y)\right\|_p \leq M_k M_{\alpha,k,y_0} D y^{1/2} \quad \text{for all } y > 0. \tag{6.15}$$

Summing up, we obtain from inequalities (6.12) and (6.15) that we have, for some $\beta \in (0,1)$,

$$\left\|u_y(\cdot,y)\right\|_p \leq (M'_{\alpha,k,y_0} D) y^{\beta-1} \quad \text{for all } y > 0, \tag{6.16}$$

In particular, it follows that

$$\left\|u_y(\cdot,y)\right\|_p = o(y^{-1}) \quad \text{as } y \downarrow 0.$$

This proves the desired assertion (a).

Step (II): Secondly, we prove assertion (b). To do this, we show that there exists a function $f \in B^{\alpha}_{p,q}(\mathbf{R}^n)$ such that

$$u(x,y) = P(x,y) * f(x).$$

We remark that

$$u(x,y) = u(x,y_0) + \int_y^{y_0} u_s(x,s)\,ds \quad \text{for all } 0 < y \leq y_0.$$

Hence we have, by inequality (6.16),

$$\begin{aligned} \|u(\cdot,y)\|_p &\leq \|u(\cdot,y_0)\|_p + \int_y^{y_0} \|u_s(\cdot,s)\|_p \,ds \\ &\leq D + M'_{\alpha,k,y_0}\, D \int_y^{y_0} s^{\beta-1}\,ds \leq D\left(1 + M'_{\alpha,k,y_0}\frac{y_0^\beta}{\beta}\right) \\ &:= M''_{\alpha,k,y_0}\, D \quad \text{for all } 0 < y \leq y_0, \end{aligned} \tag{6.17}$$

where M''_{α,k,y_0} is a positive constant given by the formula

$$M''_{\alpha,k,y_0} := 1 + M'_{\alpha,k,y_0}\frac{y_0^\beta}{\beta}, \quad 0 < \beta < 1.$$

By a similar argument, it follows that we have, for some positive constant M'''_{α,k,y_0},

$$\|u(\cdot,y_1) - u(\cdot,y_2)\|_p \leq \left|\int_{y_2}^{y_1} \|u_s(\cdot,s)\|_p \,ds\right| \leq M'''_{\alpha,k,y_0} D \left|y_2^\beta - y_1^\beta\right|.$$

This implies that $\{u(\cdot,y)\}$ is a Cauchy sequence in $L^p(\mathbf{R}^n)$ as $y \downarrow 0$. Thus, by arguing as in the proof of Theorem 5.10 we find that the function $u(x,y)$ is the Poisson integral of a function $f \in L^p(\mathbf{R}^n)$:

$$u(x,y) = P(x,y) * f(x).$$

Furthermore, by letting $y \downarrow 0$ in inequality (6.17) we obtain that

$$\|f\|_p = \lim_{y\downarrow 0} \|u(\cdot,y)\|_p \leq M''_{\alpha,k,y_0} D. \tag{6.18}$$

Therefore, in view of Proposition 6.2, it follows from inequalities (6.9) and (6.18) that

$$\|f\|_{\alpha;p,q} = \left\|y^{k-\alpha} u_y^{(k)}(\cdot,y)\right\|_{pq} + \|f\|_p \leq \left(1 + M''_{\alpha,k,y_0}\right) D.$$

This proves the desired assertion (b), with

$$M_{\alpha,k,y_0} := 1 + M''_{\alpha,k,y_0}.$$

The proof of Lemma 6.4 is complete. □

Step 7-2: Now we recall that

$$E = \left\| y^{2-\alpha} u_{yy}(\cdot,\cdot) \right\|_{pq},$$

and that

$$\|u(\cdot,y)\|_\infty \le \|P(\cdot,y)\|_{p'} \|f\|_p \le \left(\frac{1}{c_n}\right)^{1/p} y^{-n/p} \|f\|_p \quad \text{for all } y > 0.$$

Hence, by applying Lemma 6.4 to our situation we find that

$$\varepsilon \|u_y(\cdot,\varepsilon)\|_p \longrightarrow 0 \quad \text{as } \varepsilon \downarrow 0.$$

Furthermore, we remark that

$$\begin{aligned} &u(\cdot+h,\varepsilon) - 2u(\cdot,\varepsilon) + u(\cdot-h,\varepsilon) \\ &= P(\cdot,\varepsilon) * (f(\cdot+h) - 2f(\cdot) + f(\cdot-h)) \longrightarrow f(\cdot+h) - 2f(\cdot) + f(\cdot-h) \\ &\quad \text{in } L^p(\mathbf{R}^n) \text{ as } \varepsilon \downarrow 0. \end{aligned}$$

Therefore, by letting $\varepsilon \downarrow 0$ in inequality (6.8) we obtain that

$$\begin{aligned} &\|f(\cdot+h) - 2f(\cdot) + f(\cdot-h)\|_p \\ &\le 4\int_0^t y \|u_{yy}(\cdot,y)\|_p \, dy + 2t^2 \sum_{i=1}^n \|u_{yx_i}(\cdot,t)\|_p + t^2 \sum_{i,j=1}^n \left\|u_{x_i x_j}(\cdot,t)\right\|_p. \end{aligned} \tag{6.19}$$

However, we have, by Theorems 5.14 and 5.19,

$$t^2 \|u_{yx_i}(\cdot,t)\|_p \le Mt^2 \|u_{yy}(\cdot,t)\|_p \quad \text{for all } t > 0, \tag{6.20}$$

$$t^2 \left\|u_{x_i x_j}(\cdot,t)\right\|_p \le Mt^2 \|u_{yy}(\cdot,t)\|_p \quad \text{for all } t > 0. \tag{6.21}$$

Therefore, by combining inequalities (6.19), (6.20) and (6.21) we obtain that

$$\begin{aligned} &\|f(\cdot+h) - 2f(\cdot) + f(\cdot-h)\|_p \\ &\le 4\int_0^t y \|u_{yy}(\cdot,y)\|_p \, dy + 2Mt^2 \|u_{yy}(\cdot,t)\|_p. \end{aligned} \tag{6.22}$$

Step 7-2a: The case $q = \infty$. Since we have the inequality

$$\|u_{yy}(\cdot,y)\|_p \le y^{\alpha-2} E \quad \text{for all } y > 0,$$

it follows from inequality (6.22) that

$$\|f(\cdot+h) - 2f(\cdot) + f(\cdot-h)\|_p \le 4E \int_0^t y^{\alpha-1} dy + MEt^\alpha$$

$$\le M_\alpha E t^\alpha \quad \text{for all } 0 < |h| = s \le t.$$

This proves that

$$\begin{aligned} G &= \sup_{t>0} \left(t^{-\alpha} \cdot \sup_{0<|h|\le t} \|f(\cdot + h) - 2f(\cdot) + f(\cdot - h)\|_p \right) \\ &\le M_\alpha E, \quad 0 < \alpha < 2. \end{aligned}$$

Step 7-2b: The case $1 \le q < \infty$. By applying Hardy's inequality (Theorem 3.18), we obtain from inequality (6.22) that

$$\begin{aligned} G &= \left(\int_0^\infty \left(t^{-\alpha} \cdot \sup_{0<|h|\le t} \|f(\cdot + h) - 2f(\cdot) + f(\cdot - h)\|_p \right)^q \frac{dt}{t} \right)^{1/q} \\ &\le \left(\int_0^\infty t^{-\alpha q} \left(4 \int_0^t y \|u_{yy}(\cdot, y)\|_p \, dy + 2Mt^2 \|u_{yy}(\cdot, t)\|_p \right)^q \frac{dt}{t} \right)^{1/q} \\ &\le \left[2^{3-3/q} \left(\int_0^\infty \left(t^{-\alpha} \int_0^t y \|u_{yy}(\cdot, y)\|_p \, dy \right)^q \frac{dt}{t} \right)^{1/q} \right. \\ &\quad \left. + 2^{1-1/q} M \left(\int_0^\infty \left(t^{2-\alpha} \|u_{yy}(\cdot, t)\|_p \right)^q \frac{dt}{t} \right)^{1/q} \right] \\ &\le \left(\frac{2^{3-3/q}}{\alpha} + 2^{1-1/q} M \right) \left(\int_0^\infty \left(t^{2-\alpha} \|u_{yy}(\cdot, t)\|_p \right)^q \frac{dt}{t} \right)^{1/q} \\ &\le \left(\frac{8}{\alpha} + 2M \right) E. \end{aligned}$$

This proves that

$$G \le M_\alpha E,$$

with

$$M_\alpha := \frac{8}{\alpha} + 2M, \quad 0 < \alpha < 2.$$

Step 8: $E \le M_\alpha C$. We remark that

$$|P_{yy}(z, y)| \le \frac{My}{(|z|^2 + y^2)^{(n+3)/2}} \quad \text{for all } y > 0,$$

and that

$$\int_{\mathbf{R}^n} P_{yy}(z, y) \, dz = \frac{d^2}{dy^2} \left(\int_{\mathbf{R}^n} P(z, y) \, dz \right) = 0.$$

Hence we have the formula

$$u_{yy}(x,y) = \int_{\mathbf{R}^n} [f(x-z) - f(x)]\, P_{yy}(z,y)\, dz,$$

and so

$$\begin{aligned}
&\|u_{yy}(\cdot,y)\|_p && (6.23)\\
&\le M \int_{\mathbf{R}^n} \|f(\cdot - z) - f(\cdot)\|_p \frac{y}{\left(|z|^2 + y^2\right)^{(n+3)/2}}\, dz \\
&= M \int_0^\infty \int_{\Sigma_{n-1}} \|f(\cdot - t\sigma) - f(\cdot)\|_p \frac{y}{\left(t^2 + y^2\right)^{(n+3)/2}}\, t^{n-1}\, d\sigma\, dt \\
&= M\omega_n \int_0^\infty \|\phi_t(\cdot)\|_p \frac{y}{\left(t^2 + y^2\right)^{(n+3)/2}}\, t^{n-1}\, dt.
\end{aligned}$$

Here

$$\phi_t(x) = \frac{1}{\omega_n} \int_{\Sigma_{n-1}} (f(x+t\sigma) - f(x))\, d\sigma.$$

Step 8-a: The case $q = \infty$. Since we have, by Minkowski's inequality for integrals (Theorem 3.16),

$$\|\phi_t(\cdot)\|_p \le \frac{1}{\omega_n} \int_{\Sigma_{n-1}} \|f(\cdot + t\sigma) - f(\cdot)\|_p \le Ct^\alpha \quad \text{for all } t > 0,$$

it follows from inequality (6.23) that

$$\begin{aligned}
\|u_{yy}(\cdot,y)\|_p &\le MC\,\omega_n \int_0^\infty \frac{y}{\left(t^2+y^2\right)^{(n+3)/2}}\, t^{\alpha+n-1}\, dt \\
&\le MC\,\omega_n \left(\int_0^y \frac{y \cdot y^{\alpha+n-1}}{y^{n+3}}\, dt + \int_y^\infty \frac{y}{t^{n+3}} t^{\alpha+n-1}\, dt \right) \\
&= MC\,\omega_n \left(\int_0^y y^{\alpha-3}\, dt + \int_y^\infty y t^{\alpha-4}\, dt \right) \\
&= \frac{4-\alpha}{3-\alpha}\, M\,\omega_n\, C\, y^{\alpha-2} \quad \text{for all } y > 0.
\end{aligned}$$

This proves that

$$E = \sup_{y>0} \left\{ y^{2-\alpha} \|u_{yy}(\cdot,y)\|_p \right\} \le M_\alpha C,$$

with

$$M_\alpha := \frac{4-\alpha}{3-\alpha}\, M\,\omega_n, \quad 0 < \alpha < 2.$$

Step 8-b: The case $1 \leq q < \infty$. By applying Hardy's inequality (Theorem 3.18), we obtain that

$$\begin{aligned}
E &= \left(\int_0^\infty \left(y^{2-\alpha} \left\| u_{yy}(\cdot, y) \right\|_p \right)^q \frac{dy}{y} \right)^{1/q} \\
&\leq \left(\int_0^\infty \left(y^{\alpha-2} M\, \omega_n \int_0^\infty \left\| \phi_t(\cdot) \right\|_p \frac{y t^{n-1}}{(t^2+y^2)^{(n+3)/2}}\, dt \right)^q \frac{dy}{y} \right)^{1/q} \\
&= M\, \omega_n \left(\int_0^\infty \left(\int_0^\infty \left\| \phi_t(\cdot) \right\|_p \frac{y t^{n-1}}{(t^2+y^2)^{(n+3)/2}}\, dt \right)^q \frac{dy}{y^{1+q(\alpha-2)}} \right)^{1/q} \\
&\leq M\, \omega_n \left(\int_0^\infty \left(y^{-n-\alpha} \int_0^y t^{n-1} \left\| \phi_t(\cdot) \right\|_p \, dt \right)^q \frac{dy}{y} \right)^{1/q} \\
&\quad + M\, \omega_n \left(\int_0^\infty \left(y^{3-\alpha} \int_y^\infty t^{-4} \left\| \phi_t(\cdot) \right\|_p \, dt \right)^q \frac{dy}{y} \right)^{1/q} \\
&\leq \frac{M\omega_n}{n+\alpha} \left(\int_0^\infty \left(y^{-n-\alpha+1} y^{n-1} \left\| \phi_y(\cdot) \right\|_p \right)^q \frac{dy}{y} \right)^{1/q} \\
&\quad + \frac{M\omega_n}{3-\alpha} \left(\int_0^\infty \left(y^{3-\alpha+1} y^{-4} \left\| \phi_y(\cdot) \right\|_p \right)^q \frac{dy}{y} \right)^{1/q} \\
&= M\, \omega_n \left(\frac{1}{n+\alpha} + \frac{1}{3-\alpha} \right) \left(\int_0^\infty \left(y^{-\alpha} \left\| \phi_y(\cdot) \right\|_p \right)^q \frac{dy}{y} \right)^{1/q} \\
&= M\, \omega_n \left(\frac{1}{n+\alpha} + \frac{1}{3-\alpha} \right) C.
\end{aligned}$$

This proves that

$$E \leq M_\alpha C,$$

with

$$M_\alpha := M\omega_n \left(\frac{1}{n+\alpha} + \frac{1}{3-\alpha} \right), \quad 0 < \alpha < 2.$$

Step 9: $H \leq (1/\alpha) D$. If x is a point of the Lebesgue set of $f \in L^p(\mathbf{R}^n)$, it follows that

$$u(x,y) - f(x) = \int_0^y u_y(x,s)\, ds.$$

Hence, by applying Minkowski's inequality for integrals (Theorem 3.16) we obtain that

$$\left\| u(\cdot, y) - f(\cdot) \right\|_p \leq \int_0^y \left\| u_y(\cdot, s) \right\|_p \, ds. \tag{6.24}$$

Step 9-a: The case $q = \infty$. Since we have the inequality

$$s^{1-\alpha} \left\| u_y(\cdot, s) \right\|_p \leq D \quad \text{for all } s > 0,$$

it follows from inequality (6.24) that

$$\|u(\cdot, y) - f(\cdot)\|_p \leq \int_0^y \|u_y(\cdot, s)\|_p \, ds \leq D \int_0^y s^{\alpha-1} \, ds = \frac{D}{\alpha} y^\alpha.$$

This proves that

$$H = \sup_{y>0} \left\{ y^{-\alpha} \|u(\cdot, y) - f(\cdot)\|_p \right\} \leq \frac{1}{\alpha} D.$$

Step 9-b: The case $1 \leq q < \infty$. In view of inequality (6.24), it follows from an application of Hardy's inequality (Theorem 3.18) that

$$\begin{aligned} H &= \left(\int_0^\infty \left(y^{-\alpha} \|u(\cdot, y) - f(\cdot)\|_p \right)^q \frac{dy}{y} \right)^{1/q} \\ &\leq \left(\int_0^\infty \left(y^{-\alpha} \int_0^y \|u_y(\cdot, s)\|_p \, ds \right)^q \frac{dy}{y} \right)^{1/q} \\ &\leq \frac{1}{\alpha} \left(\int_0^\infty \left(y^{1-\alpha} \|u_y(\cdot, y)\|_p \right)^q \frac{dy}{y} \right)^{1/q} \\ &= \frac{1}{\alpha} D, \quad 0 < \alpha < 1. \end{aligned}$$

Step 10: $D \leq M_\alpha H$ $(0 < \alpha < 1)$. We recall that

$$u_y(x, y) = u(x, y/2) * P_y(x, y/2),$$

and that

$$|P_y(x, y/2)| \leq \frac{n+1}{y} P(x, y/2) \quad \text{for all } y > 0.$$

Hence we have, by Young's inequality (Theorem 3.23),

$$\begin{aligned} \|u_y(\cdot, y)\|_p &\leq \|u(\cdot, y/2)\|_p \, \|P_y(\cdot, y/2)\|_1 \\ &\leq \frac{n+1}{y} \|P(\cdot, y/2)\|_1 \, \|u(\cdot, y/2)\|_p \\ &= \frac{n+1}{y} \|u(\cdot, y/2)\|_p = \frac{n+1}{y} \|P(\cdot, y/2) * f(\cdot)\|_p \\ &\leq \frac{n+1}{y} \|f\|_p \quad \text{for all } y > 0. \end{aligned}$$

