MATH 201C Lecture Notes
ANALYSIS

Steve Shkoller
Department of Mathematics
University of California
Davis, CA 95616

May 26, 2017

Contents

Chapter 1

A Short Introduction to L^p Spaces

1.1 Notation

We will usually use Ω to denote an open and smooth domain in $\mathbb{R}^{\mathrm{n}}$, for $\mathrm{n} = 1, 2, 3, \ldots$ In this chapter on L^p spaces, we will sometimes use Ω to denote a more general measure space, but the reader can usually think of Ω as an open subset of Euclidean space. The *support* of a function f is the closure of the set $\left\{x \in \Omega \,\middle|\, f(x) \neq 0\right\}$.

Definition 1.1 (Continuous functions and compact support). For $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$, we let $\mathscr{C}^0(\Omega)$ denote the collection of continuous functions on Ω, and we denote by $\mathscr{C}_c^0(\Omega)$ the collection of those functions in $\mathscr{C}^0(\Omega)$ with compact support contained in Ω.

Definition 1.2 (Uniformly continuous functions). For $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ we set

$$\mathscr{C}^0(\bar{\Omega}) := \left\{u : \Omega \to \mathbb{R} \,\middle|\, u \text{ is uniformly continuous}\right\}.$$

For integers $k \geqslant 0$, we let $\mathscr{C}^k(\bar{\Omega})$ denote the collection of functions possessing partial derivatives to all orders up to k which are uniformly continuous on $\bar{\Omega}$. We use $\mathscr{C}^k_{\mathrm{loc}}(\Omega)$ to denote the functions in $\mathscr{C}^k(\bar{B})$ for all bounded balls $\bar{B}$ contained in Ω.

Definition 1.3 (Bounded continuous functions). For $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ we set

$$\mathscr{C}_b(\Omega) := \left\{u : \Omega \to \mathbb{R} \,\middle|\, u \text{ is bounded and continuous}\right\},$$

with norm $\|u\|_{\mathscr{C}_b(\Omega)} = \max_{x \in \Omega} |u(x)|$. For integers $k \geqslant 0$, we let $\mathscr{C}_b^k(\Omega)$ denote the collection of functions possessing partial derivatives to all orders up to k which are bounded and continuous on Ω.

Remark 1.4. In the case that $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ is bounded, $\mathscr{C}^0(\bar{\Omega}) \subseteq \mathscr{C}_b(\Omega)$ and $\mathscr{C}^0(\bar{\Omega})$ is a Banach space with norm $\|u\|_{\mathscr{C}(\bar{\Omega})} = \max_{x\in\bar{\Omega}} |u(x)|$.[1]

$\mathscr{C}^k(\Omega)$ is the space of functions which are k times differentiable in Ω for integers $k \geqslant 0$.

$\mathscr{C}^0(\Omega)$ then coincides with $\mathscr{C}(\Omega)$, the space of continuous functions on Ω.

$\mathscr{C}^\infty(\Omega) = \bigcap_{k\geqslant 0} \mathscr{C}^k(\Omega)$.

$\mathrm{spt}(f)$ denotes the support of a function f, and is the closure of the set $\left\{x \in \Omega \,\middle|\, f(x) \neq 0\right\}$.

$\mathscr{C}_c(\Omega) = \left\{u \in \mathscr{C}(\Omega) \,\middle|\, \mathrm{spt}\, u \text{ compact in } \Omega\right\}$.

$\mathscr{C}_c^k(\Omega) = \mathscr{C}^k(\Omega) \cap \mathscr{C}_c(\Omega)$.
$\mathscr{C}_c^\infty(\Omega) = \mathscr{C}^\infty(\Omega) \cap \mathscr{C}_c(\Omega)$. We will also use $\mathscr{D}(\Omega)$ to denote this space, which is known as the *space of test functions* in the theory of distributions.

1.2 Lebesgue Measure and Lebesgue Integral

Let $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ denote an open and smooth subset. The domain Ω is called smooth whenever its boundary $\partial\Omega$ is a smooth $(\mathrm{n}-1)$-dimensional hypersurface.

The theory of L^p spaces is founded upon the so-called Lebesgue integral (which requires some basic knowledge of the Lebesgue measure). We define the set $L^p(\Omega)$ as

$$L^p(\Omega) \equiv \left\{ f : \Omega \to \mathbb{R} \text{ measurable} \,\middle|\, \int_\Omega |f(x)|^p dx < \infty \right\}$$

where the integral is interpreted in the sense of Lebesgue. [2] We will assume that all

[1]Suppose that $\Omega = (0,1)$ and let $u(x) = \sin(1/x)$. Then clearly $u \in \mathscr{C}_b(\Omega)$ but u is not uniformly continuous, as the limit $\lim_{x\to 0^+} \sin(1/x)$ does not exist.

[2]The following theorem is usually presented in an undergraduate course in analysis:

Theorem 1.5. *Let $\Omega \subseteq$ be a domain with positive measure (length, area, volume, etc.). Then, a bounded function is Riemann integrable over Ω if and only it is continuous a.e. in Ω. The notation "a.e." denotes almost everywhere, which means up to a set of measure zero,*

In a first course on measure theory, the following theorem is established:

Theorem 1.6. *If f is non-negative Riemann (improper) integrable over Ω, then f is measurable and the Riemann (improper) integral of f over Ω is the same as the Lebesgue integral.*

Therefore, the Lebesgue integral is a generalization of the Riemann integral. See Appendix A.2.2 for a review of this material.

functions and sets are Lebesgue measurable. The Lebesgue measure is often denoted by μ so that $\mu(\Omega)$ denotes the length if n = 1, the area if n = 2, the volume if n = 3, and so on. We shall also use the notation $|\Omega|$ to mean $\mu(\Omega)$.

1.2.1 The three pillars of analysis

A function $f : \Omega \to \mathbb{R}$ is Lebesgue integrable if $\int_\Omega f(x)dx < \infty$. (We shall often write that f is integrable to mean that $f : \Omega \to \mathbb{R}$ is Lebesgue integrable.)

The following three theorems will be used throughout the course.

Theorem 1.7 (Monotone Convergence Theorem). *Let $f_k : \Omega \to \mathbb{R} \cup \{+\infty\}$ denote a sequence of non-negative functions, and suppose that the sequence f_k is monotonically increasing; that is,*

$$f_1 \leqslant f_2 \leqslant f_3 \leqslant \cdots$$

Then

$$\lim_{k\to\infty} \int_\Omega f_k(x)dx = \int_\Omega \lim_{k\to\infty} f_k(x)dx\,.$$

Theorem 1.8 (Fatou's Lemma). *Suppose the sequence $f_k : \Omega \to \mathbb{R} \cup \{+\infty\}$ and $f_k \geqslant 0$. Then*

$$\int_\Omega \liminf_{k\to\infty} f_k(x)dx \leqslant \liminf_{k\to\infty} \int_\Omega f_k(x)dx\,.$$

Example 1.9. Consider $\Omega = (0,1) \subseteq \mathbb{R}$ and suppose that $f_k = k\mathbf{1}_{(0,1/k)}$. Then $\int_0^1 f_k(x)dx = 1$ for all $k \in \mathbb{N}$, but $\int_0^1 \liminf_{k\to\infty} f_k(x)dx = 0$.

Theorem 1.10 (Dominated Convergence Theorem). *Suppose the sequence $f_k : \Omega \to \mathbb{R}$, $f_k \to f$ almost everywhere (with respect to Lebesgue measure), and furthermore, $|f_k| \leqslant g \in L^1(\Omega)$. Then $f \in L^1(\Omega)$ and*

$$\lim_{k\to\infty} \int_\Omega f_k(x)dx = \int_\Omega f(x)dx\,.$$

Equivalently, $f_k \to f$ in $L^1(\Omega)$ so that $\lim_{k\to\infty} \|f_k - f\|_{L^1(\Omega)} = 0$.

In the exercises, you will be asked to prove that the Monotone Convergence Theorem implies Fatou's Lemma which, in turn, implies the Dominated Convergence Theorem. See Appendix A.2.2 for a review of basic integration theory.

1.2.2 Iterated integrals

Let $\mathrm{I}_1 \subseteq \mathbb{R}^{\mathrm{n}}$ and $\mathrm{I}_2 \subseteq \mathbb{R}^{\mathrm{m}}$ denote open subsets.

THEOREM 1.11 (Fubini). *Let $f : \mathrm{I}_1 \times \mathrm{I}_2 \to \mathbb{R}$ be an integrable function. Then both iterated integrals exist and*

$$\int_{\mathrm{I}_1\times \mathrm{I}_1} f = \int_{\mathrm{I}_2}\Big(\int_{\mathrm{I}_1} f(x,y)\,dx\Big)dy = \int_{\mathrm{I}_1}\Big(\int_{\mathrm{I}_2} f(x,y)\,dy\Big)dx\,.$$

The existence of the iterated integrals is by no means enough to ensure that the function is integrable over the product space. As an example, let $\mathrm{I}_1 = \mathrm{I}_2 = [0,1]$. Set

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

Then a standard computation shows that

$$\int_0^1\int_0^1 f(x,y)dxdy = -\frac{\pi}{4}, \qquad \int_0^1\int_0^1 f(x,y)dydx = \frac{\pi}{4}\,.$$

Fubini's theorem shows, of course, that f is not integrable over $[0,1]^2$.