This proves that, as $N \to \infty$,

$$\left\| u_y(\cdot, 2^N y) \right\|_p \longrightarrow 0 \quad \text{for each } y > 0.$$

Hence we have the assertion

$$\|u_y(\cdot,y)\|_p = \lim_{N\to\infty} \left\|u_y(\cdot,y) - u_y(\cdot,2^N y)\right\|_p, \quad \text{for each } y>0. \tag{6.25}$$

On the other hand, we have the following claim:

Claim 6.2. There exists a positive constant M_α such that we have, for any integer $N>1$,

$$\left\|y^{1-\alpha}(u_y(\cdot,y) - u_y(\cdot,2^N y))\right\|_{pq} \le M_\alpha H, \quad 0<\alpha<1. \tag{6.26}$$

Proof. By Minkowski's inequality (Theorem 3.15), we obtain that

$$\begin{aligned} & y^{1-\alpha}\left\|u_y(\cdot,y) - u_y(\cdot,2^N y)\right\|_p \qquad (6.27)\\ &\le y^{1-\alpha}\sum_{k=1}^{N}\left\|u_y(\cdot,2^{k-1}y) - u_y(\cdot,2^k y)\right\|_p \\ &= y^{1-\alpha}\sum_{k=1}^{N}\left\|\left[f(\cdot) - u(\cdot,2^{k-1}y)\right] * P_y(\cdot,2^{k-1}y)\right\|_p \\ &\le y^{1-\alpha}\sum_{k=1}^{N}\left\|f(\cdot) - u(\cdot,2^{k-1}y)\right\|_p \left\|P_y(\cdot,2^{k-1}y)\right\|_1 \\ &\le y^{1-\alpha}\sum_{k=1}^{N}\frac{n+1}{y}2^{1-k}\left\|f(\cdot) - u(\cdot,2^{k-1}y)\right\|_p \\ &= (n+1)\sum_{k=1}^{N}(2^{\alpha-1})^{k-1}(2^{k-1}y)^{-\alpha}\left\|f(\cdot) - u(\cdot,2^{k-1}y)\right\|_p . \end{aligned}$$

However, we have, for $1\le k\le N$,

$$\begin{aligned} & \left(\int_0^\infty \left((2^{k-1}y)^{-\alpha}\left\|f(\cdot) - u(\cdot,2^{k-1}y)\right\|_p\right)^q \frac{dy}{y}\right)^{1/q} \qquad (6.28)\\ &= \left(\int_0^\infty \left(z^{-\alpha}\left\|f(\cdot) - u(\cdot,z)\right\|_p\right)^q \frac{dz}{z}\right)^{1/q}. \end{aligned}$$

Hence it follows from inequalities (6.27) and (6.28) that

$$\begin{aligned} & \left\|y^{1-\alpha}(u_y(\cdot,y) - u_y(\cdot,2^N y))\right\|_{pq} \\ &= \left(\int_0^\infty \left(y^{1-\alpha}\left\|u_y(\cdot,y) - u(\cdot,2^N y)\right\|_p\right)^q \frac{dy}{y}\right)^{1/q} \\ &\le (n+1) \\ &\quad \times \left(\int_0^\infty \left(\sum_{k=1}^{N}(2^{\alpha-1})^{k-1}(2^{k-1}y)^{-\alpha}\left\|f(\cdot) - u(\cdot,2^{k-1}y)\right\|_p\right)^q \frac{dy}{y}\right)^{1/q} \end{aligned}$$

$$= (n+1)\left(\int_0^\infty \left(\sum_{k=1}^{N}(2^{\alpha-1})^{k-1}\right)^q \left(z^{-\alpha}\left\|f(\cdot)-u(\cdot,z)\right\|_p\right)^q \frac{dz}{z}\right)^{1/q}$$

$$= (n+1)\left(\sum_{k=1}^{N}(2^{\alpha-1})^{k-1}\right)\left\|y^{-\alpha}(f(\cdot)-u(\cdot,y))\right\|_{pq}$$

$$\leq (n+1)\left(\sum_{k=1}^{\infty}\left(\frac{1}{2^{1-\alpha}}\right)^{k-1}\right)H, \quad 0<\alpha<1.$$

This proves the desired inequality (6.26), with

$$M_\alpha := \frac{n+1}{1-2^{\alpha-1}}, \quad 0<\alpha<1.$$

The proof of Claim 6.2 is complete. □

Therefore, by applying Fatou's lemma (Theorem 3.7) we obtain from formula (6.25) and inequality (6.26) that

$$\begin{aligned} D^q &= \left\|y^{1-\alpha}u_y(\cdot,y)\right\|_{pq}^q = \int_0^\infty \left(y^{1-\alpha}\left\|u_y(\cdot,y)\right\|_p\right)^q \frac{dy}{y} \\ &= \int_0^\infty \lim_{N\to\infty}\left(y^{1-\alpha}\left\|u_y(\cdot,y)-u_y(\cdot,2^N y)\right\|_p\right)^q \frac{dy}{y} \\ &\leq \liminf_{N\to\infty}\int_0^\infty \left(y^{1-\alpha}\left\|u_y(\cdot,y)-u_y(\cdot,2^N y)\right\|_p\right)^q \frac{dy}{y} \\ &= \liminf_{N\to\infty}\left\|y^{1-\alpha}(u_y(\cdot,y)-u_y(\cdot,2^N y))\right\|_{pq}^q \\ &\leq (M_\alpha H)^q. \end{aligned}$$

This proves that

$$D \leq M_\alpha H, \quad 0<\alpha<1.$$

The proof of Theorem 6.3 is now complete. □

Theorem 6.5. *Let $\alpha > 0$, $1 \leq p,\ q \leq \infty$ and let k be a positive integer less than α. Then the norm*

$$\|f\|_p + \sum_{|s|=k}\|D_x^s f\|_{\alpha-k;p,q}$$

is equivalent to the norm $\|f\|_{\alpha;p,q}$.

Proof. The proof of Theorem 6.5 is divided into two steps.

Step (I): We assume that

$$\|f\|_{\alpha;p,q} = \|f\|_p + \left\|y^{\overline{\alpha}-\alpha}u_y^{(\overline{\alpha})}(\cdot,y)\right\|_{pq} < \infty,$$

where

$$u(x,y) = P(x,y) * f(x).$$

Then we have the following three assertions (a), (b) and (c):

(a) The function $D_x^s u(x,y)$ is harmonic in the half-space $\mathbf{R}_+^{n+1}$ and bounded in each proper sub-half space of $\mathbf{R}_+^{n+1}$.
(b) $\|D_x^s u(\cdot,y)\|_p = \|D_x^s P(\cdot,y) * f(\cdot)\|_p \leq M_s \|f\|_p$, $y \geq 1$.
(c) It follows from a repeated application of Lemma 5.18 that

$$\begin{aligned}\left\| y^{\overline{\alpha}-(\alpha-k)} D_x^s (u_y^{(\overline{\alpha})}(\cdot,y)) \right\|_{pq} &\leq M_\alpha \left\| y^{\overline{\alpha}-\alpha} u_y^{(\overline{\alpha})}(\cdot,y) \right\|_{pq} \\ &\leq M_\alpha \|f\|_{\alpha;p,q}\,.\end{aligned}$$

Hence, by applying Lemma 6.4 to our situation we obtain that there exists a function $g \in B_{pq}^{\alpha-k}(\mathbf{R}^n)$ such that

$$D_x^s u(x,y) = P(x,y) * g(x), \quad |s| = k,$$

and that

$$\|g\|_{\alpha-k;p,q} \leq M \|f\|_{\alpha;p,q}\,.$$

However, it is easy to see that

$$g = D_x^s f.$$

Indeed, we have the formulas

$$P(x,y) * g(x) = D_x^s u(x,y) = D_x^s \left(P(x,y) * f(x)\right) = P(x,y) * D_x^s f(x).$$

Therefore, we obtain that

$$\begin{aligned}\|D_x^s f\|_{\alpha-k;p,q} &= \|D_x^s f\|_p + \left\| y^{\overline{\alpha}-(\alpha-k)} D_y^{\overline{\alpha}} (D_x^s u(\cdot,y)) \right\|_{pq} \\ &= \|g\|_p + \left\| y^{\overline{\alpha}-(\alpha-k)} D_y^{\overline{\alpha}} (P(\cdot,y) * g(\cdot)) \right\|_{pq} \\ &\leq \|g\|_{\alpha-k;p,q} \\ &\leq M \|f\|_{\alpha;p,q}\,.\end{aligned}$$

We are done with the proof in one direction:

$$\|f\|_p + \sum_{|s|=k} \|D_x^s f\|_{\alpha-k;p,q} \leq (M+1) \|f\|_{\alpha;p,q}\,.$$

Step (II): Conversely, we assume that

$$f \in L^p(\mathbf{R}^n),$$

and that

$$\sum_{|s|=k} \|D_x^s f\|_{\alpha-k;p,q} < \infty.$$

Then we have, for all $|s| = k$,

$$D_x^s u(x,y) = D_x^s \left(P(x,y) * f(x)\right) = P(x,y) * D_x^s f(x).$$

By applying Proposition 6.2 with $f(x) := D_x^s f(x)$, we obtain that

$$\begin{aligned} \left\| y^{k+\overline{\alpha}-\alpha} D_x^s (u_y^{(\overline{\alpha})}(\cdot,y)) \right\|_{pq} &= \left\| y^{\overline{\alpha}-(\alpha-k)} D_y^{\overline{\alpha}} (P(\cdot,y) * D_x^s f(\cdot)) \right\|_{pq} \\ &\leq \|D_x^s f\|_{\alpha-k;p,q} . \end{aligned} \tag{6.29}$$

Moreover, we have, by a repeated application of part (b) of Theorem 5.19,

$$\left\| y^{k+\overline{\alpha}-\alpha} u_y^{(k+\overline{\alpha})}(\cdot,y)) \right\|_{pq} \leq M_k \sum_{|s|=k} \left\| y^{k+\overline{\alpha}-\alpha} D_x^s (u_y^{(\overline{\alpha})}(\cdot,y)) \right\|_{pq} . \tag{6.30}$$

Hence, in view of Proposition 6.2, it follows from inequalities (6.29) and (6.30) that

$$\begin{aligned} \left\| y^{\overline{\alpha}-\alpha} u_y^{(\overline{\alpha})}(\cdot,y)) \right\|_{pq} &\leq M_k' \left\| y^{k+\overline{\alpha}-\alpha} u_y^{(k+\overline{\alpha})}(\cdot,y)) \right\|_{pq} \\ &\leq M_k' M_k \sum_{|s|=k} \left\| y^{k+\overline{\alpha}-\alpha} D_x^s (u_y^{(\overline{\alpha})}(\cdot,y)) \right\|_{pq} \\ &\leq M_k' M_k \sum_{|s|=k} \|D_x^s f\|_{\alpha-k;p,q} . \end{aligned} \tag{6.31}$$

Therefore, we obtain from inequality (6.31) that

$$\begin{aligned} \|f\|_{\alpha;p,q} &= \|f\|_p + \left\| y^{\overline{\alpha}-\alpha} u_y^{(\overline{\alpha})}(\cdot,y)) \right\|_{pq} \\ &\leq (1 + M_k' M_k) \left(\|f\|_p + \sum_{|s|=k} \|D_x^s f\|_{\alpha-k;p,q} \right) . \end{aligned}$$

This completes the proof in the other direction.

The proof of Theorem 6.5 is complete. □

By combining Theorems 6.5 and 6.3, we obtain the following:

Theorem 6.6. *Let $\alpha > 0$, $1 \le p$, $q \le \infty$ and let k be a non-negative integer such that $k < \alpha \le k+1$. Then the norm*

$$\|f\|_p + \sum_{|s|=k} \left(\int_{\mathbf{R}^n} \frac{\|D^s f(\cdot + h) - 2D^s f(\cdot) + D^s f(\cdot - h)\|_p^q}{|h|^{n+(\alpha-k)q}} \, dh \right)^{1/q}$$

is equivalent to the norm $\|f\|_{\alpha;p,q}$ defined by formula (6.1)*:*

$$\|f\|_{\alpha;p,q} = \|f\|_p + \left(\int_0^\infty \left(y^{k+1-\alpha} \left\| u_y^{(k+1)}(\cdot, y) \right\|_p \right)^q \frac{dy}{y} \right)^{1/q}.$$

6.3 Notes and Comments

The results discussed in this chapter are adapted from Taibleson [72] and Stein [68].

7
Sobolev and Besov Spaces

This chapter is devoted to the precise definitions and statements of function spaces of L^p type with some detailed proofs. The function spaces we shall treat are the following:

(i) The generalized Sobolev spaces $W^{s,p}(\Omega)$ and $H^{s,p}(\Omega)$. When Ω is a Lipschitz domain, these spaces coincide with each other.
(ii) The Besov spaces $B^{s,p}(\partial\Omega)$ on the boundary $\partial\Omega$ of a Lipschitz domain Ω are function spaces defined in terms of the L^p modulus of continuity, and enter naturally in connection with boundary value problems in the framework of L^p Sobolev spaces.

In fact, we need to make sense of the restriction $u|_{\partial\Omega}$ to the boundary $\partial\Omega$ as an element of a Besov space on $\partial\Omega$ when u belongs to a Sobolev space on the domain Ω. In particular, we formulate an important trace theorem (Theorem 7.5) that will be used in the study of boundary value problems in Parts III and IV.

In the last Section 7.4 we prove part (iv) of Examples 4.1 (Proposition 7.7):

$$W^{\theta,n/\theta}(\mathbf{R}^n) \subset \mathrm{VMO} \quad \text{for } 0 < \theta \leq 1.$$

7.1 Sobolev Spaces

In this section we present a brief description of the basic concepts and results of L^p Sobolev spaces which will be used in subsequent chapters. Many problems in partial differential equations may be formulated in terms of abstract operators acting between suitable Sobolev spaces, and these operators are then analyzed by the methods of functional analysis.

7.1.1 First Definition of Sobolev Spaces

Let Ω be an open subset of $\mathbf{R}^n$. If $1 < p < \infty$ and if s is a non-negative integer, then the Sobolev space $W^{s,p}(\Omega)$ is defined to be the space of those functions $u \in L^p(\Omega)$ such that $D^\alpha u \in L^p(\Omega)$ for $|\alpha| \le s$, and the norm $\|u\|_{W^{s,p}(\Omega)}$ is defined by the formula

$$\|u\|_{W^{s,p}(\Omega)} = \left(\sum_{|\alpha| \le s} \int_\Omega |D^\alpha u(x)|^p dx \right)^{1/p}. \tag{7.1}$$

If $1 < p < \infty$ and if $s = m + \theta$ with a non-negative integer m and $0 < \theta < 1$, then the Sobolev space $W^{s,p}(\Omega)$ is defined to be the space of those functions $u \in W^{m,p}(\Omega)$ such that, for $|\alpha| = m$, the integral (Slobodeckiĭ seminorm)

$$\iint_{\Omega \times \Omega} \frac{|D^\alpha u(x) - D^\alpha u(y)|^p}{|x-y|^{n+p\theta}} \, dx \, dy \tag{7.2}$$

is finite. The norm $\|u\|_{W^{s,p}(\Omega)}$ of $W^{s,p}(\Omega)$ is defined by the formula

$$\begin{aligned} \|u\|_{W^{s,p}(\Omega)} = \Bigg(& \sum_{|\alpha| \le m} \int_\Omega |D^\alpha u(x)|^p \, dx \\ & + \sum_{|\alpha| = m} \iint_{\Omega \times \Omega} \frac{|D^\alpha u(x) - D^\alpha u(y)|^p}{|x-y|^{n+p\theta}} \, dx \, dy \Bigg)^{1/p}. \end{aligned} \tag{7.3}$$

7.1.2 Second Definition of Sobolev Spaces

Next, we introduce a second family of Sobolev spaces, by using the Fourier transform.