When the integrand f is non-negative (and whether f is integrable or not), one can compute the integral of f over a product space using iterated integrals; this is due to Tonelli's theorem which we state as follows:

THEOREM 1.12 (Tonelli). *Let $f : \mathrm{I}_1 \times \mathrm{I}_2 \to \mathbb{R}$ be non-negative and measurable. Then*

$$\int_{\mathrm{I}_1\times \mathrm{I}_2} f = \int_{\mathrm{I}_2}\Big(\int_{\mathrm{I}_1} f(x,y)\,dx\Big)dy = \int_{\mathrm{I}_1}\Big(\int_{\mathrm{I}_2} f(x,y)\,dy\Big)dx\,.$$

There is a converse to Fubini's theorem; however, according to which the existence of one of the iterated integrals is sufficient for the integrability of the function over the product space. This converse statement is a direct consequence of the Fubini and Tonelli theorems, and is stated as the following

COROLLARY 1.13. *Let $f : \mathrm{I}_1\times\mathrm{I}_2 \to \mathbb{R}$. If one of the iterated integrals $\int_{\mathrm{I}_1}\Big(\int_{\mathrm{I}_2} |f(x,y)|dy\Big)dx$ or $\int_{\mathrm{I}_2}\Big(\int_{\mathrm{I}_1} |f(x,y)|dx\Big)dy$ exists, then the function f is integrable on the product space $\mathrm{I}_1 \times \mathrm{I}_2$, and hence, the other iterated integral exists and*

$$\int_{\mathrm{I}_1\times \mathrm{I}_2} f = \int_{\mathrm{I}_2}\Big(\int_{\mathrm{I}_1} f(x,y)\,dx\Big)dy = \int_{\mathrm{I}_1}\Big(\int_{\mathrm{I}_2} f(x,y)\,dy\Big)dx\,.$$

1.3 L^p Spaces

Now, we turn to the definition and basic properties of L^p spaces.

1.3.1 Definitions and basic properties

Definition 1.14. Let $0 < p < \infty$ and let Ω denote and open subset of $\mathbb{R}^n$. If $f : \Omega \to \mathbb{R}$ is a measurable function, then we define

$$\|f\|_{L^p(\Omega)} := \left(\int_\Omega |f|^p dx \right)^{\frac{1}{p}} \quad \text{and} \quad \|f\|_{L^\infty(\Omega)} := \text{ess sup}_{x\in\Omega} |f(x)| .$$

Note that $\|f\|_{L^p(\Omega)}$ may take the value ∞. (Unless stated otherwise, we will assume that all functions under consideration are measurable with respect to Lebesgue measure.)

Definition 1.15. The space $L^p(\Omega)$ is the set

$$L^p(\Omega) = \left\{ f : \Omega \to \mathbb{R} \,\middle|\, \|f\|_{L^p(\Omega)} < \infty \right\} .$$

The space $L^p(\Omega)$ satisfies the following vector space properties:

1. For each $\alpha \in \mathbb{R}$, if $f \in L^p(\Omega)$ then $\alpha f \in L^p(\Omega)$;

2. If $f, g \in L^p(\Omega)$, then
$$|f + g|^p \leqslant 2^{p-1}(|f|^p + |g|^p) ,$$
so that $f + g \in L^p(\Omega)$.

3. The triangle inequality is valid if $p \geqslant 1$.

Pehaps the most interesting cases are $p = 1, 2, \infty$, while all of the L^p spaces arise often in *nonlinear* estimates, and can play an important role in scaling arguments.

Definition 1.16. The space ℓ^p, called "little L^p", will be useful when we introduce Sobolev spaces on the torus and the Fourier series. For $1 \leqslant p < \infty$, we set

$$\ell^p = \left\{ \{x_n\}_{n\in\mathbb{Z}} \,\middle|\, \sum_{n=-\infty}^{\infty} |x_n|^p < \infty \right\} ,$$

where $\mathbb{Z}$ denotes the integers.

1.3.2 Basic inequalities

Convexity is fundamental to L^p spaces for $p \in [1, \infty)$.

LEMMA 1.17. *For $\lambda \in (0,1)$, $x^\lambda \leqslant (1-\lambda) + \lambda x$.*

Proof. Set $f(x) = (1-\lambda) + \lambda x - x^\lambda$; hence, $f'(x) = \lambda - \lambda x^{\lambda-1} = 0$ if and only if $\lambda(1 - x^{\lambda-1}) = 0$ so that $x = 1$ is the critical point of f. In particular, the minimum occurs at $x = 1$ with value

$$f(1) = 0 \leqslant (1-\lambda) + \lambda x - x^\lambda \,. \qquad \square$$

LEMMA 1.18. *For $a, b \geqslant 0$ and $\lambda \in (0,1)$, $a^\lambda b^{1-\lambda} \leqslant \lambda a + (1-\lambda) b$ with equality if $a = b$.*

Proof. If either $a = 0$ or $b = 0$, then this is trivially true, so assume that $a, b > 0$. Set $x = a/b$, and apply Lemma 1.17 to obtain the desired inequality. $\square$

THEOREM 1.19 (Hölder's inequality)**.** *Suppose that $1 \leqslant p \leqslant \infty$ and $1 < q < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$, then $fg \in L^1(\Omega)$. Moreover,*

$$\|fg\|_{L^1(\Omega)} \leqslant \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)} \,.$$

Note that if $p = q = 2$, then this is the Cauchy-Schwarz inequality since $|(f,g)_{L^2}| \leqslant \|fg\|_{L^1}$.

Proof. We use Lemma 1.18. Let $\lambda = \frac{1}{p}$ and set

$$a = \frac{|f|^p}{\|f\|^p_{L^p(\Omega)}} \,, \quad \text{and} \quad b = \frac{|g|^q}{\|g\|^q_{L^q(\Omega)}}$$

for all $x \in \Omega$. Then $a^\lambda b^{1-\lambda} = a^{1/p} b^{1-1/p} = a^{1/p} b^{1/q}$ so that

$$\frac{|f| \cdot |g|}{\|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)}} \leqslant \frac{1}{p} \frac{|f|^p}{\|f\|^p_{L^p(\Omega)}} + \frac{1}{q} \frac{|g|^q}{\|g\|^q_{L^q(\Omega)}} \,.$$

Integrating this inequality yields

$$\int_\Omega \frac{|f| \cdot |g|}{\|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)}} dx \leqslant \int_\Omega \Big(\frac{1}{p} \frac{|f|^p}{\|f\|^p_{L^p(\Omega)}} + \frac{1}{q} \frac{|g|^q}{\|g\|^q_{L^q(\Omega)}} \Big) dx = \frac{1}{p} + \frac{1}{q} = 1 \,. \qquad \square$$

DEFINITION 1.20. The exponent $q = \frac{p}{p-1}$ $\big($or $\frac{1}{q} = 1 - \frac{1}{p}\big)$ is called the *conjugate exponent* of p.

Lemma 1.21 (Interpolation inequality)**.** *Let* $1 \leqslant r \leqslant s \leqslant t \leqslant \infty$, *and suppose that* $u \in L^r(\Omega) \cap L^t(\Omega)$. *Then for* $\dfrac{1}{s} = \dfrac{a}{r} + \dfrac{1-a}{t}$

$$\|u\|_{L^s(\Omega)} \leqslant \|u\|^a_{L^r(\Omega)} \|u\|^{1-a}_{L^t(\Omega)} .$$

Proof. By Hölder's inequality,

$$\begin{aligned}\int_\Omega |u|^s dx &= \int_\Omega |u|^{as} |u|^{(1-a)s} dx \\ &\leqslant \left(\int_\Omega |u|^{as\frac{r}{as}} dx\right)^{\frac{as}{r}} \left(\int_\Omega |u|^{(1-a)s\frac{t}{(1-a)s}} dx\right)^{\frac{(1-a)s}{t}} = \|u\|^{as}_{L^r(\Omega)} \|u\|^{(1-a)s}_{L^t(\Omega)}.\end{aligned}$$ □

Theorem 1.22 (Minkowski's inequality)**.** *If* $1 \leqslant p \leqslant \infty$ *and* $f, g \in L^p(\Omega)$ *then*

$$\|f+g\|_{L^p(\Omega)} \leqslant \|f\|_{L^p(\Omega)} + \|g\|_{L^p(\Omega)} .$$

Proof. If $f + g = 0$ a.e., then the statement is trivial. Assume that $f + g \neq 0$ a.e. Consider the equality

$$|f+g|^p = |f+g| \cdot |f+g|^{p-1} \leqslant \big(|f| + |g|\big)|f+g|^{p-1} ,$$

and integrate over Ω to find that

$$\begin{aligned}\int_\Omega |f+g|^p dx &\leqslant \int_\Omega \left[(|f|+|g|)|f+g|^{p-1}\right] dx \\ &\overset{\text{Hölder's}}{\leqslant} \big(\|f\|_{L^p(\Omega)} + \|g\|_{L^p(\Omega)}\big) \left\||f+g|^{p-1}\right\|_{L^q(\Omega)} .\end{aligned}$$