Let $\mathcal{S}(\mathbf{R}^n)$ be the *Schwartz space* or space of C^∞ functions on $\mathbf{R}^n$ rapidly decreasing at infinity. We recall that the (direct) Fourier transform $\mathcal{F}$ and the inverse Fourier transform $\mathcal{F}^*$ are isomorphisms of $\mathcal{S}(\mathbf{R}^n)$ onto itself. The dual space $\mathcal{S}'(\mathbf{R}^n)$ of $\mathcal{S}(\mathbf{R}^n)$ are called the space of *tempered distributions* on $\mathbf{R}^n$. Roughly speaking, the tempered distributions are those distributions which grow at most polynomially at infinity, since the functions in $\mathcal{S}(\mathbf{R}^n)$ die out faster than any power of x at infinity. The importance of tempered distributions lies in the fact that they have Fourier transforms. More precisely, if $u \in \mathcal{S}'(\mathbf{R}^n)$, we define its (direct) Fourier transform $\mathcal{F}u$ by the formula

$$\langle \mathcal{F}u, \varphi \rangle = \langle u, \mathcal{F}\varphi \rangle \quad \text{for all } \varphi \in \mathcal{S}(\mathbf{R}^n),$$

where $\langle \cdot, \cdot \rangle$ is the pairing of $\mathcal{S}'(\mathbf{R}^n)$ and $\mathcal{S}(\mathbf{R}^n)$. Similarly, if $v \in \mathcal{S}'(\mathbf{R}^n)$, we define its inverse Fourier transform $\mathcal{F}^* v$ by the formula

$$\langle \mathcal{F}^* v, \psi \rangle = \langle v, \mathcal{F}^* \psi \rangle \quad \text{for all } \psi \in \mathcal{S}(\mathbf{R}^n).$$

It should be emphasized that the Fourier transforms $\mathcal{F}$ and $\mathcal{F}^*$ are isomorphisms of $\mathcal{S}'(\mathbf{R}^n)$ onto itself.

If $s \in \mathbf{R}$, we define a linear map

$$\mathcal{G}^s = (I - \Delta)^{-s/2} \colon \mathcal{S}'(\mathbf{R}^n) \longrightarrow \mathcal{S}'(\mathbf{R}^n)$$

by the formula

$$\mathcal{G}^s u = \mathcal{F}^* \left((1 + |\xi|^2)^{-s/2} \mathcal{F} u \right) \quad \text{for } u \in \mathcal{S}'(\mathbf{R}^n). \tag{7.4}$$

The operator $\mathcal{G}^s \colon \mathcal{S}'(\mathbf{R}^n) \to \mathcal{S}'(\mathbf{R}^n)$ can be visualized as follows:

$$\begin{array}{ccc} u \in \mathcal{S}'(\mathbf{R}^n) & \xrightarrow{\mathcal{G}^s} & \mathcal{S}'(\mathbf{R}^n) \ni \mathcal{G}^s u \\ {\scriptstyle \mathcal{F}}\big\downarrow & & \big\uparrow{\scriptstyle \mathcal{F}^*} \\ \mathcal{F} u \in \mathcal{S}'(\mathbf{R}^n) & \xrightarrow[(1+|\xi|^2)^{-s/2}]{} & \mathcal{S}'(\mathbf{R}^n) \ni (1 + |\xi|^2)^{-s/2} \mathcal{F} u \end{array}$$

Fig. 7.1. The operator $\mathcal{G}^s$ defined by formula (7.4)

Then it is easy to see that the map $\mathcal{G}^s$ is an isomorphism of $\mathcal{S}'(\mathbf{R}^n)$ onto itself, and its inverse is the map $\mathcal{G}^{-s}$. The function $\mathcal{G}^s u$ is called the *Bessel potential* of order s of u.

We can calculate the convolution kernel $G_s(x)$ of the Bessel potential $\mathcal{G}^s u$ for all $s > 0$. More precisely, we have the following:

Theorem 7.1. *Let $s > 0$. (i) The inverse Fourier transform*

$$\mathcal{F}^* \left((1 + |\xi|^2)^{-s/2} \right)$$

is equal to the function

$$G_s(x) = \frac{1}{(4\pi)^{s/2}} \frac{1}{\Gamma(s/2)} \int_0^\infty e^{-\pi |x|^2/\delta} \, e^{-\delta/(4\pi)} \, \delta^{(s-n)/2} \, \frac{d\delta}{\delta}. \tag{7.5}$$

In other words, we have, by the Fourier inversion formula,

$$\mathcal{F}(G_s)(\xi) = \left(1 + |\xi|^2\right)^{-s/2}. \tag{7.6}$$

Moreover, we have the formula

$$\int_{\mathbf{R}^n} G_s(x)\,dx = 1,$$

and so

$$G_s \in L^1(\mathbf{R}^n).$$

(ii) Let $1 \leq p \leq \infty$. The Bessel potential $\mathcal{G}^s$ can be expressed as follows:

$$\mathcal{G}^s u(x) = G_s * u(x) = \int_{\mathbf{R}^n} G_s(x-y)u(y)\,dy \quad \textit{for } u \in L^p(\mathbf{R}^n). \tag{7.7}$$

Furthermore, the Bessel potential $\mathcal{G}^s$ is bounded from $L^p(\mathbf{R}^n)$ into itself. More precisely, we have the inequality

$$\|\mathcal{G}^s u\|_{L^p(\mathbf{R}^n)} \leq \|u\|_{L^p(\mathbf{R}^n)} \quad \textit{for } u \in L^p(\mathbf{R}^n). \tag{7.8}$$

One of the most important facts concerning Bessel potentials is that they can be used to define the generalized Sobolev spaces $H^{s,p}(\mathbf{R}^n)$ in the following way: If $s \in \mathbf{R}$ and $1 < p < \infty$, we let

$$\begin{aligned} H^{s,p}(\mathbf{R}^n) &= \text{the image of } L^p(\mathbf{R}^n) \text{ under the mapping } \mathcal{G}^s \\ &= \{\mathcal{G}^s v : v \in L^p(\mathbf{R}^n)\}. \end{aligned}$$

We equip $H^{s,p}(\mathbf{R}^n)$ with the norm

$$\|u\|_{H^{s,p}(\mathbf{R}^n)} = \left\|\mathcal{G}^{-s} u\right\|_{L^p(\mathbf{R}^n)} \quad \text{for } u \in H^{s,p}(\mathbf{R}^n). \tag{7.9}$$

The space $H^{s,p}(\mathbf{R}^n)$ is called the Bessel-potential space of order s or the generalized *Sobolev space* of order s.

We list some basic topological properties of $H^{s,p}(\mathbf{R}^n)$:

(1) The Schwartz space $\mathcal{S}(\mathbf{R}^n)$ is dense in each $H^{s,p}(\mathbf{R}^n)$.
(2) The space $H^{-s,p'}(\mathbf{R}^n)$ is the dual space of $H^{s,p}(\mathbf{R}^n)$, where $p' = p/(p-1)$ is the exponent conjugate to p.
(3) If $s > t$, then we have the inclusions

$$\mathcal{S}(\mathbf{R}^n) \subset H^{s,p}(\mathbf{R}^n) \subset H^{t,p}(\mathbf{R}^n) \subset \mathcal{S}'(\mathbf{R}^n),$$

with continuous injections.
(4) If s is a non-negative integer, then the space $H^{s,p}(\mathbf{R}^n)$ is isomorphic to the Sobolev space $W^{s,p}(\mathbf{R}^n)$, and the norm (7.9) is equivalent to the norm (7.1).

7.1.3 Definition of General Sobolev Spaces

Now we define the generalized Sobolev spaces $H^{s,p}(\Omega)$ for general domains Ω.

For each $s \in \mathbf{R}$ and $1 < p < \infty$, we let

$$H^{s,p}(\Omega) = \text{ the space of restrictions to } \Omega \text{ of functions in } H^{s,p}(\mathbf{R}^n).$$

We equip the space $H^{s,p}(\Omega)$ with the norm

$$\|u\|_{H^{s,p}(\Omega)} = \inf \|U\|_{H^{s,p}(\mathbf{R}^n)},$$

where the infimum is taken over all $U \in H^{s,p}(\mathbf{R}^n)$ which equal u in Ω. The space $H^{s,p}(\Omega)$ is a Banach space with respect to the norm $\|\cdot\|_{s,p}$. It should be noticed that

$$H^{0,p}(\Omega) = L^p(\Omega); \quad \|\cdot\|_{H^{0,p}(\Omega)} = \|\cdot\|_{L^p(\Omega)}.$$

Then we have the following relationships between the spaces $H^{s,p}(\Omega)$ and $W^{s,p}(\Omega)$ (see [2, Theorem 5.24]):

Theorem 7.2. *If Ω is a bounded, Lipschitz domain, then we have, for all $s \geq 0$ and $1 < p < \infty$,*

$$H^{s,p}(\Omega) = W^{s,p}(\Omega).$$

7.1.4 Sobolev Imbedding Theorems

In Chapter 13, we shall need Sobolev's imbedding theorems (see [2, Theorem 4.12, Part I and Part II]; [80, Chapter 4]).

Definition 7.1. An open subset Ω of $\mathbf{R}^n$ satisfies the *cone condition* if there exists a finite cone C such that each point $x \in \Omega$ is the vertex of a finite cone C_x contained in Ω and congruent to C.

We remark that the cone C_x need not be obtained from C by parallel translation, but simply by rigid motion.

Example 7.1. Any Lipschitz domain satisfies the cone condition (see [2, Paragraph 4.11]).

Theorem 7.3 (Sobolev)**.** *Let $j \geq 0$ and $m \geq 1$ be integers, and let $1 \leq p < \infty$.*

***Part I**: Assume that Ω is an open subset of $\mathbf{R}^n$ which satisfies the* cone condition*. Then we have the imbeddings*

$$W^{m+j,p}(\Omega)$$

$$\subset \begin{cases} W^{j,q}(\Omega) & \text{for all } p \le q \le np/(n-mp) \text{ if } 1 < p < n/m, \\ W^{j,q}(\Omega) & \text{for all } p \le q < \infty \text{ if } p = n/m, \\ C^j_B(\Omega) & \text{if } n/m < p < \infty. \end{cases}$$

Here

$$C^j_B(\Omega) = \text{the space of those functions } u \in C^j(\Omega) \text{ for which } D^\alpha u \text{ is bounded for } 0 \le |\alpha| \le j,$$

with norm given by the formula

$$\|u\|_{C^j_B(\Omega)} = \max_{0 \le |\alpha| \le j} \sup_{x \in \Omega} |D^\alpha u(x)| .$$

Part II: *Assume that* Ω *is a bounded,* Lipschitz domain *in* $\mathbf{R}^n$*. Then we have the imbeddings*

$$W^{m+j,p}(\Omega)$$
$$\subset \begin{cases} C^{j+\lambda}(\overline{\Omega}) & \text{for all } 0 < \lambda \le m - n/p \text{ if } (m-1)p < n < mp, \\ C^{j+\lambda}(\overline{\Omega}) & \text{for all } 0 < \lambda < 1 \text{ if } n = (m-1)p, \\ C^{j+1}(\overline{\Omega}) & \text{if } n = (m-1) \text{ and } p = 1. \end{cases}$$

7.1.5 The Rellich–Kondrachov Theorem

In Chapter 15 we shall need the following Rellich–Kondrachov theorem (see [2, Theorem 6.3, Parts I and II], [33, Section 7.12, Theorem 7.26]):

Theorem 7.4 (Rellich–Kondrachov)**.** *Let* Ω *be a bounded, open subset of* $\mathbf{R}^n$ *that satisfies the* cone condition*. Then the imbedding*

$$W^{j+1,p}(\Omega) \longrightarrow W^{j,q}(\Omega)$$

is compact *for any integer* $j \ge 0$ *if either* $n > p$ *and* $1 \le q < np/(n-p)$ *or* $1 \le n \le p$ *and* $1 \le q < \infty$*.*

The Rellich–Kondrachov theorem is an L^p Sobolev space version of the Bolzano–Weierstrass theorem and the Ascoli–Arzelà theorem in calculus:

7.2 Besov Spaces on the Boundary

In studying boundary value problems, we shall need to make sense of the restriction $u|_{\partial\Omega}$ as an element of a function space on the boundary $\partial\Omega$ when u belongs to an L^p Sobolev space on the Lipschitz domain Ω.

Subjects	**Sequences**	**Compactness theorems**
Theory of real numbers	Sequences of real numbers	The Bolzano–Weierstrass theorem
Calculus	Sequences of continuous functions	The Ascoli–Arzerà theorem
Theory of distributions	Sequences of distributions	The Rellich–Kondrachov theorem

Table 7.1. *A bird's-eye view of three compactness theorems in calculus*

In this way, the Besov spaces $B^{s,p}(\partial\Omega)$ on the boundary $\partial\Omega$ enter naturally in connection with boundary value problems. The Besov spaces $B^{s,p}(\partial\Omega)$ are defined to be locally the Besov spaces $B^{s,p}(\mathbf{R}^{n-1})$ on $\mathbf{R}^{n-1}$, upon using local coordinate systems flattening out $\partial\Omega$, together with a partition of unity.

An open set Ω in $\mathbf{R}^n$ is called a *Lipschitz hypograph* if its boundary $\partial\Omega$ can be represented as the graph of a Lipschitz continuous function. Namely, there exists a Lipschitz continuous function $\zeta : \mathbf{R}^{n-1} \to \mathbf{R}$ such that (see Figure 5.1)

$$\Omega = \left\{x = (x', x_n) \in \mathbf{R}^n : x_n < \zeta(x'), \quad x' \in \mathbf{R}^{n-1}\right\}. \tag{5.1}$$

We define Besov spaces $B^{s,p}(\partial\Omega)$ on the boundary $\partial\Omega$ of a Lipschitz domain Ω, upon using local coordinate systems flattening out $\partial\Omega$, together with a partition of unity, in the following way.

Step 1: If Ω is a Lipschitz hypograph defined by formula (5.1), then we remark that its boundary

$$\partial\Omega = \left\{x = (x', \zeta(x')) : x' \in \mathbf{R}^{n-1}\right\}$$

is an $(n-1)$-dimensional, $C^{0,1}$ submanifold of $\mathbf{R}^n$. Hence we find that $\partial\Omega$ has a surface measure $d\sigma$ and a unit outward normal $\boldsymbol{\nu}$ which exists $d\sigma$-almost everywhere in $\mathbf{R}^{n-1}$. Indeed, by Rademacher's theorem (Theorem 5.1) it follows that the function $\zeta(x')$ is Fréchet differentiable almost everywhere in $\mathbf{R}^{n-1}$ with

$$\|\nabla\zeta\|_{L^\infty(\mathbf{R}^{n-1})} \leq C,$$

where C is any Lipschitz constant for the function $\zeta(x')$. Then we have the following formulas for $d\sigma$ and $\boldsymbol{\nu}$

$$d\sigma = \sqrt{1+|\nabla\zeta(x')|^2}\, dx',$$
$$\boldsymbol{\nu} = \frac{(-\nabla\zeta(x'), 1)}{\sqrt{1+|\nabla\zeta(x')|^2}}.$$

Step 1-1: Now we can define the Besov spaces $B^{s,p}(\partial\Omega)$ for $0 < s \leq 1$ in the following way: For any function $\varphi \in L^p(\partial\Omega) = L^p(\partial\Omega, d\sigma)$, we define a function

$$\varphi_\zeta(x') := \varphi(x', \zeta(x')), \quad x' \in \mathbf{R}^{n-1},$$

and let, for $0 < s < 1$,

$$B^{s,p}(\partial\Omega) = \left\{\varphi \in L^p(\partial\Omega) : \varphi_\zeta \in B^{s,p}(\mathbf{R}^{n-1})\right\}.$$

We equip this space with the norm (see the norm A in Theorem 6.3 with $n := n-1$, $\alpha := s$, $p = q$)

$$|\varphi|_{B^{s,p}(\partial\Omega)} = |\varphi_\zeta|_{B^{s,p}(\mathbf{R}^{n-1})} \tag{7.10}$$
$$= \left(\int_{\mathbf{R}^{n-1}} |\varphi_\zeta(x')|^p\, dx' + \iint_{\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}} \frac{|\varphi_\zeta(x') - \varphi_\zeta(y')|^p}{|x'-y'|^{(n-1)+ps}} dx'\, dy'\right)^{1/p}.$$

For $s = 1$, we let

$B^{1,p}(\partial\Omega) =$ the space of (equivalence classes of) functions $\varphi \in L^p(\partial\Omega, d\sigma)$ for which the integral

$$\iint_{\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}} \frac{|\varphi_\zeta(x'+h') - 2\varphi_\zeta(x') + \varphi_\zeta(x'-h')|^p}{|h'|^{(n-1)+p}} dh'\, dx'.$$

is finite.