Since $q = \dfrac{p}{p-1}$,

$$\left\||f+g|^{p-1}\right\|_{L^q(\Omega)} = \left(\int_\Omega |f+g|^p dx\right)^{\frac{1}{q}} ,$$

from which it follows that

$$\left(\int_\Omega |f+g|^p dx\right)^{1-\frac{1}{q}} \leqslant \|f\|_{L^p(\Omega)} + \|g\|_{L^q(\Omega)} ,$$

which completes the proof, since $\dfrac{1}{p} = 1 - \dfrac{1}{q}$. □

Corollary 1.23. *For* $1 \leqslant p \leqslant \infty$, $L^p(\Omega)$ *is a normed linear space.*

Example 1.24 (Concavity)**.** Let Ω denote a subset of $\mathbb{R}^n$ whose Lebesgue measure is equal to one. If $f \in L^1(\Omega)$ satisfies $f(x) \geqslant M > 0$ for almost all $x \in \Omega$, then $\log(f) \in L^1(\Omega)$ and satisfies

$$\int_\Omega \log f dx \leqslant \log\Big(\int_\Omega f dx\Big).$$

To see this, consider the function $g(t) = t-1-\log t$ for $t > 0$. Compute $g'(t) = 1-\frac{1}{t} = 0$ so $t = 1$ is a minimum (since $g''(1) > 0$). Thus, $\log t \leqslant t-1$ and letting $t \mapsto \frac{1}{t}$ we see that

$$1-\frac{1}{t} \leqslant \log t \leqslant t-1\,. \tag{1.1}$$

Since $\log x$ is continuous and f is measurable, then $\log f$ is measurable for $f > 0$. Let $t = \frac{f(x)}{\|f\|_{L^1(\Omega)}}$ in (1.1) to find that

$$1-\frac{\|f\|_{L^1(\Omega)}}{f(x)} \leqslant \log f(x) - \log\|f\|_{L^1(\Omega)} \leqslant \frac{f(x)}{\|f\|_{L^1(\Omega)}} - 1\,. \tag{1.2}$$

Since $g(x) \leqslant \log f(x) \leqslant h(x)$ for two integrable functions g and h, it follows that $\log f(x)$ is integrable. Next, integrate (1.2) to finish the proof, as $\int_\Omega \Big(\frac{f(x)}{\|f\|_{L^1(\Omega)}}-1\Big)dx = 0$.

1.3.3 The space $(L^p(\Omega), \|\cdot\|_{L^p(\Omega)})$ is complete

Recall the a normed linear space is a Banach space if every Cauchy sequence has a limit in that space; furthermore, recall that a sequence $x_k \to x$ in $\mathbb{B}$ if $\lim_{k\to\infty} \|x_k - x\|_{\mathbb{B}} = 0$.

The proof of completeness makes use of the following two lemmas which are restatements of the Monotone Convergence Theorem and the Dominated Convergence Theorem, respectively.

Lemma 1.25 (MCT)**.** *If $f_k \in L^1(\Omega)$, $0 \leqslant f_1(x) \leqslant f_2(x) \leqslant \cdots$, and $\|f_k\|_{L^1(\Omega)} \leqslant C < \infty$, then $\lim_{k\to\infty} f_k(x) = f(x)$ with $f \in L^1(\Omega)$ and $\|f_k - f\|_{L^1(\Omega)} \to 0$ as $k \to 0$.*

Lemma 1.26 (DCT)**.** *If $f_k \in L^1(\Omega)$, $\lim_{k\to\infty} f_k(x) = f(x)$ a.e., and if $\exists\ g \in L^1(\Omega)$ such that $|f_k(x)| \leqslant |g(x)|$ a.e. for all n, then $f \in L^1(\Omega)$ and $\|f_k - f\|_{L^1(\Omega)} \to 0$.*

Proof. Apply the Dominated Convergene Theorem to the sequence $h_k = |f_k - f| \to 0$ a.e., and note that $|h_k| \leqslant 2g$. □

Theorem 1.27. *If $1 \leqslant p < \infty$, then $L^p(\Omega)$ is a Banach space.*

Proof. **Step 1. The Cauchy sequence.** Let $\{f_k\}_{k=1}^\infty$ denote a Cauchy sequence in $L^p(\Omega)$, and assume without loss of generality (by extracting a subsequence if necessary) that $\|f_{k+1} - f_k\|_{L^p(\Omega)} \leqslant 2^{-k}$.

Step 2. Conversion to a convergent monotone sequence. Define the sequence $\{g_k\}_{k=1}^\infty$ as

$$g_1 = 0, \quad g_k = |f_1| + |f_2 - f_1| + \cdots + |f_k - f_{k-1}| \quad \text{for} \quad k \geqslant 2\,.$$

It follows that

$$0 \leqslant g_1 \leqslant g_2 \leqslant \cdots \leqslant g_k \leqslant \cdots$$

so that g_k is a monotonically increasing sequence. Furthermore, $\{g_k\}_{k=1}^\infty$ is uniformly bounded in $L^p(\Omega)$ as

$$\int_\Omega g_k^p dx = \|g_k\|^p_{L^p(\Omega)} \leqslant \Big(\|f_1\|_{L^p(\Omega)} + \sum_{i=2}^\infty \|f_i - f_{i-1}\|_{L^p(\Omega)}\Big)^p \leqslant \big(\|f_1\|_{L^p(\Omega)} + 1\big)^p\,;$$

thus, by the Monotone Convergence Theorem, $g_k^p \nearrow g^p$ a.e., $g \in L^p(\Omega)$, and $g_k \leqslant g$ a.e.

Step 3. Pointwise convergence of $\{f_k\}_{k=1}^\infty$. For all $k \geqslant 1$,

$$\begin{aligned} |f_{k+\ell} - f_\ell| &= |f_{k+\ell} - f_{k+\ell-1} + f_{k+\ell-1} + \cdots - f_{\ell+1} + f_{\ell+1} - f_\ell| \\ &\leqslant \sum_{i=\ell+1}^{k+\ell} |f_i - f_{i-1}| = g_{k+\ell} - g_\ell \to 0 \text{ a.e. as } \ell \to \infty. \end{aligned}$$

Therefore, $f_k \to f$ a.e. Since

$$|f_k| \leqslant |f_1| + \sum_{i=2}^k |f_i - f_{i-1}| \leqslant g_k \leqslant g \text{ for all } k \in \mathbb{N}\,,$$

it follows that $|f| \leqslant g$ a.e. Hence, $|f_k|^p \leqslant g^p$, $|f|^p \leqslant g^p$, and $|f - f_k|^p \leqslant 2g^p$, and by the Dominated Convergence Theorem,

$$\lim_{k\to\infty} \int_\Omega |f - f_k|^p dx = \int_\Omega \lim_{k\to\infty} |f - f_k|^p dx = 0\,.$$

□

1.3.4 Convergence criteria for L^p functions

If $\{f_k\}_{k=1}^\infty$ is a sequence in $L^p(\Omega)$ which converges to f in $L^p(\Omega)$, then there exists a subsequence $\big\{f_{k_j}\big\}_{j=1}^\infty$ such that $f_{k_j} \to f$ a.e., but it is in general *not true* that the entire sequence itself will converge pointwise a.e. to the limit f, without some further conditions holding.

EXAMPLE 1.28. Let $\Omega = [0,1]$, and consider the subintervals

$$[0,\tfrac{1}{2}]\,,[\tfrac{1}{2},1]\,,[0,\tfrac{1}{3}]\,,[\tfrac{1}{3},\tfrac{2}{3}]\,,[\tfrac{2}{3},1]\,,[0,\tfrac{1}{4}]\,,[\tfrac{1}{4},\tfrac{2}{4}]\,,[\tfrac{2}{4},\tfrac{3}{4}]\,,[\tfrac{3}{4},1]\,,[0,\tfrac{1}{5}]\,,\cdots$$

Let f_k denote the indicator function of the k^{th} interval of the above sequence. Then $\|f_k\|_{L^p(\Omega)} \to 0$, but $f_k(x)$ does not converge for any $x \in [0,1]$.

EXAMPLE 1.29. Set $\Omega = [0,1]$, and for $k \in \mathbb{N}$, set $f_k = k\mathbf{1}_{[0,\frac{1}{k}]}$. Then $f_k \to 0$ a.e. as $k \to \infty$, but $\|f_k\|_{L^1(\Omega)} = 1$; thus, $f_k \to 0$ pointwise, but not in the L^1 sense.

THEOREM 1.30. *For $1 \leqslant p < \infty$, suppose that $\{f_k\}_{k=1}^{\infty} \subseteq L^p(\Omega)$ and that $f_k \to f$ a.e. If $\lim\limits_{k\to\infty} \|f_k\|_{L^p(\Omega)} = \|f\|_{L^p(\Omega)}$, then $f_k \to f$ in $L^p(\Omega)$.*

Proof. Given $a,b \geqslant 0$, convexity implies that $\big(\frac{a+b}{2}\big)^p \leqslant \frac{1}{2}(a^p + b^p)$ so that $(a+b)^p \leqslant 2^{p-1}(a^p + b^p)$, and hence $|a-b|^p \leqslant 2^{p-1}(|a|^p + |b|^p)$. Set $a = f_k$ and $b = f$ to obtain the inequality

$$0 \leqslant 2^{p-1}\left(|f_k|^p + |f|^p\right) - |f_k - f|^p\,.$$

Since $f_k(x) \to f(x)$ a.e.,

$$2^p \int_\Omega |f|^p dx = \int_\Omega \lim_{k\to\infty} \Big(2^{p-1}(|f_k|^p + |f|^p) - |f_k - f|^p\Big) dx\,.$$

Thus, Fatou's lemma asserts that

$$\begin{aligned}
2^p \int_\Omega |f|^p dx &\leqslant \liminf_{k\to\infty} \int_\Omega \big(2^{p-1}(|f_k|^p + |f|^p) - |f_k - f|^p\big)\, dx \\
&= 2^{p-1}\int_\Omega |f|^p dx + 2^{p-1}\lim_{k\to\infty}\int_\Omega |f_k|^p dx + \liminf_{k\to\infty}\Big(-\int_\Omega |f_k - f|^p dx\Big) \\
&= 2^p \int_\Omega |f|^p dx - \limsup_{k\to\infty}\int_\Omega |f_k - f|^p dx\,.
\end{aligned}$$

As $\int_\Omega |f|^p dx < \infty$, the last inequality shows that $\limsup\limits_{k\to\infty} \int_\Omega |f_k - f|^p dx \leqslant 0$. It follows that $\limsup\limits_{k\to\infty} \int_\Omega |f_k - f|^p dx = \liminf\limits_{k\to\infty} \int_\Omega |f_k - f|^p dx = 0$, so that $\lim\limits_{k\to\infty} \int_\Omega |f_k - f|^p dx = 0$. □

1.3.5 The space $L^\infty(\Omega)$

DEFINITION 1.31. With $\|f\|_{L^\infty(\Omega)} = \inf\big\{M \geqslant 0 \,\big|\, |f(x)| \leqslant M \ \text{ a.e.}\big\}$, we set

$$L^\infty(\Omega) = \big\{f : \Omega \to \mathbb{R} \,\big|\, \|f\|_{L^\infty(\Omega)} < \infty\big\}\,.$$

Theorem 1.32. *$(L^\infty(\Omega), \|\cdot\|_{L^\infty(\Omega)})$ is a Banach space.*

Proof. Let $\{f_k\}_{k=1}^\infty$ be a Cauchy sequence in $L^\infty(\Omega)$. It follows that $|f_k - f_\ell| \leqslant \|f_k - f_\ell\|_{L^\infty(\Omega)}$ a.e. and hence $f_k \to f$ a.e., where f is measurable and essentially bounded.

Choose $\epsilon > 0$ and $N(\epsilon)$ such that $\|f_k - f_\ell\|_{L^\infty(\Omega)} < \epsilon$ for all $k, \ell \geqslant N(\epsilon)$. Since $|f_k(x) - f(x)| = \lim_{\ell\to\infty} |f_k(x) - f_\ell(x)| \leqslant \epsilon$ holds for almost every $x \in \Omega$, it follows that $\|f_k - f\|_{L^\infty(\Omega)} \leqslant \epsilon$ for $k \geqslant N(\epsilon)$, so that $\|f_k - f\|_{L^\infty(\Omega)} \to 0$. ◻

1.3.6 Comparison

Remark 1.33. In general, there is no relation of the type $L^p(\Omega) \subseteq L^q(\Omega)$. For example, suppose that $\Omega = (0,1)$ and set $f(x) = x^{-\frac{1}{2}}$. Then $f \in L^1(0,1)$, but $f \notin L^2(0,1)$. On the other hand, if $\Omega = (1,\infty)$ and $f(x) = x^{-1}$, then $f \in L^2(1,\infty)$, but $f \notin L^1(1,\infty)$.

Lemma 1.34 (L^p comparisons)**.** *If $1 \leqslant p < q < r \leqslant \infty$, then* (a) $L^p(\Omega) \cap L^r(\Omega) \subseteq L^q(\Omega)$, *and* (b) $L^q(\Omega) \subseteq L^p(\Omega) + L^r(\Omega)$.

Proof. We begin with (b). Suppose that $f \in L^q$, define the set $E = \{x \in \Omega \ : \ |f(x)| \geqslant 1\}$, and write f as

$$f = \underbrace{f\mathbf{1}_E}_{\equiv g} + \underbrace{f\mathbf{1}_{E^\complement}}_{\equiv h} \ .$$

Our goal is to show that $g \in L^p(\Omega)$ and $h \in L^r(\Omega)$. Since $|g|^p = |f|^p\mathbf{1}_E \leqslant |f|^q\mathbf{1}_E$ and $|h|^r = |f|^r\mathbf{1}_{E^\complement} \leqslant |f|^q\mathbf{1}_{E^\complement}$, assertion (b) is proven.

For (a), we prove Lyapunov's inequality:

$$\int_\Omega |f|^q dx \leqslant \left(\int_\Omega |f|^p dx\right)^{\frac{r-q}{r-p}} \left(\int_\Omega |f|^r dx\right)^{\frac{q-p}{r-p}} .$$

Hölder's inequality can be stated as follows: $\int_\Omega g^s h^t dx \leqslant \left(\int_\Omega g dx\right)^s \left(\int_\Omega h dx\right)^t$ for $s + t = 1$. We thus set $g = |f|^p$, $h = |f|^r$ with $s = \frac{r-q}{r-p}$ and $t = \frac{q-p}{r-p}$. Notice, then, that $g^s h^t = |f|^q$, which completes the proof. ◻

Theorem 1.35. *If $\mu(\Omega) < \infty$ and $q > p$, then $L^q(\Omega) \subseteq L^p(\Omega)$.*

Proof. Consider the case that $q = 2$ and $p = 1$. Then by the Cauchy-Schwarz inequality,

$$\int_\Omega |f| dx = \int_\Omega |f| \cdot 1\, dx \leqslant \|f\|_{L^2(\Omega)} \sqrt{\mu(\Omega)}\,.$$

The general case follows from Hölder's inequality. ◻

1.3.7 Approximation of $L^p(\Omega)$ by simple functions

Lemma 1.36. *If $p \in [1, \infty)$, then the set of simple functions $f = \sum_{i=1}^{n} a_i \mathbf{1}_{E_i}$, where each E_i is a subset of $\mathbb{R}^n$ with $\mu(E_i) < \infty$, is dense in $L^p(\Omega)$. (Note that $\mathbf{1}_E$ denotes the indicator function for the set E, so that $\mathbf{1}_E(x) = 1$ for $x \in E$ and $\mathbf{1}_E(x) = 0$ for $x \in E^c$.)*

Proof. If $f \in L^p(\Omega)$, then f is measurable; thus, there exists a sequence $\{\phi_k\}_{k=1}^{\infty}$ of simple functions, such that $\phi_k \to f$ a.e. with

$$0 \leqslant |\phi_1| \leqslant |\phi_2| \leqslant \cdots \leqslant |f|\,;$$

that is, ϕ_k approximates f from below.

Recall that $|\phi_k - f|^p \to 0$ a.e. and $|\phi_k - f|^p \leqslant 2^p |f|^p \in L^1(\Omega)$, so by the Dominated Convergence Theorem, $\|\phi_k - f\|_{L^p(\Omega)} \to 0$.

Now, suppose that the set E_i are disjoint; then by the definition of the Lebesgue integral,

$$\int_\Omega \phi_k^p dx = \sum_{i=1}^{k} |a_i|^p \mu(E_i) < \infty\,.$$

If $a_i \neq 0$, then $\mu(E_i) < \infty$. ◻

1.3.8 Approximation of $L^p(\Omega)$ by continuous functions

Lemma 1.37. *Suppose that $\Omega \subseteq \mathbb{R}^n$ is bounded. Then $\mathscr{C}^0(\Omega)$ is dense in $L^p(\Omega)$ for $p \in [1, \infty)$.*

Proof. Let K be any compact subset of Ω. The functions

$$F_{K,\ell}(x) = \frac{1}{1 + \ell \operatorname{dist}(x, K)} \in \mathscr{C}^0(\Omega) \quad \text{satisfy} \quad F_{K,\ell} \leqslant 1\,,$$

and decrease monotonically to the characteristic function $\mathbf{1}_K$. The Monotone Convergence Theorem shows that

$$F_{K,\ell} \to \mathbf{1}_K \text{ in } L^p(\Omega), \quad 1 \leqslant p < \infty\,.$$

Next, let $A \subseteq \Omega$ be any measurable set. Then

$$\mu(A) = \sup\left\{\mu(K) \,\middle|\, K \subseteq A, K \text{ compact}\right\}.$$

It follows that there exists an increasing sequence of K_j of compact subsets of A such that $\lambda(A \backslash \bigcup_j K_j) = 0$. By the Monotone Convergence Theorem, $\mathbf{1}_{K_j} \to \mathbf{1}_A$ in $L^p(\Omega)$ for $p \in [1, \infty)$. According to Lemma 1.36, each function in $L^p(\Omega)$ is a norm limit of simple functions, so the lemma is proved. □

1.3.9 Approximation of $L^p(\Omega)$ by smooth functions

For $\Omega \subseteq \mathbb{R}^n$ open, for $\epsilon > 0$ taken sufficiently small, define the open subset of Ω by

$$\Omega_\epsilon := \left\{x \in \Omega \,\middle|\, \operatorname{dist}(x, \partial\Omega) > \epsilon\right\}.$$

Definition 1.38 (Mollifiers)**.** Define $\eta \in \mathscr{C}^\infty(\mathbb{R}^n)$ by

$$\eta(x) := \begin{cases} Ce^{(|x|^2-1)^{-1}} & \text{if } |x| < 1 \\ 0 & \text{if } |x| \geqslant 1 \end{cases},$$

with constant $C > 0$ chosen such that $\int_{\mathbb{R}^n} \eta(x)dx = 1$.