The space $B^{1,p}(\partial\Omega)$ is a Banach space with respect to the norm (see the norm B in Theorem 6.3 with $n := n-1$, $\alpha := 1$, $p = q$)

$$|\varphi|_{B^{1,p}(\partial\Omega)} = |\varphi_\zeta|_{B^{1,p}(\mathbf{R}^{n-1})} \tag{7.11}$$
$$= \left(\int_{\mathbf{R}^{n-1}} |\varphi_\zeta(x')|^p\, dx' \right.$$
$$\left. + \iint_{\mathbf{R}^{n-1}\times\mathbf{R}^{n-1}} \frac{|\varphi_\zeta(x'+h') - 2\varphi_\zeta(x') + \varphi_\zeta(x'-h')|^p}{|h'|^{(n-1)+p}}\, dh'\, dx'\right)^{1/p}.$$

Step 1-2: If $\kappa(\Omega)$ is a Lipschitz hypograph for some rigid motion $\kappa : \mathbf{R}^n \to \mathbf{R}^n$, then we can define the Besov spaces $B^{s,p}(\partial\Omega)$ for $0 < s \leq 1$ in the same way except that

$$\varphi_\zeta(x') := \varphi\left(\kappa^{-1}(x', \zeta(x'))\right) \quad \text{for } x' \in \mathbf{R}^{n-1}.$$

Step 2: We consider the general case where Ω is a bounded Lipschitz domain. By using the notation of Definition 5.1, we choose a partition of unity $\{\phi_j\}_{j=1}^J$ subordinate to the open covering $\{U_j\}_{j=1}^J$ of $\partial\Omega$ (see Figure 5.4). Then we define the Besov spaces $B^{s,p}(\partial\Omega)$ for $0 < s \leq 1$ as follows:

$$B^{s,p}(\partial\Omega) = \{\varphi \in L^p(\partial\Omega) : \phi_j\varphi \in B^{s,p}(\partial\Omega_j),\ 1 \leq j \leq J\},$$

where the norm $|\varphi|_{B^{s,p}(\partial\Omega)}$ is defined by the formula

$$|\varphi_{B^{s,p}(\partial\Omega)} = \sum_{j=1}^{J} |\phi_j\varphi|_{B^{s,p}(\partial\Omega_j)}.$$

It should be emphasized that the Besov spaces $B^{s,p}(\partial\Omega)$ are independent of the open covering $\{U_j\}$ and the partition of unity $\{\phi_j\}$ used.

Step 3: Furthermore, we shall require Besov spaces $B^{s,p}(\partial\Omega)$ for $1 < s < 2$ defined on a bounded $C^{1,1}$ domain Ω.

Step 3-1: If Ω is a $C^{1,1}$ hypograph defined by formula (5.1) for some function $\zeta \in C^{1,1}(\mathbf{R}^{n-1})$, then we define the Besov spaces $B^{s,p}(\partial\Omega)$ for $s = 1 + \theta$ with $0 < \theta < 1$ in the same way by replacing the norm (7.10) by the norm

$$\begin{aligned} &|\varphi|_{B^{s,p}(\partial\Omega)} = |\varphi_\zeta|_{B^{s,p}(\mathbf{R}^{n-1})} \\ &= \Bigg(\sum_{|\alpha| \leq 1} \int_{\mathbf{R}^{n-1}} |D^\alpha \varphi_\zeta(x')|^p \, dx' \\ &\quad + \sum_{|\alpha|=1} \iint_{\mathbf{R}^{n-1} \times \mathbf{R}^{n-1}} \frac{|D^\alpha \varphi_\zeta(x') - D^\alpha \varphi_\zeta(y')|^p}{|x' - y'|^{(n-1)+p\theta}} \, dx' \, dy' \Bigg)^{1/p}. \end{aligned}$$

Step 3-2: If Ω is a bounded $C^{1,1}$ domain, then the Besov spaces $B^{s,p}(\partial\Omega)$ for $1 < s < 2$ are defined to be locally the Besov spaces $B^{s,p}(\partial\Omega_j)$, $1 \leq j \leq J$, just as in Step 3. Here it should be emphasized that the boundary $\partial\Omega$ is an $(n-1)$-dimensional, $C^{1,1}$ submanifold of $\mathbf{R}^n$.

The norm of $B^{s,p}(\partial\Omega)$ for $0 \leq s < 2$ will be denoted by $|\cdot|_{s,p}$.

7.3 Trace Theorems

In this section we prove an important trace theorem which will be used in the study of boundary value problems in the framework of L^p Sobolev spaces (cf. [2], [7], [67], [72], [92]):

Theorem 7.5. *Let $1 < p < \infty$. For every function $f \in H^{s,p}(\mathbf{R}^n)$ with $s > 1/p$, the restriction*

$$g := Rf = f|_{\mathbf{R}^{n-1}}$$

is well defined almost everywhere in $\mathbf{R}^{n-1}$, and $g \in B^{s-1/p,p}(\mathbf{R}^{n-1})$. Furthermore, the restriction mapping R so defined is continuous, that is, there exists a positive constant C such that

$$\|Rf\|_{B^{s-1/p,p}(\mathbf{R}^{n-1})} \leq C \, \|f\|_{H^{s,p}(\mathbf{R}^n)} \quad \text{for all } f \in H^{s,p}(\mathbf{R}^n).$$

An elementary proof of Theorem 7.5 is given in [79, Theorem 6.6].

Under certain hypotheses on the domain Ω, functions in Sobolev spaces $H^{s,p}(\Omega)$ may be extended as functions in $H^{s,p}(\mathbf{R}^n)$. In this way, the trace theorem remains valid for $H^{s,p}(\Omega)$ and $B^{s-1/p,p}(\partial\Omega)$. More precisely, we have the following (see [2, Remarks 7.45]):

Theorem 7.6 (the trace theorem)**.** *Let Ω be a bounded, $C^{1,1}$ domain of $\mathbf{R}^n$. If $1 < p < \infty$, then the trace map*

$$\begin{aligned} \gamma = (\gamma_0, \gamma_1) : W^{2,p}(\Omega) &\longrightarrow B^{2-1/p,p}(\partial\Omega) \oplus B^{1-1/p,p}(\partial\Omega) \\ u &\longmapsto \left(u|_{\partial\Omega}, \left.\frac{\partial u}{\partial \boldsymbol{\nu}}\right|_{\partial\Omega} \right) \end{aligned}$$

is continuous and surjective. Here $\boldsymbol{\nu} = -\mathbf{n}$ is the unit outward normal to the boundary $\partial\Omega$ (see Figure 7.2).

Indeed, it suffices to note that we have, by Theorem 7.2,

$$H^{2,p}(\Omega) = W^{2,p}(\Omega).$$

7.4 VMO Functions Revisited

In this last section we prove part (iv) of Examples 4.1. Namely, we prove the following:

Proposition 7.7. *We have, for $0 < \theta \leq 1$,*

$$W^{\theta,n/\theta}(\mathbf{R}^n) \subset \mathrm{VMO}\,.$$

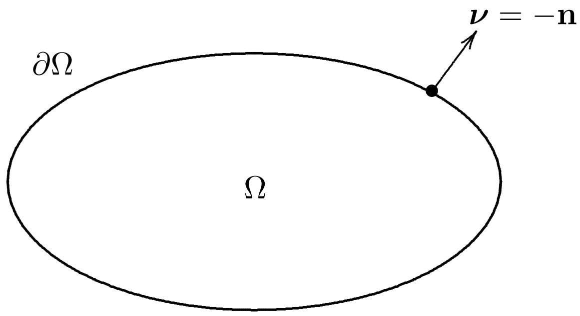

Fig. 7.2. The unit outward normal $\boldsymbol{\nu} = -\mathbf{n}$ to the boundary $\partial\Omega$

Proof. The proof of Proposition 7.7 is divided into two steps.

Step (1): First, we consider the case $\theta = 1$. The proof is based on the following *Poincaré inequality* (see [33, Chapter 7, formula (7.45)], [45, Chapter 1, Theorem 1.51]):

Lemma 7.8 (Poincaré)**.** *Let B be an open ball with diameter d. Let $1 \leq p < \infty$. Then we have, for all $u \in W^{1,p}(\mathbf{R}^n)$,*

$$\|u - u_B\|_p \leq 2^{n-1}\, d\, \|\nabla u\|_p\,, \tag{7.12}$$

where

$$u_B = \frac{1}{|B|} \int_B u(x)\, dx$$

is the integral average of u over B.

By using Poincaré's inequality (7.12) with $p := n$, we obtain that

$$\int_B |u(x) - u_B|^n \, dx \leq \left(2^{n-1} d\right)^n \int_B |\nabla u(x)|^n \, dx.$$

Hence it follows from an application of Hölder's inequality $p := n$ and $q := n/(n-1)$ (Theorem 3.14) that

$$\begin{aligned}
&\int_B |u(x) - u_B| \, dx \\
&\leq \left(\int_B |u(x) - u_B|^n \, dx\right)^{1/n} \left(\int_B dx\right)^{1-1/n} \\
&= \left(\int_B |u(x) - u_B|^n \, dx\right)^{1/n} |B|^{1-1/n}
\end{aligned}$$

$$\leq 2^{n-1}\, d \cdot \left(\int_B |\nabla u(x)|^n \, dx \right)^{1/n} \left(\frac{2^n n \Gamma(n/2)}{2 d^n \pi^{n/2}} \right)^{1/n} |B| \, ,$$

so that

$$\frac{1}{|B|} \int_B |u(x) - u_B| \, dx \leq \frac{2^n}{\sqrt{\pi}} \left(\frac{n}{2} \Gamma\left(\frac{n}{2}\right) \right)^{1/n} \left(\int_B |\nabla u(x)|^n \, dx \right)^{1/n} .$$

However, we find that the integral

$$\int_B |\nabla u(x)|^n \, dx$$

is *absolutely continuous*.

Therefore, we have proved that $u \in \mathrm{VMO}$ if $u \in W^{1,n}(\mathbf{R}^n)$.

Step (2): Secondly, we consider the case where $0 < \theta < 1$. By applying Hölder's inequality for $p := n/\theta$ and $q := n/(n-\theta)$, we obtain that

$$\begin{aligned} \int_B |u(x) - u_B| \, dx &\leq \left(\int_B |u(x) - u_B|^{n/\theta} \, dx \right)^{\theta/n} \left(\int_B dx \right)^{1-\theta/n} \\ &= \left(\int_B |u(x) - u_B|^{n/\epsilon} \, dx \right)^{\theta/n} |B|^{1-\theta/n} , \end{aligned}$$

so that

$$\frac{1}{|B|} \int_B |u(x) - u_B| \, dx \leq \left(\frac{1}{|B|} \int_B |u(x) - u_B|^{n/\epsilon} \, dx \right)^{\theta/n} .$$

Moreover, we have, by Fubini's theorem (Theorem 3.10),

$$\begin{aligned} &\frac{1}{|B|} \int_B |u(x) - u_B| \, dx \leq \left(\frac{1}{|B|} \int_B |u(x) - u_B|^{n/\epsilon} \, dx \right)^{\theta/n} \qquad (7.13) \\ &= \left(\frac{1}{|B|} \left(\int_B \frac{1}{|B|} \left| \int_B (u(x) - u(y)) \, dy \right| \right)^{n/\theta} dx \right)^{\theta/n} \\ &\leq \left(\frac{1}{|B|^{1+n/\theta}} \int_B \left[\int_B |u(x) - u(y)| \, dy \right]^{n/\theta} dx \right)^{\theta/n} \\ &\leq \left(\frac{1}{|B|^{1+n/\theta}} \int_B \left[\left(\int_B |u(x) - u(y)|^{n/\theta} \, dy \right)^{\theta/n} \right. \right. \\ &\qquad \left. \left. \times \left(\int_B dy \right)^{1-\theta/n} \right]^{n/\theta} dx \right)^{\theta/n} \\ &= \left(\frac{1}{|B|^2} \int_B \int_B |u(x) - u(y)|^{n/\theta} \, dy \, dx \right)^{\theta/n} . \end{aligned}$$

However, it is easy to see that we have, for the open ball B with diameter d,

$$|x-y|^{2n} \leq c(n)\,|B|^2 \quad \text{for all } x,\, y \in B,$$

where

$$c(n) = \frac{4^{n-1} n^2 \Gamma\left(\frac{n}{2}\right)^2}{\pi^n}.$$

Hence, by combining this inequality with inequality (7.13) we obtain that

$$\begin{aligned}\frac{1}{|B|}\int_B |u(x)-u_B|\,dx &\leq \left(\frac{1}{|B|^2}\int_B\int_B |u(x)-u(y)|^{n/\theta}\,dy\,dx\right)^{\theta/n} \\ &\leq c(n)\left(\int\int_{B\times B}\frac{|u(x)-u(y)|^{n/\theta}}{|x-y|^{2n}}dx\,dy\right)^{\theta/n}.\end{aligned}$$

However, by formula (7.2) with $p := n/\theta$ we find that the integral

$$\int\int_{B\times B}\frac{|u(x)-u(y)|^{n/\theta}}{|x-y|^{2n}}dx\,dy$$

is *absolutely continuous*.

Therefore, we have proved that $u \in \mathrm{VMO}$ if $u \in W^{\theta,n/\theta}(\mathbf{R}^n)$ for $0 < \theta < 1$.

The proof of Proposition 7.7 is complete. □

7.5 Notes and Comments

For more thorough treatments of the subject in this chapter, the reader might be referred to Adams–Fournier [2], Aronszajn–Smith [5], Bergh–Löfström [7], Stein [68], Taibleson [72] and Triebel [92].

8

Maximum Principles in Sobolev spaces

In this chapter we prove various maximum principles for second-order, elliptic differential operators with discontinuous coefficients such as the weak and strong maximum principles (Theorems 8.5 and 8.9) and Hopf's boundary point lemma (Lemma 8.8) in the framework of L^p Sobolev spaces that will play an important role in the proof of uniqueness theorems for the Dirichlet problem in Part IV.

Let Ω be a bounded domain in Euclidean space $\mathbf{R}^n$, $n \geq 2$, with boundary $\partial\Omega$ of class $C^{1,1}$. We consider a second-order, elliptic Waldenfels integro-differential operator W with real *discontinuous* coefficients of the form

$$Wu(x) = Au(x) + Su(x) \quad \text{for } x \in \Omega, \tag{8.1}$$

where

$$Au(x) = \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{n} b_i(x) \frac{\partial u}{\partial x_i}(x) + c(x)u(x)$$

and

$$Su(x) = \int_{\mathbf{R}^n \setminus \{0\}} \left(u(x+z) - u(x) - \sum_{j=1}^{n} z_j \frac{\partial u}{\partial x_j}(x) \right) K(x,z)\, \mu(dz).$$

More precisely, we assume that the coefficients $a^{ij}(x)$, $b^i(x)$ and $c(x)$ of the differential operator A satisfy the following three conditions (1), (2) and (3):

(1) $a^{ij}(x) \in L^\infty(\Omega)$, $a^{ij}(x) = a^{ji}(x)$ for all $1 \leq i,j \leq n$ and for almost all $x \in \Omega$ and there exist a constant $\lambda > 0$ such that

$$\frac{1}{\lambda}\, |\xi|^2 \leq \sum_{i,j=1}^{n} a^{ij}(x) \xi_i \xi_j \leq \lambda\, |\xi|^2 \tag{8.2}$$

for almost all $x \in \Omega$ and all $\xi \in \mathbf{R}^n$.

(2) $b^i(x) \in L^\infty(\Omega)$ for all $1 \leq i \leq n$.
(3) $c(x) \in L^\infty(\Omega)$ and $c(x) \leq 0$ for almost all $x \in \Omega$.

Moreover, we assume that the integral kernel $K(x,z)$ and the measure $\mu(\cdot)$ of the integro-differential operator S satisfy the following two conditions (4) and (5):

(4) $K \in L^\infty(\mathbf{R}^n \times \mathbf{R}^n)$ with $K(x,y) \geq 0$ almost everywhere in $\mathbf{R}^n \times \mathbf{R}^n$, and satisfies the condition

$$K(x,z) = 0 \quad \text{if } x \in \Omega \text{ and } x + z \notin \overline{\Omega}. \tag{8.3}$$

Probabilistically, the condition (8.3) implies that all jumps from Ω are within $\overline{\Omega}$. Analytically, the condition (8.3) guarantees that the operator S can be interpreted as a mapping acting on functions u which are defined in $\overline{\Omega}$.

(5) $\mu(dz)$ is a Radon measure on $\mathbf{R}^n \setminus \{0\}$ which has a density with respect to the Lebesgue measure dz on $\mathbf{R}^n$ and satisfies the *moment condition*

$$\int_{0<|z|\leq 1} |z|^2 \, \mu(dz) + \int_{|z|\geq 1} |z| \, \mu(dz) < \infty. \tag{8.4}$$

The moment condition (8.4) implies that the measure $\mu(\cdot)$ admits a singularity of order 2 at the origin, and this singularity at the origin is produced by the accumulation of *small jumps* of Markovian particles, while the measure $\mu(\cdot)$ admits a singularity of order 1 at infinity, and this singularity at infinity is produced by the accumulation of *large jumps* of Markovian particles.

Example 8.1. A typical example of the Radon measure $\mu(dz)$ which satisfies the moment condition (8.4) is given by the formula

$$\mu(dz) = \begin{cases} \dfrac{1}{|z|^{n+2-\varepsilon}}\, dz & \text{for } 0 < |z| \leq 1, \\ \dfrac{1}{|z|^{n+1+\varepsilon}}\, dz & \text{for } |z| > 1, \end{cases}$$

where $\varepsilon > 0$.

The operator W is called a second-order, *Waldenfels integro-differential operator* (cf. [96], [11], [79]). The differential operator A is called a *diffusion operator* which describes analytically a strong Markov process with continuous paths (diffusion process) in the interior Ω. In fact, we remark that the differential operator A is *local*, that is, the value $Au(x^0)$ at an interior point $x^0 \in \Omega$ is determined by the values of u in an arbitrary small neighborhood of x^0. Moreover, it is known from Peetre's theorem ([56]) that a linear operator is local if and only if it is a differential operator. The operator S is called a second-order, *Lévy integro-differential operator* which is supposed to correspond to the jump phenomenon in the interior Ω; a Markovian particle moves by jumps to a random point, chosen with kernel $K(x, y)$, in the interior Ω. Therefore, the Waldenfels integro-differential operator W is supposed to correspond to such a diffusion phenomenon that a Markovian particle moves both by jumps and continuously in the state space Ω.

8.1 Mapping properties of Lévy operators

In this section, we consider the Lévy integro-differential operator

$$Su(x) = \int_{\mathbf{R}^n \setminus \{0\}} \left(u(x+z) - u(x) - \sum_{j=1}^{n} z_j \frac{\partial u}{\partial x_j}(x) \right) K(x, z)\, \mu(dz),$$

in the framework of L^p Sobolev spaces. The essential point in the proof is how to estimate the Lévy integral operator Su in terms of Sobolev norms (Theorem 8.1). We show that the operator $\mathcal{W} = (W, \gamma_0)$ may be considered as a perturbation of a compact operator to the operator $\mathcal{A} = (A, \gamma_0)$ in the framework of Sobolev spaces (Theorem 8.3).

8.1.1 Boundedness of Lévy operators defined on $\mathbf{R}^n$

Our main result in this subsection is stated as follows:

Theorem 8.1. *Assume that the moment condition* (8.4) *is satisfied. Then the Lévy operator*

$$S : W^{2,p}(\mathbf{R}^n) \longrightarrow L^p(\mathbf{R}^n)$$

is bounded for all $1 < p < \infty$. *More recisely, there exists a constant* $C > 0$ *such that we have the inequality*

$$\begin{aligned} &\|Su\|_{L^p(\mathbf{R}^n)} \qquad (8.5) \\ &\leq C \, \|K\|_\infty \left(\int_{|z| \geq 1} |z| \; \mu(dz) + \int_{0<|z|\leq 1} |z|^2 \; \mu(dz) \right) \|u\|_{W^{2,p}(\mathbf{R}^n)} \end{aligned}$$

for all $u \in W^{2,p}(\mathbf{R}^n)$.

Proof. In order to prove the L^p-boundedness, we write the integral term $Su(x)$ in the form

$$Su(x) = S_1u(x) + S_2u(x).$$

Here:

$$S_1u(x) := \int_{|z|\geq 1} \left(u(x+z) - u(x) - \sum_{j=1}^{n} z_j \frac{\partial u}{\partial x_j}(x) \right) K(x,z)\, \mu(dz),$$

$$S_2u(x) := \int_{0<|z|\leq 1} \left(u(x+z) - u(x) - \sum_{j=1}^{n} z_j \frac{\partial u}{\partial x_j}(x) \right) K(x,z)\, \mu(dz).$$

First, we need the the following well-known lemma (see Ziemer [100, p. 45-46]):

Lemma 8.2. *Let $0 \leq \gamma \leq 1$ and $1 < p < \infty$. If we let*

$$T_h(f)(x) := f(x+h) - f(x) \quad \text{for } h \in \mathbf{R}^n,$$

then there exists a constant $C > 0$ such that

$$\|T_h(f)\|_{H^{\gamma,p}} \leq C\,|h|\,\|f\|_{H^{1+\gamma,p}} \tag{8.6}$$

for all $f \in H^{1+\gamma,p}(\mathbf{R}^n)$.

(1) First, we estimate the norm $\|S_1u\|_{L^p(\mathbf{R}^n)}$: If we let

$$g_1(x) := \int_{|z|\geq 1} |u(x+z) - u(x)|\; \mu(dz)$$

and

$$g_2(x) := \sum_{j=1}^{n} \left| \frac{\partial u}{\partial x_j}(x) \right| \int_{|z|\geq 1} |z_j|\; \mu(dz),$$

then we have the inequality

$$|S_1u(x)| \leq \|K\|_\infty \left(g_1(x) + g_2(x)\right). \tag{8.7}$$

However, it follows from an application of Minkowskii's inequality for integrals and inequality (8.6) with $\gamma := 0$ that

$$\begin{aligned} \|g_1\|_{L^p(\mathbf{R}^n)} &\leq \int_{|z|\geq 1} \left(\int_{\mathbf{R}^n} |u(x+z) - u(x)|^p\; dx \right)^{1/p} \mu(dz) \qquad (8.8) \\ &\leq C \left(\int_{|z|\geq 1} |z|\, \mu(dz) \right) \|u\|_{W^{1,p}(\mathbf{R}^n)}. \end{aligned}$$

Here and in the following the letter C denotes a generic positive constant independent of $u \in W^{2,p}(\mathbf{R}^n)$.

Moreover, we have the inequality

$$\begin{aligned} \|g_2\|_{L^p(\mathbf{R}^n)} &\leq \sum_{j=1}^n \left\| \frac{\partial u}{\partial x_j} \right\|_{L^p(\mathbf{R}^n)} \left(\sum_{j=1}^n \int_{|z|\geq 1} |z_j| \ \mu(dz) \right) \\ &\leq C \left(\int_{|z|\geq 1} |z| \, \mu(dz) \right) \|u\|_{W^{1,p}(\mathbf{R}^n)} . \end{aligned} \tag{8.9}$$

By combining inequalities (8.7), (8.8) and (8.9), we obtain that

$$\begin{aligned} \|S_1 u\|_{L^p(\mathbf{R}^n)} &\leq \|K\|_\infty \left(\|g_1\|_{L^p(\mathbf{R}^n)} + \|g_2\|_{L^p(\mathbf{R}^n)} \right) \\ &\leq C \, \|K\|_\infty \left(\int_{|z|\geq 1} |z| \ \mu(dz) \right) \|u\|_{W^{1,p}(\mathbf{R}^n)} . \end{aligned} \tag{8.10}$$

(2) In order to estimate the norm $\|S_2 u\|_p$, by using Taylor's formula we obtain that

$$\begin{aligned} & u(x+z) - u(x) - \sum_{j=1}^n z_j \frac{\partial u}{\partial x_j}(x) \\ &= \sum_{j=1}^n z_j \int_0^1 \left(\frac{\partial u}{\partial x_j}(x+tz) - \frac{\partial u}{\partial x_j}(x) \right) \, dt. \end{aligned}$$

Hence we have the inequality

$$\begin{aligned} & |S_2 u(x)| \\ &\leq \|K\|_\infty \sum_{j=1}^n \int_0^1 \int_{0<|z|\leq 1} |z| \left| \frac{\partial u}{\partial x_j}(x+tz) - \frac{\partial u}{\partial x_j}(x) \right| \mu(dz) \, dt. \end{aligned}$$

By Minkowskii's inequality for integrals, it follows that

$$\begin{aligned} & \|S_2 u\|_{L^p(\mathbf{R}^n)} \\ &\leq C \, \|K\|_\infty \int_0^1 \int_{0<|z|\leq 1} |z| \sum_{j=1}^n \left(\int_{\mathbf{R}^n} \left| \frac{\partial u}{\partial x_j}(x+tz) - \frac{\partial u}{\partial x_j}(x) \right|^p \, dx \right)^{1/p} \\ & \quad \times \mu(dz) \, dt. \end{aligned} \tag{8.11}$$

However, by applying inequality (8.6) with $\gamma := 0$ we obtain that

$$\sum_{j=1}^n \left(\int_{\mathbf{R}^n} \left| \frac{\partial u}{\partial x_j}(x+tz) - \frac{\partial u}{\partial x_j}(x) \right|^p \, dx \right)^{1/p}$$

$$\leq Ct\,|z| \sum_{j=1}^{n} \left\| \frac{\partial u}{\partial x_j} \right\|_{W^{1,p}(\mathbf{R}^n)}$$

$$\leq Ct\,|z|\, \|u\|_{W^{2,p}(\mathbf{R}^n)} \quad \text{for all } 0 \leq t \leq 1.$$

Therefore, we have, by inequality (8.11),

$$\begin{aligned} &\|S_2 u\|_{L^p(\mathbf{R}^n)} \\ &\leq C\,\|K\|_\infty\, \|u\|_{W^{2,p}(\mathbf{R}^n)} \int_0^1 \int_{0<|z|\leq 1} t\,|z|^2\, \mu(dz)\, dt \\ &= \frac{1}{2}\, C\, \|K\|_\infty \left(\int_{0<|z|\leq 1} |z|^2\, \mu(dz) \right) \|u\|_{W^{2,p}(\mathbf{R}^n)}. \end{aligned} \tag{8.12}$$

The desired inequality (8.5) follows by combining inequalities (8.10) and (8.12):

$$\begin{aligned} &\|Su\|_{L^p(\mathbf{R}^n)} \\ &\leq \|S_1 u\|_{L^p(\mathbf{R}^n)} + \|S_2 u\|_{L^p(\mathbf{R}^n)} \\ &\leq C\,\|K\|_\infty \left(\int_{|z|\geq 1} |z|\, \mu(dz) + \int_{0<|z|\leq 1} |z|^2\, \mu(dz) \right) \|u\|_{W^{2,p}(\mathbf{R}^n)}. \end{aligned}$$

The proof of Theorem 8.1 is complete. □

8.1.2 The case of a bounded domain

In this subsection, we prove the following version of Theorem 8.1 with respect to the bounded domain Ω:

Theorem 8.3. *Assume that the moment condition* (8.4) *is satisfied. Then the operator*

$$S : W^{2,p}(\Omega) \longrightarrow L^p(\Omega)$$

is bounded *for all* $1 < p < \infty$. *Moreover, if* $n < p < \infty$, *then the operator*

$$S : W^{2,p}(\Omega) \longrightarrow L^p(\Omega)$$

is compact.

Proof. The proof of Theorem 8.3 is divided into two steps.

Step 1: By restriction arguments, we find from inequality (8.5) that the operator

$$S : W^{2,p}(\Omega) \longrightarrow W^{2,p}(\mathbf{R}^n) \xrightarrow{S} L^p(\mathbf{R}^n) \longrightarrow L^p(\Omega)$$

is bounded.

Step 2: In order to prove the *compactness*, we make use of Bony–Courrège–Priouret [11, Théorème XXI]. We show that the operator

$$S : W^{2,p}(\Omega) \longrightarrow L^p(\Omega)$$

is compact for $n < p < \infty$.

We recall that $K(x,z) = 0$ if $x + z \notin \overline{\Omega}$. Hence there exists a compact set $M \subset \mathbf{R}^n$ such that

$$Su(x) = \int_M \left(u(x+z) - u(x) - \sum_{j=1}^n z_j \frac{\partial u}{\partial x_j}(x) \right) K(x,z)\, \mu(dz)$$

may be interpreted as a mapping acting on functions u defined $\overline{\Omega}$. Indeed, we may take

$$M = \overline{\cup_{x \in \overline{\Omega}} \left\{ \overline{\Omega} - x \right\}}.$$

First, we take a smooth function χ in $C_0^\infty(\mathbf{R})$ such that

$$\chi(t) = \begin{cases} 1 & \text{if } |t| \leq 1, \\ 0 & \text{if } |t| \geq 2. \end{cases}$$

For every $0 < \varepsilon < 1$, we let

$$\Phi_\varepsilon(z) := 1 - \chi\left(\frac{|z|}{\varepsilon} \right).$$

We remark that

$$\Phi_\varepsilon(z) = \begin{cases} 0 & \text{if } |z| \leq \varepsilon, \\ 1 & \text{if } |z| \geq 2\varepsilon. \end{cases}$$

Moreover, by Dini's theorem it follows that the sequence $\Phi_\varepsilon(x,z)$ converges *uniformly* to 1 for $z \neq 0$, as $\varepsilon \downarrow 0$.

Now we introduce a family of truncation operators given by the formula

$$S_{\Phi_\varepsilon} u(x) = \int_M \left(u(x+z) - u(x) - \sum_{j=1}^n z_j \frac{\partial u}{\partial x_j}(x) \right) K(x,z) \Phi_\varepsilon(z)\, \mu(dz).$$

Then it is easy to see that the operator

$$S_{\Phi_\varepsilon} : C^1(\overline{\Omega}) \longrightarrow L^\infty(\Omega)$$

is bounded. Indeed, we have, by the mean value theorem,

$$|S_{\Phi_\varepsilon} u(x)|$$

$$\begin{aligned}
&\leq \int_M \left| u(x+z) - u(x) - \sum_{j=1}^n z_j \frac{\partial u}{\partial x_j}(x) \right| |K(x,z)| \, \Phi_\varepsilon(z) \, \mu(dz) \\
&\leq \int_{z \in M, |z| \geq \varepsilon} |u(x+z) - u(x)| \, |K(x,z)| \, \mu(dz) \\
&\quad + \sum_{j=1}^n \int_{z \in M, |z| \geq \varepsilon} |z_j| \, |K(x,z)| \, \mu(dz) \left| \frac{\partial u}{\partial x_j}(x) \right| \\
&\leq C \, \|K\|_\infty \left(\int_{z \in M, |z| \geq \varepsilon} |z| \, \mu(dz) \right) \|u\|_{C^1(\overline{\Omega})} \\
&\leq C \, \|K\|_\infty \left(\frac{1}{\varepsilon} \int_{\varepsilon \leq |z| < 1} |z|^2 \, \mu(dz)) + \int_{|z| \geq 1} |z| \, \mu(dz) \right) \|u\|_{C^1(\overline{\Omega})} \\
&\qquad \text{for all } x \in \overline{\Omega}.
\end{aligned}$$

Since the embedding $W^{2,p}(\Omega) \hookrightarrow C^1(\overline{\Omega})$ is compact for $n < p < \infty$ and since the embedding $L^\infty(\Omega) \hookrightarrow L^p(\Omega)$ is continuous, we obtain that the operator

$$S_{\Phi_\varepsilon} : W^{2,p}(\Omega) \longrightarrow L^p(\Omega)$$

is *compact*. The situation can be visualized as follows:

$$\boxed{S_{\Phi_\varepsilon} : W^{2,p}(\Omega) \underset{\text{compactly}}{\hookrightarrow\hookrightarrow} C^1(\overline{\Omega}) \xrightarrow{S_{\Phi_\varepsilon}} L^\infty(\Omega) \longrightarrow L^p(\Omega).}$$

On the other hand, we obtain from inequality (8.5) that

$$\|S_{\Phi_\varepsilon} u\|_{L^p(\Omega)} \leq C \, \|u\|_{W^{2,p}(\Omega)} \quad \text{for all } u \in W^{2,p}(\Omega) \text{ and all } 0 < \varepsilon < 1.$$

Furthermore, it follows from an application of Lebesgue's dominated convergence theorem that

$$S_{\Phi_\varepsilon} \longrightarrow S \quad \text{as } \varepsilon \downarrow 0 \tag{8.13}$$

with respect to the operator norm in the space $\mathcal{L}\left(W^{2,p}(\Omega), L^p(\Omega)\right)$ of bounded linear operators on $W^{2,p}(\Omega)$ into $L^p(\Omega)$. Indeed, just as in the proof of inequality (8.5) we obtain that

$$\begin{aligned}
&\|Su - S_{\Phi_\varepsilon} u\|_{L^p(\Omega)} \\
&\leq C \, \|K\|_\infty \left(\int_{|z| \geq 1} |z| \, \chi\left(\frac{|z|}{\varepsilon}\right) \mu(dz) + \int_{0 < |z| \leq 1} |z|^2 \, \chi\left(\frac{|z|}{\varepsilon}\right) \mu(dz) \right) \\
&\qquad \times \|u\|_{W^{2,p}(\Omega)} \quad \text{for all } u \in W^{2,p}(\Omega).
\end{aligned}$$

However, by condition (8.4) it follows from the Lebesgue dominated convergence theorem that

$$\lim_{\varepsilon\downarrow 0}\left(\int_{|z|\geq 1}|z|\,\chi\left(\frac{|z|}{\varepsilon}\right)\mu(dz)+\int_{0<|z|\leq 1}|z|^2\,\chi\left(\frac{|z|}{\varepsilon}\right)\,\mu(dz)\right)=0.$$

Summing up, we obtain from assertion (8.13) that the operator

$$S: W^{2,p}(\Omega)\longrightarrow L^p(\Omega)$$

is *compact* for $n<p<\infty$.