For $\epsilon > 0$, the standard sequence of mollifiers on $\mathbb{R}^n$ is defined by

$$\eta_\epsilon(x) = \epsilon^{-n}\eta(x/\epsilon)\,,$$

and satisfy $\int_{\mathbb{R}^n} \eta_\epsilon(x)dx = 1$ and $\operatorname{spt}(\eta_\epsilon) \subseteq \overline{B(0,\epsilon)}$. Note that $\eta_\epsilon \in \mathscr{C}_c^\infty(\mathbb{R}^n)$.

Definition 1.39. For $\Omega \subseteq \mathbb{R}^n$ open, set

$$L^p_{\text{loc}}(\Omega) = \left\{u : \Omega \to \mathbb{R} \,\middle|\, u \in L^p(\widetilde{\Omega})\ \forall\, \widetilde{\Omega} \subset\subset \Omega\right\},$$

where $\widetilde{\Omega} \subset\subset \Omega$ means that $\widetilde{\Omega}$ is compactly contained in Ω, i.e., there exists a compact set K such that $\widetilde{\Omega} \subseteq K \subseteq \Omega$. For example, K could be $\overline{\widetilde{\Omega}}$.

Definition 1.40 (Mollification of L^p functions for $1 \leqslant p < \infty$)**.** For $f \in L^p_{\text{loc}}(\Omega)$, we define its mollification by

$$f^\epsilon = \eta_\epsilon * f \quad \text{in} \quad \Omega_\epsilon \,,$$

so that

$$f^\epsilon(x) = \int_\Omega \eta_\epsilon(x-y) f(y) dy = \int_{B(0,\epsilon)} \eta_\epsilon(y) f(x-y) dy \quad \forall\, x \in \Omega_\epsilon \,.$$

Lemma 1.41 (Commuting the derivative with the integral)**.** *Let $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ denote an open and smooth subset. Let $(a,b) \subseteq \mathbb{R}$ be an open interval, and let $f : (a,b) \times \Omega \to \mathbb{R}$ be a function such that for each $t \in (a,b)$, $f(t,\cdot) : \Omega \to \mathbb{R}$ is integrable and $\frac{\partial f}{\partial t}(t,x)$ exists for each $(t,x) \in (a,b) \times \Omega$. Furthermore, assume that there is an integrable function $g : \Omega \to [0,\infty)$ such that $\sup_{t \in (a,b)} \left|\frac{\partial f}{\partial t}(t,x)\right| \leqslant g(x)$ for all $x \in \Omega$. Then the function h defined by $h(t) \equiv \int_\Omega f(t,x)dx$ is differentiable and the derivative is given by*

$$\frac{dh}{dt}(t) = \frac{d}{dt} \int_\Omega f(t,x)dx = \int_\Omega \frac{\partial f}{\partial t}(t,x)dx$$

for each $t \in (a,b)$.

Proof. Let $t_0 \in (a,b)$. To show that $\frac{dh}{dt}(t_0)$ exists, consider the limit of the sequence of difference quotients

$$\lim_{n\to\infty} \frac{h(t_n) - h(t_0)}{t_n - t_0} \,,$$

where $t_n \to t_0$ as $n \to \infty$. We see that

$$\frac{h(t_n) - h(t_0)}{t_n - t_0} = \int_\Omega \frac{f(t_n,x) - f(t_o,x)}{t_n - t_0} dx \,.$$

With

$$F_n(x) \equiv \frac{f(t_n,x) - f(t_0,x)}{t_n - t_0} \,,$$

it follows that $\lim_{n\to\infty} F_n(x) = \frac{\partial f}{\partial t}(t_0,x)$ for all $x \in \Omega$.

By the mean value theorem, there exists a point $\xi_n \in (t_0,t_n)$ such that

$$F_n(x) = \frac{\partial f}{\partial t}(\xi_n,x)$$

and since $\left|\frac{\partial f}{\partial t}(\xi_n,x)\right| \leqslant \sup_{t\in(a,b)} \left|\frac{\partial f}{\partial t}(t,x)\right|$, we have (by hypothesis) our dominating function; hence, by the dominated convergence theorem, it follows that

$$\lim_{n\to\infty} \frac{h(t_n)-h(t_0)}{t_n-t_0} = \int_\Omega \lim_{n\to\infty} F_n(x)dx = \int_\Omega \frac{\partial f}{\partial t}(t_0,x)dx\,. \qquad \square$$

Theorem 1.42 (Mollification of $L^p(\Omega)$ functions). *If for $p \in [1,\infty)$, $f \in L^p_{\text{loc}}(\Omega)$ and $f^\epsilon = \eta_\epsilon * f$ denotes the mollified function, then*

(A) $f^\epsilon \in \mathscr{C}^\infty(\Omega_\epsilon)$;

(B) $f^\epsilon \to f$ *a.e. as* $\epsilon \to 0$;

(C) *if in addition* $f \in \mathscr{C}^0(\Omega)$, *then* $f^\epsilon \to f$ *uniformly on compact subsets of* Ω;

(D) $f^\epsilon \to f$ *in* $L^p_{\text{loc}}(\Omega)$.

Proof. **Part (A).** Continuity of f^ϵ follows from the Dominated Convergence Theorem, and the fact that $\eta_\epsilon(x-y)|f(y)|1_{B(x,\epsilon)}$ is integrable. The fact that all partial derivatives of u^ϵ of all orders are continuous follows from repeated application of Lemma 1.41. To see that $\frac{\partial f^\epsilon}{\partial x_i}(x)$ exists and is continuous for each $x \in \Omega_\epsilon$ and $i = 1, ..., \text{n}$, we show that

$$\frac{\partial f^\epsilon}{\partial x_i}(x) = \int_{\mathbb{R}^\text{n}} \frac{\partial}{\partial x_i}\eta_\epsilon(x-y)f(y)dy\,.$$

From Definition 1.38, η_ϵ is a smooth function; hence, since $f \in L^1_{\text{loc}}(\Omega)$, we see that $y \mapsto \frac{\partial}{\partial x_i}\eta_\epsilon(x-y)u(y) \in L^1_{\text{loc}}(\Omega)$ uniformly in $x \in \omega$ for any set $\omega \subset\subset \Omega$. Application of Lemma 1.41 then shows that $f^\epsilon \in \mathscr{C}^1(\Omega_\epsilon)$. A similar argument shows that all higher-order partial derivatives of f^ϵ are continuous, and hence that $f^\epsilon \in \mathscr{C}^\infty(\Omega_\epsilon)$.

Step 2. Part (B). By the Lebesgue differentiation theorem,

$$\lim_{\epsilon\to 0} \frac{1}{|B(x,\epsilon)|}\int_{B(x,\epsilon)} |f(y)-f(x)|dy = 0 \text{ for a.e. } x \in \Omega\,. \tag{1.3}$$

Choose $x \in \Omega$ for which this limit holds. Then

$$\begin{aligned} |f_\epsilon(x)-f(x)| &\leqslant \int_{B(x,\epsilon)} \eta_\epsilon(x-y)|f(y)-f(x)|dy \\ &= \frac{1}{\epsilon^n}\int_{B(x,\epsilon)} \eta((x-y)/\epsilon)|f(y)-f(x)|dy \\ &\leqslant \frac{C}{|B(x,\epsilon)|}\int_{B(x,\epsilon)} |f(x)-f(y)|dy \longrightarrow 0 \quad \text{as} \quad \epsilon \to 0\,. \end{aligned} \tag{1.4}$$

Step 3. Part (C). We choose another set ω such that $\widetilde{\Omega} \subset\subset \omega \subset\subset \Omega$. Since f is continuous on Ω, it follows that f is uniformly continuous on ω. We choose $\epsilon > 0$ small enough so that f^ϵ is well defined on $\widetilde{\Omega}$. Then the limit in (1.3) holds uniformly for $x \in \widetilde{\Omega}$. The inequality (1.4) then shows that $f^\epsilon(x) \to f(x)$ uniformly on $\widetilde{\Omega}$.