The proof of Theorem 8.3 is complete. □

Furthermore, we can obtain an L^∞-version of Theorem 8.1 as follows:

Lemma 8.4. *For every $\varepsilon>0$, there exists a constant $C(\varepsilon)>0$ such that we have, for all $u\in C^2(\overline{\Omega})$,*

$$\begin{aligned}&\|Su\|_{L^\infty(\Omega)} \\ &\leq \frac{1}{2}\sigma(\varepsilon)\,\|K\|_\infty\left\|\nabla^2 u\right\|_{L^\infty(\Omega)}+C(\varepsilon)\,\|K\|_\infty\,\|\nabla u\|_{L^\infty(\Omega)}.\end{aligned}\tag{8.14}$$

Here

$$\sigma(\varepsilon)=\int_{0<|z|\leq\varepsilon}|z|^2\ \mu(dz).$$

Proof. For each $\varepsilon>0$, we decompose the integral term $Su(x)$ into the two terms $S_\varepsilon^{(1)}u$ and $S_\varepsilon^{(2)}u$ as follows:

$$Su(x)=S_\varepsilon^{(1)}u(x)+S_\varepsilon^{(2)}u(x).$$

Here:

$$\begin{aligned}S_\varepsilon^{(1)}u(x)&:=\int_{0<|z|\leq\varepsilon}\left(u(x+z)-u(x)-\sum_{j=1}^n z_j\frac{\partial u}{\partial x_j}(x)\right)K(x,z)\,\mu(dz)\\ &=\int_0^1(1-t)\,dt\int_{0<|z|\leq\varepsilon}z\cdot\nabla^2 u(x+tz)z\,K(x,z)\,\mu(dz)\end{aligned}$$

and

$$S_\varepsilon^{(2)}u(x):=\int_{|z|>\varepsilon}(u(x+z)-u(x)-z\cdot\nabla u)\,\mu(dz).$$

(1) First, we have the inequality

$$\left\|S_\varepsilon^{(1)}u\right\|_{L^p(\Omega)}\leq\frac{1}{2}\,\|K\|_\infty\left(\int_{0<|z|\leq\varepsilon}|z|^2\ \mu(dz)\right)\left\|\nabla^2 u\right\|_{L^\infty(\Omega)}\tag{8.15}$$

$$= \frac{1}{2}\sigma(\varepsilon)\,\|K\|_\infty \left\|\nabla^2 u\right\|_{L^\infty(\Omega)}.$$

By condition (8.4), it follows from an application of Lebesgue's dominated convergence theorem that

$$\lim_{\varepsilon\downarrow 0}\sigma(\varepsilon) = 0. \tag{8.16}$$

(2) Secondly, we rewrite the term $S_\varepsilon^{(2)}u$ in the form

$$\begin{aligned} S_\varepsilon^{(2)}u(x) &= \int_{|z|>\varepsilon} K(x,z)\,(u(x+z)-u(x))\,\mu(dz) \\ &\quad + \sum_{j=1}^{n}\int_{|z|>\varepsilon} K(x,z)z_j\,\frac{\partial u}{\partial x_j}(x)\,\mu(dz) \\ &:= A(x) + B(x). \end{aligned}$$

Then, by using condition (8.4) we can estimate the term $B(x)$ as follows:

$$\begin{aligned} |B(x)| &\le \int_{|z|>\varepsilon} K(x,z)|z|\cdot\sum_{j=1}^{n}\left|\frac{\partial u}{\partial x_j}(x)\right|\mu(dz) \\ &\le \|K\|_\infty\left(\int_{|z|>\varepsilon}|z|\,\mu(dz)\right)\|\nabla u\|_{L^\infty(\Omega)} \\ &= \delta(\varepsilon)\,\|K\|_\infty\,\|\nabla u\|_{L^\infty(\Omega)} \quad \text{for all } x\in\Omega, \end{aligned}$$

where

$$\delta(\varepsilon) := \int_{|z|>\varepsilon}|z|\,\mu(dz).$$

However, the term $\delta(\varepsilon)$ can be estimated as follows:

$$\begin{aligned} \delta(\varepsilon) &= \int_{|z|>1}|z|\,\mu(dz) + \int_{\varepsilon<|z|\le 1}|z|\,\mu(dz) \\ &\le \int_{|z|>1}|z|\,\mu(dz) + \frac{1}{\varepsilon}\int_{\varepsilon<|z|\le 1}|z|^2\,\mu(dz) \\ &\le \int_{|z|>1}|z|\,\mu(dz) + \frac{1}{\varepsilon}\int_{0<|z|\le 1}|z|^2\,\mu(dz) \\ &= C_2 + \frac{C_1}{\varepsilon}. \end{aligned}$$

Hence we obtain the inequality

$$\|B\|_{L^\infty(\Omega)} \le \left(\frac{C_1}{\varepsilon} + C_2\right)\|K\|_\infty\,\|\nabla u\|_{L^\infty(\Omega)}. \tag{8.17}$$

On the other hand, by Morrey's imbedding theorem (see [2, Lemma 4.28], [33, Theorem 7.17], [80, Lemma 4.7]) we can find a constant $C > 0$ such that

$$|u(x+z) - u(x)| \le C\,|z|^{1-n/p}\,\|\nabla u\|_{L^p(\Omega)}\,.$$

Hence it follows that

$$\begin{aligned}
|A(x)| &\le \int_{|z|>\varepsilon} K(x,z)\,(u(x+z)-u(x))\,\mu(dz)\\
&\le C\,\|K\|_\infty \int_{|z|>\varepsilon} |z|^{1-n/p}\,\mu(dz)\cdot\|\nabla u\|_{L^p(\Omega)}\\
&= C\,\|K\|_\infty \int_{|z|>\varepsilon} |z|\cdot\frac{1}{|z|^{n/p}}\,\mu(dz)\,\|\nabla u\|_{L^p(\Omega)}\\
&\le \frac{C\,\|K\|_\infty}{\varepsilon^{n/p}}\left(\int_{|z|>\varepsilon}|z|\,\mu(dz)\right)\|\nabla u\|_{L^p(\Omega)}\\
&= \delta(\varepsilon)\,\frac{C\,\|K\|_\infty}{\varepsilon^{n/p}}\,\|\nabla u\|_{L^p(\Omega)}\\
&\le \left(\frac{C_1}{\varepsilon}+C_2\right)\frac{C\,\|K\|_\infty}{\varepsilon^{n/p}}\,\|\nabla u\|_{L^p(\Omega)} \quad \text{for all } x\in\Omega.
\end{aligned}$$

This proves that

$$\begin{aligned}
\|A\|_{L^\infty(\Omega)} &\le \left(\frac{C_1}{\varepsilon}+C_2\right)\frac{C\,\|K\|_\infty}{\varepsilon^{n/p}}\,\|\nabla u\|_{L^p(\Omega)}\\
&= C_\varepsilon\,\|K\|_\infty\,\|\nabla u\|_{L^p(\Omega)}\\
&\le C_\varepsilon\,\|K\|_\infty\,C_3\,|\Omega|^{1/p}\,\|\nabla u\|_{L^\infty(\Omega)}\,,
\end{aligned} \tag{8.18}$$

where

$$C_\varepsilon = \frac{C}{\varepsilon^{n/p}}\left(\frac{C_1}{\varepsilon}+C_2\right)$$

and $|\Omega|$ is the volume of the domain Ω.

By combining estimates (8.17) and (8.18), we obtain that

$$\begin{aligned}
\left\|S_\varepsilon^{(2)}u\right\|_{L^\infty(\Omega)} &\le \|A\|_{L^\infty(\Omega)} + \|B\|_{L^\infty(\Omega)}\\
&\le C(\varepsilon)\,\|K\|_\infty\,\|\nabla u\|_{L^\infty(\Omega)}\,,
\end{aligned} \tag{8.19}$$

with some constant $C(\varepsilon) > 0$ depending on ε.

(3) The desired estimate (8.14) follows by combining estimates (8.15) and (8.19):

$$\|Su\|_{L^\infty(\Omega)} \le \|S_\varepsilon^{(1)}u\|_{L^\infty(\Omega)} + \|S_\varepsilon^{(2)}u\|_{L^\infty(\Omega)}$$

$$\leq \frac{1}{2}\sigma(\varepsilon)\,\|K\|_\infty\,\left\|\nabla^2 u\right\|_{L^\infty(\Omega)} + C(\varepsilon)\,\|K\|_\infty\,\|\nabla u\|_{L^\infty(\Omega)}\,.$$

The proof of Lemma 8.4 is complete. □

8.2 Weak Maximum Principle

The purpose of this section is to prove a variant of the weak maximum principle in the framework of L^p Sobolev spaces, essentially due to Bony [9] (cf. [9, Théorème 2], [33, Section 9.1, Theorem 9.1], [93, Chapter 3, Lemma 3.25]):

Theorem 8.5 (the weak maximum principle)**.** *Assume that a function $u \in W^{2,p}(\Omega)$, with $n < p < \infty$, satisfies the condition*

$$Wu(x) \geq 0 \quad \textit{for almost all } x \in \Omega. \tag{8.20}$$

Then we have the inequality

$$\max_{\overline{\Omega}} u \leq \max_{\partial\Omega} u^+,$$

where

$$u^+(x) = \max\{u(x), 0\} \quad \textit{for } x \in \overline{\Omega}.$$

Here it should be noticed that we have, by Sobolev's imbedding theorem (see Theorem 7.3),

$$W^{2,p}(\Omega) \subset C^1(\overline{\Omega}),$$

since $2 - n/p > 1$ for $n < p < \infty$.

The proof of Theorem 8.5 is divided into Step (I) and Step (II).

Step (I): The next lemma, due to Bony [9, Théorème 1], plays an essential role in the proof of Theorem 8.5 (cf. [93, Chapter 3, Lemma 3.24]):

Proposition 8.6 (Bony)**.** *If a function $u \in W^{2,p}(\Omega)$, with $n < p < \infty$, attains a local maximum at a point $x_0 \in \Omega$, then we have the assertion*

$$\begin{aligned}
&\lim \operatorname{ess\,inf}_{x\to x_0}\left(\sum_{i,j=1}^n a^{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j}(x)\right)\\
&= \lim_{\rho\downarrow 0}\left\{\operatorname{ess\,inf}_{B(x_0,\rho)}\left(\sum_{i,j=1}^n a^{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j}(x)\right)\right\}\\
&\leq 0,
\end{aligned}$$

where $B(x_0, \rho)$ denote the open ball of radius ρ about x_0. More precisely, for every neighborhood $U(x_0)$ of x_0 there exists a subset M of $U(x_0)$, with positive Lebesgue measure*, such that*

$$\sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}(x) \le 0 \quad \textit{for all } x \in M.$$

Proof. The proof of Proposition 8.6 is divided into four steps.

Step 1: First, we may assume that $u \in W^{2,p}(\Omega)$ attains a *strict* local maximum at a point $x_0 \in \Omega$. Indeed, if we consider the function

$$\widetilde{u}(x) := u(x) - \varphi(x) = u(x) - |x - x_0|^4 \quad \text{for } x \in \Omega,$$

then it is easy to verify that $\widetilde{u} \in W^{2,p}(\Omega)$ attains a strict local maximum at a point x_0 and further that

$$\begin{aligned} &\lim \operatorname{ess\,inf}_{x \to x_0} \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2 \widetilde{u}}{\partial x_i \partial x_j}(x) \\ &= \lim \operatorname{ess\,inf}_{x \to x_0} \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}(x), \end{aligned}$$

since we have the assertion

$$\lim \operatorname{ess\,inf}_{x \to x_0} \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(x) = 0.$$

Step 2: Now we assume that $u \in W^{2,p}(\Omega)$ attains a *strict* local maximum at the point x_0 in an open neighborhood $U(x_0$ in Ω. We define the *upper contact set* $\mathcal{W}$ of u by the formula

$$\begin{aligned} &\mathcal{W} \qquad (8.21) \\ &:= \{x \in U(x_0) : u(y) \le u(x) + \langle \nabla u(x), y - x \rangle \quad \text{for all } y \in \Omega\}. \end{aligned}$$

Then we can find a constant $r > 0$ such that we have, for each $\eta \in \mathbf{R}^n$ satisfying $|\eta| < r$,

$$u(x) < u(x_0) + \langle \eta, x - x_0 \rangle \quad \text{for all } x \in \Omega. \qquad (8.22)$$

If we let

$$\varphi(x, \eta) = u(x) - \langle x, \eta \rangle \quad \text{for } x \in \Omega,$$

then inequality (8.22) can be rewritten in the form

$$\varphi(x, \eta) < \varphi(x_0, \eta) \quad \text{for all } x \in \Omega.$$

This implies that the function $\varphi(x,\eta)$ takes its maximum in the interior Ω. Hence there exists a point $\widetilde{x} \in \Omega$, depending on η, such that

$$\varphi(y,\eta) \le \varphi(\widetilde{x},\eta) \quad \text{for all } y \in \Omega,$$

or equivalently, we have, for each $\eta \in \mathbf{R}^n$ satisfying $|\eta| < r$,

$$u(y) \le u(\widetilde{x}) + \langle \eta, y - \widetilde{x} \rangle \quad \text{for all } y \in \Omega. \tag{8.23}$$

In other words, we may translate vertically the plane

$$z = u(x) + \langle \eta, y - x \rangle$$

to the highest such position $\widetilde{x}$, that is, the surface

$$z = u(y) \quad \text{for } y \in \Omega,$$

lies below the plane.

More precisely, we have the following claim (see [9, Lemme 2]):

Claim 8.1. $\eta = \nabla u(\widetilde{x})$.

Proof. If we let

$$\Phi(y) := u(\widetilde{x}) - u(y) + \langle \eta, y - \widetilde{x} \rangle \quad \text{for } y \in \Omega,$$

then it follows from inequality (8.23) that

$$\begin{aligned} &\Phi(y) \ge 0 \quad \text{for all } y \in \Omega, \\ &\Phi(\widetilde{x}) = 0. \end{aligned}$$

This implies that $\nabla_y \Phi(\widetilde{x}) = 0$, so that $\eta = \nabla u(\widetilde{x})$.

The proof of Claim 8.1 is complete. □

By Claim 8.1, we can rewrite inequality (8.23) as follows:

$$u(y) \le u(\widetilde{x}) + \langle \nabla u(\widetilde{x}), y - \widetilde{x} \rangle \quad \text{for all } y \in \Omega.$$

This proves that $\widetilde{x} \in \mathcal{W}$.

Summing up, we have proved the following claim:

Claim 8.2. Let $\mathcal{W}$ be the *upper contact set* of u defined by formula (8.21). Then we have, for some constant r,

$$B(0,r) \subset \{\nabla u(\widetilde{x}) : \widetilde{x} \in \mathcal{W}\}.$$

Step 3: We prove that the Lebesgue measure of $|\mathcal{W}|$ is *positive*. To do this, we introduce a mapping

$$\begin{aligned} F : U(x_0) &\longrightarrow \mathbf{R}^n \\ x &\longmapsto \nabla u(x), \end{aligned}$$

where $u \in W^{2,p}(\Omega)$ with $n < p < \infty$. Then we have, by Claim 8.2,

$$B(0,r) \subset F(\mathcal{W}).$$

This proves that $|F(\mathcal{W})| > 0$, since the set $F(\mathcal{W})$ contains the open ball $B(0,r)$.

Furthermore, the next lemma, due to Bony [9, Lemme 1], proves that $|\mathcal{W}| > 0$:

Lemma 8.7 (Bony). *Let $F = (f_1, f_2, \dots, f_n)$ be a $W^{1,p}$ mapping of an open subset Ω of $\mathbf{R}^n$ into $\mathbf{R}^n$, where $n < p < \infty$. If $\mathcal{S}$ is a subset of zero Lebesgue measure, then the Lebesgue measure of the image $F(\mathcal{S})$ is equal to zero, that is,*

$$|\mathcal{S}| = 0 \implies |F(\mathcal{S})| = 0.$$

Proof. (1) If $\gamma > 0$, we let (see Figure 8.1)

$$\begin{aligned} C_\gamma := &\text{a compact cube with sides parallel to the axes} \\ &\text{and non-empty interior, with side length } \gamma, \end{aligned}$$

and we estimate the Lebesgue measure $\mu(F(C_\gamma))$ of the image $F(C_\gamma)$:

$$F(C_\gamma) = \{(f_1(x), f_2(x), \dots, f_n(x) : x \in C_\gamma\}.$$

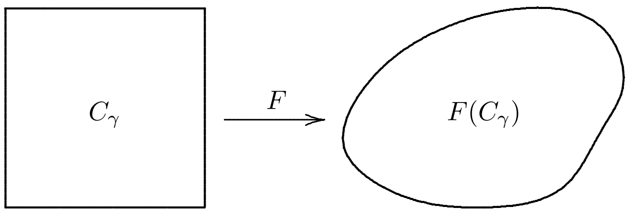

Fig. 8.1. The compact cube C_γ with sides parallel to the axes and the image $F(C_\gamma)$

To do this, we make use of a special case of Morrey's theorem (see [2, Lemma 4.28], [33, Theorem 7.17]; [80, Lemma 4.7]):

Claim 8.3 (Morrey). Let Ω be a cube in $\mathbf{R}^n$ and $n < p < \infty$. Then there exists a positive constant $C = C(p,n)$ such that we have, for all $u \in W^{1,p}(\Omega)$,

$$\sup_{\substack{x,y\in\Omega \\ x\neq y}} \frac{|u(x) - u(y)|}{|x-y|^{1-n/p}} \leq C \, \|\nabla u\|_{L^p(\Omega)} . \tag{8.24}$$

Proof. By a density argument, it suffices to show that we have, for all $u \in C^1(\overline{\Omega})$,

$$|u(x) - u(y)| \leq C \, \|\nabla u\|_{L^p(\Omega)} \, |x - y|^{1-n/p} \quad \text{for } x,\, y \in \Omega. \tag{8.24$'$}$$

Furthermore, by a non-singular linear transformation we may assume that Ω is a cube having unit side length.

For $0 < t < 1$, we denote by Q_t a subset of Ω which is a closed cube having side length t and faces parallel to those to Ω. If $x,\, y \in \Omega$ and

$$\sigma := |x - y| < 1,$$

we can find a cube Q_σ such that $x,\, y \in Q_\sigma$. Then we have, for all $z \in Q_\sigma$,

$$\begin{aligned} u(x) - u(z) &= -\int_0^1 \frac{d}{dt} u(x + t(z-x))\, dt \\ &= -\int_0^1 \sum_{j=1}^n (z_j - x_j) \cdot \frac{\partial u}{\partial x_j}(x + t(z-x))\, dt, \end{aligned}$$

and so, by Schwarz's inequality (Theorem 3.14 with $p = q := 2$),

$$|u(x) - u(z)| \leq \sqrt{n}\, \sigma \int_0^1 |\nabla u(x + t(z-x))|\, dt.$$

Hence it follows from an application of Hölder's inequality (Theorem 3.14) with $q := p/(p-1)$ that

$$\begin{aligned} &\left| u(x) - \frac{1}{\sigma^n} \int_{Q_\sigma} u(z)\, dz \right| \qquad (8.25) \\ &= \left| \frac{1}{\sigma^n} \int_{Q_\sigma} (u(x) - u(z))\, dz \right| \\ &\leq \frac{1}{\sigma^n} \int_{Q_\sigma} |u(x) - u(z)|\, dz \\ &\leq \frac{\sqrt{n}}{\sigma^{n-1}} \int_{Q_\sigma} dz \int_0^1 |\nabla u(x + t(z-x))|\, dt \\ &= \frac{\sqrt{n}}{\sigma^{n-1}} \int_0^1 \left(\int_{Q_\sigma} |\nabla u(x + t(z-x))|\, dz \right) dt \\ &= \frac{\sqrt{n}}{\sigma^{n-1}} \int_0^1 \left(\int_{Q_{t\sigma} + (1-t)x} |\nabla u(\zeta)|\, d\zeta \right) t^{-n}\, dt \\ &\leq \frac{\sqrt{n}}{\sigma^{n-1}} \left(\int_\Omega |\nabla u(z)|^p\, dz \right)^{1/p} \int_0^1 t^{-n} \left(|Q_{t\sigma} + (1-t)x| \right)^{1/q}\, dt \end{aligned}$$

$$= \frac{\sqrt{n}}{\sigma^{n-1}} \left(\int_{\Omega} |\nabla u(z)|^p \, dz \right)^{1/p} \int_0^1 t^{-n} \left((t\sigma)^n \right)^{1/q} \, dt$$
$$= K\sigma^{1-n/p} \|\nabla u\|_{L^p(\Omega)} ,$$

with

$$K := K(n,p) = \sqrt{n} \int_0^1 t^{-n/p} \, dt = \frac{\sqrt{n}\, p}{p-n}.$$

Similarly, we have, with y in place of x in inequality (8.25),

$$\left| u(y) - \frac{1}{\sigma^n} \int_{Q_\sigma} u(z)\, dz \right| \leq K\sigma^{1-n/p} \|\nabla u\|_{L^p(\Omega)} . \tag{8.26}$$

Therefore, by combining inequalities (8.25) and (8.26) we obtain that

$$\begin{aligned} |u(x) - u(y)| &\leq \left| u(x) - \frac{1}{\sigma^n} \int_{Q_\sigma} u(z)\, dz \right| + \left| \frac{1}{\sigma^n} \int_{Q_\sigma} u(z)\, dz - u(y) \right| \\ &\leq 2K \|\nabla u\|_{L^p(\Omega)} \sigma^{1-n/p} \\ &= 2K \|\nabla u\|_{L^p(\Omega)} |x-y|^{1-n/p} \quad \text{for all } x,\, y \in \Omega. \end{aligned}$$

This proves the desired inequality (8.24′) with $C := 2K$.

The proof of Claim 8.3 is complete. □

By applying Claim 8.3 with

$$n := n, \quad \Omega := C_\gamma, \quad u := f_i \text{ for } 1 \leq i \leq n,$$

we obtain that, for some positive constant $K = K(p,n)$,

$$\operatorname{osc} f_i := \sup_{x,y \in C_\gamma} |f_i(x) - f_i(y)| \leq K\gamma^{1-n/p} \|\nabla f_i\|_{L^p(C_\gamma)} , \tag{8.27}$$
$$1 \leq i \leq n.$$

Since the Lebesgue measure $\mu(F(C_\gamma))$ is estimated by $\operatorname{osc} f_1 \times \operatorname{osc} f_2 \times \cdots \times \operatorname{osc} f_n$, it follows from an application of inequality (8.27) that

$$\begin{aligned} &|F(C_\gamma)| \\ &\leq K^n \gamma^{n(1-n/p)} \|\nabla f_1\|_{L^p(C_\gamma)} \times \ldots \times \|\nabla f_n\|_{L^p(C_\gamma)} . \end{aligned} \tag{8.28}$$

(2) Now we assume that

$$|\mathcal{S}| = 0.$$

Then, for any $\varepsilon > 0$ we can find a family of non-overlapping cubes C_α with side length γ_α, $\alpha \in \Lambda$, such that

$$\mathcal{S} \subset \bigcup_{\alpha \in \Lambda} C_\alpha,$$

$$\sum_{\alpha\in\Lambda} \gamma_\alpha^n \le \varepsilon.$$

By the subadditivity of the Lebesgue measure, it follows from inequality (8.28) that

$$\begin{aligned}
&|F(\mathcal{S})| \qquad (8.29)\\
&\le \sum_{\alpha\in\Lambda} |F(C_\alpha)|\\
&\le K^n \sum_{\alpha\in\Lambda} \gamma_\alpha^{n(1-n/p)} \|\nabla f_1\|_{L^p(C_\alpha)} \times \ldots \times \|\nabla f_n\|_{L^p(C_\alpha)}\\
&\le K^n \sum_{\alpha\in\Lambda} \gamma_\alpha^{n(1-n/p)} \Big(\|\nabla f_1\|_{L^p(C_\alpha)}^p + \ldots + \|\nabla f_n\|_{L^p(C_\alpha)}^p\Big)^{n/p}.
\end{aligned}$$

However, by applying a discrete version of Hölder's inequality (Theorem 3.14) to the last term of inequality (8.29) we obtain that

$$\begin{aligned}
&\sum_{\alpha\in\Lambda} \gamma_\alpha^{n(1-n/p)} \Big(\|\nabla f_1\|_{L^p(C_\alpha)}^p + \ldots + \|\nabla f_n\|_{L^p(C_\alpha)}^p\Big)^{n/p} \qquad (8.30)\\
&\le \left(\sum_{\alpha\in\Lambda} \gamma_\alpha^n\right)^{1-n/p} \left(\sum_{\alpha\in\Lambda} \Big(\|\nabla f_1\|_{L^p(C_\alpha)}^p + \ldots + \|\nabla f_n\|_{L^p(C_\alpha)}^p\Big)\right)^{n/p}\\
&\le \left(\sum_{\alpha\in\Lambda} \gamma_\alpha^n\right)^{1-n/p} \Big(\|\nabla f_1\|_{L^p(\cup_{\alpha\in\Lambda} C_\alpha)}^p + \ldots + \|\nabla f_n\|_{L^p(\cup_{\alpha\in\Lambda} C_\alpha)}^p\Big)^{n/p}\\
&\le \varepsilon^{1-n/p} \Big(\|\nabla f_1\|_{L^p(\Omega)}^p + \ldots + \|\nabla f_n\|_{L^p(\Omega)}^p\Big)^{n/p}.
\end{aligned}$$

Therefore, we have, by inequalities (8.29) and (8.30),

$$\begin{aligned}
&|F(\mathcal{S})|\\
&\le K^n \sum_{\alpha\in\Lambda} \gamma_\alpha^{n(1-n/p)} \Big(\|\nabla f_1\|_{L^p(C_\alpha)} \times \ldots \times \|\nabla f_n\|_{L^p(C_\alpha)}\Big)^{n/p}\\
&\le K^n\, \varepsilon^{1-n/p} \Big(\sum_{\alpha\in\Lambda} \Big(\|\nabla f_1\|_{L^p(C_\alpha)}^p + \ldots + \|\nabla f_n\|_{L^p(C_\alpha)}^p\Big)\Big)^{n/p}\\
&\le K^n\, \varepsilon^{1-n/p} \Big(\|\nabla f_1\|_{L^p(\Omega)}^p + \ldots + \|\nabla f_n\|_{L^p(\Omega)}^p\Big)^{n/p}\\
&= K^n\, \varepsilon^{1-n/p} \Big(\|\nabla F\|_{L^p(\Omega)}\Big)^n.
\end{aligned}$$

This implies that $|F(\mathcal{S})| = 0$, since $\varepsilon > 0$ is arbitrary.

The proof of Lemma 8.7 is complete. □

By applying Lemma 8.7 to our situation, we obtain that $|\mathcal{W}| > 0$. In other words, for every neighborhood $U(x_0)$ of x_0 there exists a subset $\mathcal{W}$ of $U(x_0)$, with positive Lebesgue measure, such that

$$u(y) \leq u(x) + \langle \nabla u(x), y - x \rangle \quad \text{for all } x \in \mathcal{W} \text{ and all } y \in \Omega. \tag{8.31}$$

Step 4: Moreover, we need the following claim (see [16, Theorem 12], [45, Theorem 1.72]):

Claim 8.4 (Calderón–Zygmund). If $u \in W^{2,p}(\Omega)$ with $n < p < \infty$, then it is almost everywhere two times differentiable in the usual sense.

By Claim 8.4, we may assume that $u(x)$ is two times differentiable almost everywhere in $\mathcal{W}$.

Now we take an arbitrary vector ξ in the unit sphere Σ_{n-1} in $\mathbf{R}^n$, and let

$$\phi(t) := u(x + t\xi) \quad \text{for } x \in \widetilde{M} \text{ and } |t| < \operatorname{dist}(x, \partial U(x_0)).$$

By the change of variables, we may assume that, at almost all $x \in W$ each first partial derivative $\partial u/\partial x_i$, $1 \leq i \leq n$, has a directional derivative with respect to ξ in the usual sense. Then we have, by Taylor's formula,

$$\phi(t) = \phi(0) + t\,\frac{\partial \phi}{\partial t}(0) + \frac{t^2}{2!}\,\frac{\partial^2 \phi}{\partial t^2}(0) + o\left(t^2\right) \quad \text{as } t \to 0,$$

that is,

$$u(x + t\xi) = u(x) + \langle \nabla u(x), y - x \rangle + \frac{t^2}{2!} \sum_{i,j=1}^{n} \frac{\partial^2 u}{\partial x_i \partial x_j}(x)\xi_i \xi_j + o\left(t^2\right) \quad \text{as } t \to 0.$$

By inequality (8.31) with $y := x + t\xi$, we have, for any given vector $\xi \in \Sigma_{n-1}$,

$$\sum_{i,j=1}^{n} \frac{\partial^2 u}{\partial x_i \partial x_j}(x)\xi_i\,\xi_j \leq 0 \quad \text{for almost all } x \in \mathcal{W}. \tag{8.32}$$

Let $\{\xi^{(n)}\}$ be a *countable* dense subset of Σ_{n-1}. By assertion (8.32), for every neighborhood $U(x_0)$ of x_0 we can find a subset $\widetilde{M}$ of $U(x_0)$, with $|\widetilde{M}| > 0$, such that

$$\sum_{i,j=1}^{n} \frac{\partial^2 u}{\partial x_i \partial x_j}(x)\xi_i^{(n)}\,\xi_j^{(n)} \leq 0 \quad \text{for all } x \in \widetilde{M} \text{ and all } \xi^{(n)} \in \Sigma_{n-1}.$$

By passing to the limit in this inequality, we obtain that

$$\sum_{i,j=1}^{n} \frac{\partial^2 u}{\partial x_i \partial x_j}(x)\xi_i\xi_j \leq 0 \quad \text{for all } x \in \widetilde{M} \text{ and all } \xi \in \Sigma_{n-1}.$$

Summing up, we have proved ([9, Proposition 1]) that, for every neighborhood $U(x_0)$ of x_0 there exists a subset $\widetilde{M}$ of $U(x_0)$, with $|\widetilde{M}| > 0$, such that the Hessian matrix

$$\left(\frac{\partial^2 u}{\partial x_i \partial x_j}(x)\right)$$

is negative semi-definite for all $x \in \widetilde{M}$.

On the other hand, condition (8.2) implies that the matrix $(a^{ij}(x))$ is positive definite for almost all $x \in \Omega$.

Therefore, for every neighborhood $U(x_0)$ of x_0 we can find a subset M of $U(x_0)$, with $|M| > 0$, such that

$$\sum_{i,j=1}^{n} a^{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j}(x) \leq 0 \quad \text{for all } x \in M.$$

The proof of Lemma 8.7 is complete. □

Step (II) (End of Proof of Theorem 8.5): We divide the proof of Theorem 8.5 into Case II-1 and Case II-2.

Case II-1: First, we consider the case where

$$A_0 u := \sum_{i,j=1}^{n} a^{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b^i(x)\frac{\partial u}{\partial x_i},$$
$$W_0 u := A_0 u + Su.$$

(1) We assume that a function $u \in W^{2,p}(\Omega)$, with $n < p < \infty$, satisfies the condition

$$W_0 u(x) = (A_0 + S)\, u(x) \geq 0 \quad \text{for almost all } x \in \Omega, \tag{8.33}$$

and takes a local maximum at an interior point $x_0 \in \Omega$. Then we have, by Proposition 8.6,

$$\lim \operatorname{ess\,inf}_{x \to x_0} \sum_{i,j=1}^{n} a^{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j}(x) \leq 0. \tag{8.34}$$

However, it should be noticed that

$$b^i(x) \in L^\infty(\Omega),$$

$$\frac{\partial u}{\partial x_i}(x_0) = 0, \quad 1 \leq i \leq n.$$

Therefore, we obtain from assertion (8.34) that

$$\lim \operatorname{ess\,inf}_{x \to x_0} A_0 u(x) \leq 0. \tag{8.35}$$

On the other hand, it follows from inequality (8.31) that

$$\begin{aligned} &Su(x) \\ &= \int_{\mathbf{R}^n \setminus \{0\}} \left(u(x+z) - u(x) - \sum_{j=1}^{n} z_j \frac{\partial u}{\partial x_j}(x) \right) K(x,z)\, \mu(dz) \\ &\leq 0 \quad \text{for all } x \in \mathcal{W}. \end{aligned} \tag{8.36}$$

(2) We choose a positive constant γ so large that

$$\gamma > \lambda \left\| b^1 \right\|_{L^\infty(\Omega)}, \tag{8.37}$$

and let, for any $\varepsilon > 0$,

$$u_\varepsilon(x) := u(x) + \varepsilon e^{\gamma x_1}.$$

Then, by conditions (8.2), inequalities (8.33), (8.36) and (8.37) we can find a constant $\eta = \eta(\varepsilon) > 0$ such that

$$\begin{aligned} A_0 u_\varepsilon(x) &= A_0 u(x) + \varepsilon \left[a^{11}(x)\gamma^2 + b^1(x)\gamma \right] e^{\gamma x_1} \\ &\geq -Su(x) + \varepsilon \left[a^{11}(x)\gamma^2 + b^1(x)\gamma \right] e^{\gamma x_1} \\ &\geq \varepsilon \left[a^{11}(x)\gamma^2 + b^1(x)\gamma \right] e^{\gamma x_1} \geq \varepsilon \left[\frac{\gamma^2}{\lambda} - \left\| b^1 \right\|_{L^\infty(\Omega)} \gamma \right] e^{\gamma x_1} \\ &\geq \eta(\varepsilon) > 0 \quad \text{for almost all } x \in \Omega. \end{aligned}$$

In view of inequality (8.35), this proves that the function $u_\varepsilon(x)$ does not take a maximum at any interior point of Ω. Hence we have the inequality

$$u_\varepsilon(x) = u(x) + \varepsilon e^{\gamma x_1} \leq \max_{y \in \partial\Omega} (u(y)) + \varepsilon e^{\gamma y_1}) \quad \text{for all } x \in \overline{\Omega}. \tag{8.38}$$

Therefore, by letting $\varepsilon \downarrow 0$ in inequality (8.38) we obtain that

$$u(x) \leq \max_{y \in \partial\Omega} u(y) \quad \text{for all } x \in \overline{\Omega}.$$

This proves that

$$\max_{\overline{\Omega}} u = \max_{\partial\Omega} u. \tag{8.39}$$

Here it should be emphasized that we do not make any regularity assumption on the boundary $\partial\Omega$.