Step 4. Part (D). For $f \in L^p_{\mathrm{loc}}(\Omega)$, $p \in [1, \infty)$, once again choose open sets $\widetilde{\Omega} \subset\subset \omega \subset\subset \Omega$; then, for $\epsilon > 0$ small enough,

$$\|f^\epsilon\|_{L^p(\tilde{\Omega})} \leqslant \|f\|_{L^p(\omega)} .$$

To see this, note that

$$\begin{aligned} |f^\epsilon(x)| &\leqslant \int_{B(x,\epsilon)} \eta_\epsilon(x-y)|f(y)|dy \\ &= \int_{B(x,\epsilon)} \eta_\epsilon(x-y)^{(p-1)/p}\eta_\epsilon(x-y)^{1/p}|f(y)|dy \\ &\leqslant \left(\int_{B(x,\epsilon)} \eta_\epsilon(x-y)dy\right)^{(p-1)/p} \left(\int_{B(x,\epsilon)} \eta_\epsilon(x-y)|f(y)|^p dy\right)^{1/p} , \end{aligned}$$

so that for $\epsilon > 0$ sufficiently small

$$\begin{aligned} \int_{\tilde{\Omega}} |f^\epsilon(x)|^p dx &\leqslant \int_{\tilde{\Omega}} \int_{B(x,\epsilon)} \eta_\epsilon(x-y)|f(y)|^p dy dx \\ &\leqslant \int_\omega |f(y)|^p \left(\int_{B(y,\epsilon)} \eta_\epsilon(x-y)dx\right) dy \leqslant \int_\omega |f(y)|^p dy . \end{aligned}$$

Since by Lemma 1.37, $\mathscr{C}^0(\omega)$ is dense in $L^p(\omega)$, choose $g \in \mathscr{C}^0(\omega)$ such that $\|f-g\|_{L^p(\omega)} < \delta$; thus

$$\begin{aligned} \|f^\epsilon - f\|_{L^p(\tilde{\Omega})} &\leqslant \|f^\epsilon - g^\epsilon\|_{L^p(\tilde{\Omega})} + \|g^\epsilon - g\|_{L^p(\tilde{\Omega})} + \|g - f\|_{L^p(\tilde{\Omega})} \\ &\leqslant 2\|f-g\|_{L^p(\omega)} + \|g^\epsilon - g\|_{L^p(\tilde{\Omega})} \leqslant 2\delta + \|g^\epsilon - g\|_{L^p(\omega)} . \end{aligned}$$

□

1.4 Convolutions and Integral Operators

If $u : \mathbb{R}^n \to \mathbb{R}$ satisfies certain integrability conditions, then we can define the operator K acting on the function u as follows:

$$(Ku)(x) = \int_{\mathbb{R}^n} k(x,y)u(y)dy ,$$

where $k(x,y)$ is called the *integral kernel.* The mollification procedure, introduced in Definition 1.40, is one example of the use of integral operators; the Fourier transform is another.

Definition 1.43. Let $\mathscr{B}(L^p(\mathbb{R}^n), L^p(\mathbb{R}^n))$ denote the collection of bounded linear operators from $L^p(\mathbb{R}^n)$ to itself. Using the Representation Theorem 1.51, the natural norm on $\mathscr{B}(L^p(\mathbb{R}^n), L^p(\mathbb{R}^n))$ is given by

$$\|K\|_{\mathscr{B}(L^p(\mathbb{R}^n),L^p(\mathbb{R}^n))} = \sup_{\|f\|_{L^p(\mathbb{R}^n)}=1} \sup_{\|g\|_{L^q(\mathbb{R}^n)}=1} \left| \int_{\mathbb{R}^n} (Kf)(x) g(x) dx \right| .$$

Theorem 1.44. *Let $1 \leqslant p < \infty$, $(Ku)(x) = \int_{\mathbb{R}^n} k(x,y)u(y)dy$, and suppose that*

$$\int_{\mathbb{R}^n} |k(x,y)| dx \leqslant C_1 \ \forall\, y \in \mathbb{R}^n \ \text{and} \ \int_{\mathbb{R}^n} |k(x,y)| dy \leqslant C_2 \ \forall\, x \in \mathbb{R}^n ,$$

where $0 < C_1, C_2 < \infty$. Then $K : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$ is bounded and

$$\|K\|_{\mathscr{B}(L^p(\mathbb{R}^n),L^p(\mathbb{R}^n))} \leqslant C_1^{\frac{1}{p}} C_2^{\frac{p-1}{p}} .$$

In order to prove Theorem 1.44, we will need another well-known inequality.

Lemma 1.45 (Cauchy-Young Inequality)**.** *If $\frac{1}{p} + \frac{1}{q} = 1$, then for all $a, b \geqslant 0$,*

$$ab \leqslant \frac{a^p}{p} + \frac{b^q}{q} .$$

Proof. Suppose that $a, b > 0$, otherwise the inequality trivially holds.

$$\begin{aligned} ab &= \exp(\log(ab)) = \exp(\log a + \log b) \quad (\text{since } a, b > 0) \\ &= \exp\left(\frac{1}{p}\log a^p + \frac{1}{q}\log b^q\right) \\ &\leqslant \frac{1}{p}\exp(\log a^p) + \frac{1}{q}\exp(\log b^q) \quad (\text{using the convexity of } \exp) \\ &= \frac{a^p}{p} + \frac{b^q}{q} \end{aligned}$$

where we have used the condition $\frac{1}{p} + \frac{1}{q} = 1$. □

Lemma 1.46 (Cauchy-Young Inequality with δ)**.** *If* $\frac{1}{p}+\frac{1}{q}=1$*, then for all* $a,b\geqslant 0$*,*

$$ab \leqslant \delta\, a^p + C_\delta b^q\,, \qquad \delta>0\,,$$

with $C_\delta=(\delta p)^{-q/p}q^{-1}$.

Proof. This is a trivial consequence of Lemma 1.45 by setting

$$ab = a\cdot(\delta p)^{1/p}\,\frac{b}{(\delta p)^{1/p}}\,. \qquad \square$$

Proof of Theorem 1.44. According to Lemma 1.45, $|f(y)g(x)|\leqslant \frac{|f(y)|^p}{p}+\frac{|g(x)|^q}{q}$ so that

$$\begin{aligned}\Big|\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} & k(x,y)f(y)g(x)dydx\Big|\\ &\leqslant \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\frac{|k(x,y)|}{p}dx|f(y)|^p dy+\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\frac{|k(x,y)|}{q}dy|g(x)|^q dx\\ &\leqslant \frac{C_1}{p}\|f\|^p_{L^p(\Omega)}+\frac{C_2}{q}\|g\|^q_{L^q(\Omega)}\,.\end{aligned}$$

To improve this bound, notice that

$$\begin{aligned}\Big|\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} & k(x,y)f(y)g(x)dydx\Big|\\ &\leqslant \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\frac{|k(x,y)|}{p}dx|tf(y)|^p dy+\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\frac{|k(x,y)|}{q}dy|t^{-1}g(x)|^q dx\\ &\leqslant \frac{C_1t^p}{p}\|f\|^p_{L^p(\Omega)}+\frac{C_2t^{-q}}{q}\|g\|^q_{L^q(\Omega)}=:F(t)\,.\end{aligned}$$

Find the value of t for which $F(t)$ has a minimum to establish the desired bounded. $\square$

Theorem 1.47 (Simple version of Young's inequality)**.** *Suppose that* $k\in L^1(\mathbb{R}^n)$ *and* $f\in L^p(\mathbb{R}^n)$*. Then*

$$\|k*f\|_{L^p(\mathbb{R}^n)}\leqslant \|k\|_{L^1(\mathbb{R}^n)}\|f\|_{L^p(\mathbb{R}^n)}\,.$$

Proof. Define

$$K_k(f)=k*f:=\int_{\mathbb{R}^n}k(x-y)f(y)dy\,.$$

Let $C_1=C_2=\|k\|_{L^1(\mathbb{R}^n)}$ in Theorem 1.44. Then according to Theorem 1.44, $K_k: L^p(\mathbb{R}^n)\to L^p(\mathbb{R}^n)$ and $\|K_k\|_{\mathscr{B}(L^p(\mathbb{R}^n),L^p(\mathbb{R}^n))}\leqslant C_1$. $\square$

Theorem 1.44 can easily be generalized to the setting of integral operators $K : L^q(\mathbb{R}^n) \to L^r(\mathbb{R}^n)$ built with kernels $k \in L^p(\mathbb{R}^n)$ such that $1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Such a generalization leads to

Theorem 1.48 (Young's inequality for convolution)**.** *Suppose that $k \in L^p(\mathbb{R}^n)$ and $f \in L^q(\mathbb{R}^n)$. Then*

$$\|k * f\|_{L^r(\mathbb{R}^n)} \leqslant \|k\|_{L^p(\mathbb{R}^n)} \|f\|_{L^q(\mathbb{R}^n)} \quad \text{for} \quad 1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}\,. \tag{1.5}$$

1.5 The Dual Space and Weak Topology

1.5.1 Continuous linear functionals on $L^p(\Omega)$

Let $L^p(\Omega)'$ denote the dual space of $L^p(\Omega)$, consisting of all continuous linear functions $\phi : L^p(\Omega) \to \mathbb{R}$. For $\phi \in L^p(\Omega)'$, the $L^p(\Omega)'$-norm of ϕ is defined by

$$\|\phi\|_{L^p(\Omega)'} = \sup_{\|f\|_{L^p(\Omega)}=1} |\phi(f)|\,.$$

This is the so-called *operator norm* which we shall sometimes denote by $\|\phi\|_{\text{op}}$

Theorem 1.49. *Let $p \in (1, \infty]$, $q = \frac{p}{p-1}$. For $g \in L^q(\Omega)$, define $F_g : L^p(\Omega) \to \mathbb{R}$ as*

$$F_g(f) = \int_\Omega fg dx\,.$$

Then $F_g \in L^p(\Omega)'$ and $\|F_g\|_{\text{op}} = \|g\|_{L^q(\Omega)}$.

Proof. The linearity of F_g follows from the linearity of the Lebesgue integral. By Hölder's inequality,

$$|F_g(f)| = \left|\int_\Omega fg dx\right| \leqslant \int_\Omega |fg|\, dx \leqslant \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)}\ ,$$

so that $\sup_{\|f\|_{L^p(\Omega)}=1} |F_g(f)| \leqslant \|g\|_{L^q(\Omega)}$.

For the reverse inequality, we first consider the case that $p \in (0, \infty]$, and set $f = |g|^{q-1} \operatorname{sgn}(g)$; then, f is measurable and in $L^p(\Omega)$ since $|f|^p = |f|^{\frac{q}{q-1}} = |g|^q$ and since $fg = |g|^q$,

$$\begin{aligned} F_g(f) &= \int_\Omega fg dx = \int_\Omega |g|^q dx = \Big(\int_\Omega |g|^q dx\Big)^{\frac{1}{p}+\frac{1}{q}} \\ &= \Big(\int_\Omega |f|^p dx\Big)^{\frac{1}{p}} \Big(\int_\Omega |g|^q dx\Big)^{\frac{1}{q}} = \|f\|_{L^p(\Omega)} \|g\|_{L^q(\Omega)} \end{aligned}$$

so that $\|g\|_{L^q(\Omega)} = \dfrac{F_g(f)}{\|f\|_{L^p(\Omega)}} = F_g\left(\dfrac{f}{\|f\|_{L^p(\Omega)}}\right) \leqslant \|F_g\|_{\mathrm{op}}$.

Next, we consider the case that $p = 1$ and $q = \infty$. We can assume that $g \neq 0$ a.e, for otherwise, the equality is trivial. For $\epsilon > 0$, let $E_\epsilon = \{x \in \Omega \ : \ |g(x)| \geqslant \|g\|_{L^\infty(\Omega)} - \epsilon\}$, and set $f(x) = \mathbf{1}_{E_\epsilon}(x)\frac{\operatorname{sgn}(g(x))}{|E_\epsilon|}$. Then $\|f\|_{L^1(\Omega)} = 1$ and

$$\left|\int_\Omega f(x)g(x)dx\right| = \frac{1}{|E_\epsilon|}\int_{E_\epsilon} |g(x)dx \geqslant \|g\|_{L^\infty(\Omega)} - \epsilon\,.$$

□

REMARK 1.50. Theorem 1.49 shows that for $1 < p \leqslant \infty$, there exists a linear isometry $g \mapsto F_g$ from $L^q(\Omega)$ into $L^p(\Omega)'$, the dual space of $L^p(\Omega)$. When $p = \infty$, $g \mapsto F_g : L^1(\Omega) \to L^\infty(\Omega)'$ is rarely onto ($L^\infty(\Omega)'$ is strictly larger than $L^1(\Omega)$); on the other hand, if the measure space Ω is σ-finite, then $L^\infty(\Omega) = L^1(\Omega)'$.

1.5.2 A theorem of F. Riesz

THEOREM 1.51 (Representation theorem)**.** *Suppose that $1 \leqslant p < \infty$ and $\phi \in L^p(\Omega)'$. Then there exists $g \in L^q(\Omega)$, $q = \dfrac{p}{p-1}$ such that*

$$\phi\,(f) = \int_\Omega fg dx \quad \forall\, f \in L^p(\Omega)\,,$$

and $\|\phi\|_{\mathrm{op}} = \|g\|_{L^q(\Omega)}$.

COROLLARY 1.52. *For $p \in (1, \infty)$ the space $L^p\,(\Omega, \mu)$ is reflexive; that is, $L^p(\Omega)'' = L^p(\Omega)$.*

The proof Theorem 1.51 crucially relies on the Radon-Nikodym theorem, whose statement requires the following definition.

DEFINITION 1.53. If μ and ν are measure on (Ω, A) then $\nu \ll \mu$ if $\nu(E) = 0$ for every set E for which $\mu(E) = 0$. In this case, we say that ν is absolutely continuous with respect to μ.

THEOREM 1.54 (Radon-Nikodym)**.** *If μ and ν are two finite measures on Ω; that is, $\mu(\Omega) < \infty, \nu(\Omega) < \infty$, and $\nu \ll \mu$, then*

$$\int_\Omega F(x)\,d\nu(x) = \int_\Omega F(x)h(x)d\mu(x) \tag{1.6}$$

holds for some non-negative function $h \in L^1(\Omega, \mu)$ and every positive measurable function F.

Proof. Define measures $\alpha = \mu + 2\nu$ and $\omega = 2\mu + \nu$, and let $H = L^2(\Omega, \alpha)$ (a Hilbert space) and suppose $\phi : L^2(\Omega, \alpha) \to \mathbb{R}$ is defined by $\phi(f) = \int_\Omega f d\omega$. We show that ϕ is a bounded linear functional since

$$\begin{aligned}|\phi(f)| = \Big|\int_\Omega f\, d(2\mu + \nu)\Big| \leqslant \int_\Omega |f|\, d(2\mu + 4\nu) = 2\int_\Omega |f|\, d\alpha \\ \leqslant \|f\|_{L^2(x,\alpha)} \sqrt{\alpha(\Omega)}\,.\end{aligned}$$

Thus, by the L^2 Riesz representation theorem[3], there exists $g \in L^2(\Omega, \alpha)$ such that

$$\phi(f) = \int_\Omega f\, d\omega = \int_\Omega f g\, d\alpha\,,$$

which implies that

$$\int_\Omega f(2g-1)d\nu = \int_\Omega f(2-g)d\mu\,. \tag{1.7}$$

Given $0 \leqslant F$ a measurable function on Ω, if we set $f = \dfrac{F}{2g-1}$ and $h = \dfrac{2-g}{2g-1}$ then $\int_\Omega F d\nu = \int_\Omega F h\, dx$ which is the desired result, if we can prove that $\frac{1}{2} \leqslant g(x) \leqslant 2$. Define the sets

$$E_k^1 = \Big\{x \in \Omega \,\Big|\, g(x) < \frac{1}{2} - \frac{1}{k}\Big\} \quad \text{and} \quad E_k^2 = \Big\{x \in \Omega \,\Big|\, g(x) > 2 + \frac{1}{k}\Big\}\,.$$

By substituting $f = \mathbf{1}_{E_k^j}$, $j = 1, 2$ in (1.7), we see that

$$\mu(E_k^j) = \nu(E_k^j) = 0 \text{ for } j = 1, 2\,,$$

from which the bounds $1/2 \leqslant g(x) \leqslant 2$ hold. Also $\mu\big(\{x \in \Omega \,|\, g(x) = 1/2\}\big) = 0$ and $\nu\big(\{x \in \Omega \,|\, g(x) = 2\}\big) = 0$. Notice that if $F = 1$, then $h \in L^1(\Omega)$. ▫

Remark 1.55 (The more general version of the Radon-Nikodym theorem)**.** Suppose that $\mu(\Omega) < \infty$, ν is a finite signed measure (by the Hahn decomposition, $\nu = \nu^- + \nu^+$) such that $\nu \ll \mu$; then, there exists $h \in L^1(\Omega, \mu)$ such that $\int_\Omega F\, d\nu = \int_\Omega F h\, d\mu$.

[3] The L^2 Riesz representation theorem is proved using the orthogonality relations that the L^2 inner-product provides, together with the Hahn-Banach Theorem.

Lemma 1.56 (Converse to Hölder's inequality)**.** *Let $\mu(\Omega) < \infty$. Suppose that g is measurable and $fg \in L^1(\Omega)$ for all simple functions f. If*

$$M(g) = \sup_{\|f\|_{L^p(\Omega)}=1} \left\{ \left| \int_\Omega fg\, d\mu \right| \Big| f \text{ is a simple function} \right\} < \infty\,, \tag{1.8}$$

then $g \in L^q(\Omega)$, and $\|g\|_{L^q(\Omega)} = M(g)$.

Proof. Let $\{\phi_k\}_{k=1}^\infty$ be a sequence of simple functions such that $\phi_k \to g$ a.e. and $|\phi_k| \leqslant |g|$. Set

$$f_k = \frac{|\phi_k|^{q-1}\operatorname{sgn}(\phi_k)}{\|\phi_k\|_{L^q(\Omega)}^{q-1}}$$

so that $\|f_k\|_{L^p(\Omega)} = 1$ for $p = \dfrac{q}{q-1}$. By Fatou's lemma,

$$\|g\|_{L^q(\Omega)} \leqslant \liminf_{k\to\infty} \|\phi_k\|_{L^q(\Omega)} = \liminf_{k\to\infty} \int_\Omega |f_k\phi_k| d\mu\,.$$

Since $\phi_k \to g$ a.e., then

$$\|g\|_{L^q(\Omega)} \leqslant \liminf_{k\to\infty} \int_\Omega |f_k\phi_k| d\mu \leqslant \liminf_{k\to\infty} \int_\Omega |f_k g| d\mu \leqslant M(g)\,.$$

The reverse inequality is implied by Hölder's inequality. □

Proof of Theorem 1.51. We have already proven that there exists a natural inclusion $\iota : L^q(\Omega) \to L^p(\Omega)'$ which is an isometry. It remains to show that ι is surjective.