Case II-2: Secondly, we consider the general case where

$$A = \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b^i(x) \frac{\partial}{\partial x_i} + c(x).$$

Here we recall that $c(x) \in L^\infty(\Omega)$ and $c(x) \leq 0$ for almost all $x \in \Omega$.

We assume that a function $u \in W^{2,p}(\Omega)$, with $n < p < \infty$, satisfies the condition

$$Wu(x) = (A + S)\, u(x) \geq 0 \quad \text{for almost all } x \in \Omega. \tag{8.20}$$

(1) If $u(x) \leq 0$ in Ω, then it follows that

$$\max_{\overline{\Omega}} u \leq 0 \leq \max_{\partial\Omega} u^+.$$

(2) We consider the case where the subdomain

$$\Omega^+ = \{x \in \Omega : u(x) > 0\}$$

is not empty. If we define a differential operator

$$A_0 := A - c(x) = \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b^i(x) \frac{\partial}{\partial x_i},$$

then it follows from condition (8.20) that

$$\begin{aligned}(A_0 + S)\, u(x) &= Wu(x) - c(x)u(x) \geq -c(x)u(x) \\ &\geq 0 \quad \text{for almost all } x \in \Omega^+.\end{aligned}$$

Hence, by applying assertion (8.39) to our situation we obtain (see Figure 8.2 below) that

$$\max_{\overline{\Omega}} u = \max_{\overline{\Omega^+}} u = \max_{\partial\Omega^+} u = \max_{\partial\Omega} u^+.$$

Summing up, we have proved the desired inequality

$$\max_{\overline{\Omega}} u \leq \max_{\partial\Omega} u^+.$$

Now the proof of Theorem 8.5 is complete. □

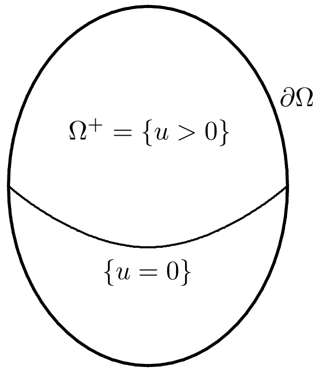

Fig. 8.2. The subdomain $\Omega^+ = \{u > 0\}$ of Ω

8.3 Hopf's Boundary Point Lemma

In this section, we study the inward normal derivative $\partial u/\partial \mathbf{n}(x_0')$ at a boundary point x_0' where the function $u(x)$ takes its non-negative maximum.

The Hopf boundary point lemma reads as follows (cf. Hopf [34], Oleĭnik [54]):

Lemma 8.8 (Hopf). *Assume that a function $u \in W^{2,p}(\Omega)$, with $n < p < \infty$, satisfies the condition*

$$Wu(x) \geq 0 \quad \textit{for almost all } x \in \Omega. \tag{8.20}$$

If $u(x)$ attains a non-negative, strict local maximum at a point x_0' of $\partial\Omega$, then we have the inequality

$$\frac{\partial u}{\partial \mathbf{n}}(x_0') < 0, \tag{8.40}$$

where $\mathbf{n} = (n_1, n_2, \ldots, n_n)$ is the unit inward normal to the boundary $\partial\Omega$ (see Figure 8.3 below).

Proof. By Theorem 8.5, it suffices to consider the case

$$\begin{cases} u(x_0') = m := \max_{x\in\overline{\Omega}} u(x) \geq 0, \\ u(y) < u(x_0') \quad \text{for all } y \in \Omega. \end{cases} \tag{8.41}$$

The proof of inequality (8.40) is divided into three steps.

Step 1: By condition (8.41), we can find an open ball $B(y, r)$ contained in the domain Ω, centered at y, such that (see Figure 8.4)

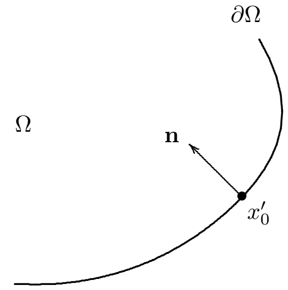

Fig. 8.3. The unit inward normal **n** to the boundary $\partial\Omega$ at x'_0

(a) The point x'_0 is on the boundary $S(y,r) = \{z \in \Omega : |z-y| = r\}$ of $B(y,r)$;
(b) $\mathbf{n} = s(y - x'_0)$ for some $s > 0$.

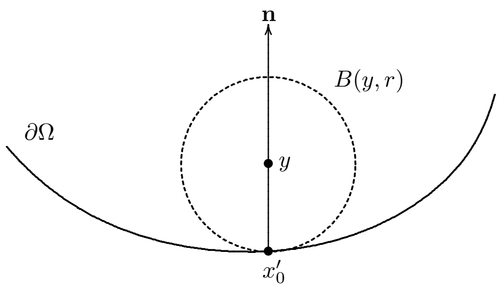

Fig. 8.4. The open ball $B(y,r)$ contained in the domain Ω centered at y

Step 2: Near the boundary point x'_0, we introduce local coordinate systems (x', x_n) such that $x' = (x_1, x_2, \ldots, x_{n-1})$ give local coordinates for the boundary $\partial\Omega$ and that (see Figure 8.5 below)

$$\begin{aligned}
&\Omega = \{(x', x_n) : x_n > 0\}\,, \\
&\partial\Omega = \{(x', x_n) : x_n = 0\}\,, \\
&x'_0 = (0, \ldots, 0, 0)\,, \\
&y = (0, \ldots, 0, r)\,, \\
&|x'_0 - y| = r.
\end{aligned}$$

Now we introduce a function $v(x)$ by the formula

$$v(x) = v(x', x_N) = \exp\left[-\gamma|x-y|^2\right] - \exp\left[-\gamma R^2\right], \tag{8.42}$$

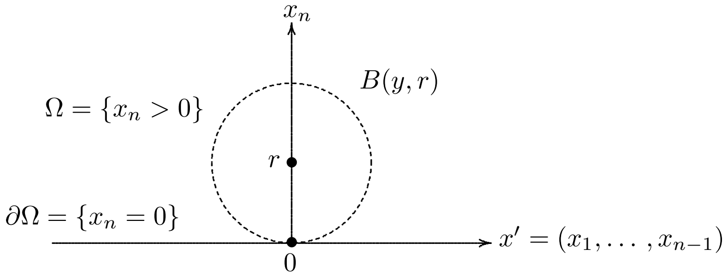

Fig. 8.5. The local coordinate systems (x', x_n) such that $\Omega = \{x_n > 0\}$ and $\partial\Omega = \{x_n = 0\}$

$$0 < R < \frac{1}{2},$$

where γ is a positive constant to be chosen later on. Then it is easy to see that

$$\begin{aligned} Av(x) &= \sum_{i,j=1}^{n} a^{ij}(x) \frac{\partial^2 v}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{n} b^i(x) \frac{\partial v}{\partial x_i}(x) + c(x)v(x) \qquad (8.43) \\ &= \exp\left[-\gamma|x-y|^2\right] \times \left[4\gamma^2 \left(\sum_{i,j=1}^{n} a^{ij}(x)(x_i - y_i)(x_j - y_j)\right)\right. \\ &\quad \left. - 2\gamma \left(\sum_{i=1}^{n} \left(a^{ii}(x) + b^i(x)(x_i - y_i)\right)\right)\right] + c(x)v(x). \end{aligned}$$

However, we have, by formula (8.42) and condition (8.2),

$$v(x) \leq \exp\left[-\gamma|x-y|^2\right] \quad \text{for all } x \in \Omega,$$

and

$$\sum_{i,j=1}^{n} a^{ij}(x)(x_i - y_i)(x_j - y_j) \geq \frac{1}{\lambda}|x-y|^2 \quad \text{for almost all } x \in \Omega.$$

Hence, it follows from formula (8.43) that

$$\begin{aligned} &Av(x) \qquad (8.44) \\ &\geq \exp\left[-\gamma|x-y|^2\right] \\ &\quad \times \left[\frac{4\gamma^2}{\lambda}|x-y|^2 - 2\gamma \left(\sum_{i=1}^{n} \left(a^{ii}(x) + |b^i(x)(x_i - y_i)|\right) + |c(x)|\right)\right] \\ &\qquad \text{for almost all } x \in \Omega. \end{aligned}$$

Moreover, for any $\rho > 0$, we can choose a constant $\gamma = \gamma(\rho)$ so large that we have, for $\rho < |x-y| < R$,

$$\begin{aligned} &\frac{4\gamma^2}{\lambda}|x-y|^2 - 2\gamma\left(\sum_{i=1}^{n}\left(a^{ii}(x) + |b^i(x)(x_i-y_i)|\right) + |c(x)|\right) \qquad (8.45) \\ &\geq \frac{4\rho^2}{\lambda}\gamma^2 - 2\gamma\left(\sum_{i=1}^{n}\left(\left\|a^{ii}\right\|_{L^\infty(\Omega)} + |\left\|b^i\right\|_{L^\infty(\Omega)} R\right) + \|c\|_{L^\infty(\Omega)}\right) \\ &\geq 0 \quad \text{for almost all } x \in \Omega. \end{aligned}$$

Therefore, by combining inequalities (8.44) and (8.45) we obtain that, for any $\rho > 0$ there exists a positive constant $\gamma = \gamma(\rho)$ such that (see Figure 8.6)

$$\begin{aligned} &Av(x) \geq 0 \qquad (8.46) \\ &\text{for almost all } x \in \Gamma_{\rho R} := \{z \in \Omega : \rho < |z-y| < R\}. \end{aligned}$$

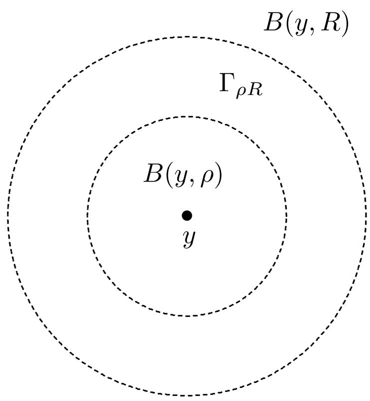

Fig. 8.6. The annular domain $\Gamma_{\rho R} = B(y,R) \setminus \overline{B(y,\rho)}$

On the other hand, by combining inequality (8.14) with assertion (8.16) we can find that, for every small $\eta > 0$ there exists a constant $C_\eta > 0$ such that

$$\begin{aligned} &|Sv(x)| \leq \left(\eta\gamma^2 + C_\eta\gamma\right)\exp\left[-\gamma|x-y|^2\right] \qquad (8.47) \\ &\qquad \text{for almost all } x \in \Gamma_{\rho r} := \{z \in \Omega : \rho < |z-y| < r\}. \end{aligned}$$

Therefore, by taking

$$\eta := \frac{2\rho^2}{\lambda},$$
$$\gamma > \frac{\lambda}{2\rho^2}\left(C + C_\eta\right),$$

we obtain from inequalities (8.46) and (8.47) that

$$\begin{aligned} Wv(x) &= Av(x) + Sv(x) \\ &\geq Av(x) - |Sv(x)| \\ &\geq \left(\frac{2\rho^2}{\lambda}\gamma^2 - \left(C + C_\eta\right)\gamma\right) \exp\left[-\gamma|x-y|^2\right] \\ &= \frac{2\rho^2}{\lambda}\gamma\left(\gamma - \frac{\lambda(C + C_\eta)}{2\rho^2}\right) \exp\left[-\gamma|x-y|^2\right] \\ &> 0 \quad \text{for almost all } x \in \Gamma_{\rho r} := \{z \in \Omega : \rho < |z-y| < r\}. \end{aligned} \tag{8.48}$$

Step 3: Without loss of generality, we may assume that, for $0 < R < 1/2$ sufficiently small,

$$u(x) < u(x_0) \quad \text{in } B(y, R). \tag{8.49}$$

If $\varepsilon > 0$, we let

$$w(x) := u(x) - u(x_0) + \varepsilon v(x).$$

(a) First, we have, by condition (8.49),

$$w(x) = u(x) - u(x_0) + \varepsilon v(x) \leq 0 \quad \text{on } S(y, \rho) = \{z \in \Omega : |z-y| = \rho\},$$

if $\varepsilon > 0$ is chosen sufficiently small.

(b) Secondly, it follows that

$$\begin{aligned} w(x) &= u(x) - u(x_0) + \varepsilon v(x) \\ &\leq 0 \quad \text{on } S(y, r) = \{z \in \Omega : |z-y| = r\}, \end{aligned}$$

since $v(x) = 0$ on $S(y, r)$.

Hence we have, by assertions (a) and (b),

$$w(x) \leq 0 \quad \text{on } \partial\Gamma_{\rho r} = S(y, \rho) \cup S(y, r).$$

On the other hand, by conditions (8.20) and (8.48) it follows that

$$\begin{aligned} Ww(x) &= Wu(x) + \varepsilon Wv(x) - u(x_0)\left(W1\right)(x) \geq \varepsilon Wv(x) - c(x)u(x_0) \\ &> -c(x)u(x_0) \\ &\geq 0 \quad \text{for almost all } x \in \Gamma_{\rho r} = B(y, r) \setminus \overline{B(y, \rho)}. \end{aligned}$$

Therefore, by applying Theorem 8.5 with $\Omega := \Gamma_{\rho r}$ we obtain that

$$w(x) \leq 0 \quad \text{in } \Gamma_{\rho r} = B(y,r) \setminus \overline{B(y,\rho)}. \tag{8.50}$$

(c) On the other hand, we have, by formula (8.38),

$$w(x_0) = \varepsilon v(x_0) = 0. \tag{8.51}$$

Therefore, it follows from assertions (8.50) and (8.51) that

$$\frac{\partial w}{\partial \mathbf{n}}(x_0') = \frac{\partial u}{\partial \mathbf{n}}(x_0') + \varepsilon \frac{\partial v}{\partial \mathbf{n}}(x_0') \leq 0. \tag{8.52}$$

However, we have, by formula (8.42),

$$\frac{\partial v}{\partial \mathbf{n}}(x_0') = 2\gamma r e^{-\gamma r^2} > 0. \tag{8.53}$$

Summing up, we obtain from inequalities (8.52) and (8.53) that

$$\frac{\partial u}{\partial \mathbf{n}}(x_0') \leq -\varepsilon \frac{\partial v}{\partial \mathbf{n}}(x_0') = -2\varepsilon\gamma r e^{-\gamma r^2} < 0.$$

Now the proof of Lemma 8.8 is complete. □

8.4 Strong Maximum Principle

Finally, we can prove the following strong maximum principle for the operator A ([9, Théorème 2]):

Theorem 8.9 (the strong maximum principle)**.** *Assume that a function $u \in W^{2,p}(\Omega)$, with $n < p < \infty$, satisfies the condition*

$$Wu(x) \geq 0 \quad \textit{for almost all } x \in \Omega. \tag{8.20}$$

If $u(x)$ attains a non-negative maximum at an interior point of Ω, then it is a (non-negative) constant *function.*

Proof. Our proof is based on a reduction to absurdity. We let

$$m := \max_{x \in \overline{\Omega}} u(x) \geq 0,$$
$$\mathcal{S} := \{x \in \Omega : u(x) = m\},$$

and assume, to the contrary, that

$$\mathcal{S} \subsetneqq \Omega.$$

Since $\mathcal{S}$ is closed in Ω, we can find a point x_0 of $\mathcal{S}$ and an open ball $B(y,R)$ contained in the set $\Omega \setminus \mathcal{S}$, centered at y, such that (see Figure 8.7)

(a) $B(y,R) \subset \Omega \setminus \mathcal{S}$;
(b) x_0 is on the boundary $S(y,R) = \{z \in \Omega : |z-y| = R\}$ of $B(y,R)$.

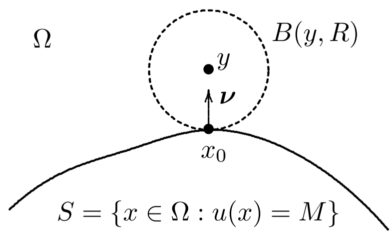

Fig. 8.7. The open ball $B(y,R)$ contained in the set $\Omega \setminus S$, centered at y

By applying Hopf's boundary point lemma (Lemma 8.8) with $\Omega := B(y,R)$, we obtain that

$$\sum_{i=1}^{n} \nu_i \, \frac{\partial u}{\partial x_i}(x_0) < 0, \tag{8.54}$$

where

$$\boldsymbol{\nu} = \frac{y - x_0}{|y - x_0|}.$$

However, since $u(x_0) = m$ for some interior point $x_0 \in \Omega$, it follows that

$$\frac{\partial u}{\partial x_i}(x_0) = 0, \quad 1 \leq i \leq n.$$

Hence we have the assertion

$$\sum_{i=1}^{n} \nu_i \, \frac{\partial u}{\partial x_i}(x_0) = 0.$$

This contradicts inequality (8.54).

The proof of Theorem 8.9 is complete. □

8.5 Notes and Comments

The results of this chapter are adapted from Bony [9], Bony–Courrège–Priouret [11], Troianiello [93] and also Taira [73], [79], [81] and [82].