Let $\phi \in L^p(\Omega)'$ and define a set function ν on measurable subsets $E \subseteq \Omega$ by

$$\nu(E) = \int_\Omega \mathbf{1}_E d\nu =: \phi(\mathbf{1}_E)\,.$$

Thus, if $\mu(E) = 0$, then $\nu(E) = 0$. Then

$$\int_\Omega f\, d\nu =: \phi(f)$$

for all simple functions f, and by Lemma 1.36, this holds for all $f \in L^p(\Omega)$. By the Radon-Nikodym theorem, there exists $0 \leqslant g \in L^1(\Omega)$ such that

$$\int_\Omega f\, d\nu = \int_\Omega fg\, d\mu \quad \forall\, f \in L^p(\Omega)\,.$$

But

$$\phi(f) = \int_\Omega f\, d\nu = \int_\Omega fg\, d\mu \tag{1.9}$$

and since $\phi \in L^p(\Omega)'$, then $M(g)$ given by (1.8) is finite, and by the converse to Hölder's inequality, $g \in L^q(\Omega)$, and $\|\phi\|_{\mathrm{op}} = M(g) = \|g\|_{L^q(\Omega)}$. □

1.5.3 Weak convergence

The importance of the Representation Theorem 1.51 is in the use of the weak-* topology on the dual space $L^p(\Omega)'$. Recall that for a Banach space $\mathbb{B}$ and for any sequence ϕ_j in the dual space $\mathbb{B}'$, $\phi_j \overset{*}{\rightharpoonup} \phi$ in $\mathbb{B}'$ weak-*, if $\langle \phi_j, f\rangle \to \langle \phi, f\rangle$ for each $f \in \mathbb{B}$, where $\langle \cdot, \cdot\rangle$ denotes the duality pairing between $\mathbb{B}'$ and $\mathbb{B}$.

Theorem 1.57 (Alaoglu's Lemma)**.** *If $\mathbb{B}$ is a Banach space, then the closed unit ball in $\mathbb{B}'$ is compact in the weak -* topology.*

Definition 1.58. For $1 \leqslant p < \infty$, a sequence $\{f_k\}_{k=1}^\infty \subseteq L^p(\Omega)$ is said to *weakly converge* to $f \in L^p(\Omega)$ if

$$\int_\Omega f_k(x)\phi(x)dx \to \int_\Omega f(x)\phi(x)dx \quad \forall\, \phi \in L^q(\Omega), q = \frac{p}{p-1}\,.$$

We denote this convergence by saying that $f_k \rightharpoonup f$ in $L^p(\Omega)$ weakly.

Given that $L^p(\Omega)$ is reflexive for $p \in (1, \infty)$, a simple corollary of Alaoglu's Lemma is the following

Theorem 1.59 (Weak compactness for $L^p(\Omega)$, $1 < p < \infty$)**.** *If $1 < p < \infty$ and $\{f_k\}_{k=1}^\infty$ is a bounded sequence in $L^p(\Omega)$, then there exists a subsequence $\{f_{k_j}\}_{j=1}^\infty$ such that $f_{k_j} \rightharpoonup f$ in $L^p(\Omega)$ weakly.*

Definition 1.60. A sequence $\{f_k\}_{k=1}^\infty \subseteq L^\infty(\Omega)$ is said to converge weak-* to $f \in L^\infty(\Omega)$ if

$$\int_\Omega f_k(x)\phi(x)dx \to \int_\Omega f(x)\phi(x)dx \quad \forall\, \phi \in L^1(\Omega)\,.$$

We denote this convergence by saying that $f_k \overset{*}{\rightharpoonup} f$ in $L^\infty(\Omega)$ weak-*.

Theorem 1.61 (Weak-* compactness for L^∞)**.** *If $\{f_k\}_{k=1}^\infty$ is a bounded sequence in $L^\infty(\Omega)$, then there exists a subsequence $\{f_{k_j}\}_{j=1}^\infty$ such that $f_{k_j} \overset{*}{\rightharpoonup} f$ in $L^\infty(\Omega)$ weak-*.*

Lemma 1.62. *If $f_k \to f$ in $L^p(\Omega)$, then $f_k \rightharpoonup f$ in $L^p(\Omega)$.*

Proof. By Hölder's inequality,

$$\left|\int_\Omega g(f_k - f)dx\right| \leqslant \|f_k - f\|_{L^p(\Omega)}\|g\|_{L^q(\Omega)}\,. \qquad \square$$

Note that if $\{f_k\}_{k=1}^\infty$ is weakly convergent, in general, this does not imply that $\{f_k\}_{k=1}^\infty$ is strongly convergent.

Example 1.63. If $p=2$, let $\{f_k\}_{k=1}^\infty$ denote any orthonormal sequence in $L^2(\Omega)$. From Bessel's inequality

$$\sum_{k=1}^\infty \left| \int_\Omega f_k g dx \right|^2 \leqslant \|g\|^2_{L^2(\Omega)} ,$$

we see that $f_k \rightharpoonup 0$ in $L^2(\Omega)$.

We can often arrive at the same conclusion by more elementary arguments.

Example 1.64. Let $u_k(x) = \sin(kx)$ and let $\Omega = (0, 2\pi)$. In this case the $u_k \rightharpoonup 0$ in $L^2(0,2\pi)$, but this sequence does not converge strongly.

So we must show that $\int_0^{2\pi} \sin(kx) v(x) dx \to 0$ as $k \to \infty$ for all $v \in L^2(0,2\pi)$. By Theorem 1.42, we see that $\mathscr{C}^1([0,2\pi])$ is dense in $L^2(0,2\pi)$ (as the interval $(0,2\pi)$ is identified with the circle $\mathbb{S}^1$ which has no boundary). Thus, we consider our test function $v \in \mathscr{C}^1([0,2\pi])$ so that for some constant $M \geqslant 0$, $\max_{x\in[0,2\pi]} \left(|v(x)| + \left|\frac{dv}{dx}(x)\right| \right) \leqslant M$. Then

$$\begin{aligned}
\int_0^{2\pi} \sin(kx) v(x) dx &= -\frac{1}{k} \int_0^{2\pi} \frac{d}{dx} \cos(kx) v(x) dx \\
&= \frac{-v(x)\cos(kx)}{k}\bigg|_0^{2\pi} + \frac{1}{k}\int_0^{2\pi} \frac{dv}{dx}(x) \cos(kx) dx \\
&\leqslant \frac{1}{k}\big(- v(2\pi) + v(0)\big) + \frac{1}{k} \|v'\|_{L^2(\Omega)} \| \cos(k\cdot)\|_{L^2(\Omega)} \\
&\leqslant \frac{C}{k} \to 0\,.
\end{aligned}$$

Employing a density argument, we see that $\int_0^{2\pi} \sin(kx) v(x) dx \to 0$ as $k \to \infty$ for all $v \in L^2(0,2\pi)$.

On the other hand,

$$\begin{aligned}
\| \sin^2(kx)\|^2_{L^2(0,2\pi)} &= \int_0^{2\pi} |\sin(kx) - 0|^2 dx = \frac{1}{k}\int_0^{2\pi k} sin^2 y\, dy \\
&= \frac{1}{2k}(y - \sin y \cos y)\bigg|_0^{2\pi k} = \pi\,,
\end{aligned}$$

so that $\sin(kx)$ does not converge strongly in $L^2(0,2\pi)$.

We have just shown that $u_k \rightharpoonup 0$ in $L^2(0,2\pi)$, and an interesting question is the following: what does u_k^2 weakly converge to? Example 1.64 is an example of a more general fact that periodic functions weakly converge to their average as the wavelength tends to zero (see Problem 1.13).

Example 1.65. Let $f_k = \sin^2(kx)$ and once again, set $\Omega = (0,2\pi)$. We will show that $f_k \rightharpoonup \frac{1}{2}$ in $L^2(0,2\pi)$, which is the same as showing that for all $v \in L^2(0,2\pi)$,

$$\int_0^{2\pi} \sin^2(kx)v(x)dx \to \int_0^{2\pi} \frac{v}{2}dx\,. \tag{1.10}$$

By Lemma 1.36, it suffices to prove (1.10) for all simple functions $v(x)$, and by linearity of the integral, we may consider $v(x) = \mathbf{1}_{(0,a)}(x)$ for some $a \in (0,2\pi)$. In this case, (1.10) reduces to

$$\int_0^a \sin^2(kx)dx \to \frac{a}{2}\,,$$

and this follows from the anti-derivative formula given in Example 1.64.

There are essentially three types of mechanisms by which a sequence $u_k \rightharpoonup u$ in $L^p(\Omega)$ but $u_k \nrightarrow u$ in $L^p(\Omega)$. We have just seen examples of the first mechanism: *oscillation*, for which $u_k(x) = \sin(kx)$ is a nice example. The second mechanism is *concentration*, and the sequence $u_k(x) = k^{1/p}h(kx)$ for any fixed function $h \in L^p(\mathbb{R})$; for example, letting $h(x) = e^{-|x|}$ for $x \in \mathbb{R}$, we see that $u_k(x)$ concentrates near the origin $x = 0$, and has unbounded amplitude as $k \to \infty$. (In fact, as we shall sees later, this sequence converges to the Dirac measure in the sense of distribution.) The third mechanism can be termed *'escape to ∞'*, wherein $u_k(x) = h(x+k)$ for some fixed $h \in L^p(\mathbb{R})$.

Returning to example 1.63, we see that the map $f \mapsto \|f\|_{L^p(\Omega)}$ is continuous, but not weakly continuous. It is, however, weakly lower-semicontinuous.

Theorem 1.66. *If $f_k \rightharpoonup f$ weakly in $L^p(\Omega)$, then $\|f\|_{L^p(\Omega)} \leqslant \liminf_{k\to\infty} \|f_k\|_{L^p(\Omega)}$.*

Proof. As a consequence of Theorem 1.51,

$$\begin{aligned}\|f\|_{L^p(\Omega)} &= \sup_{\|g\|_{L^q(\Omega)}=1} \left|\int_\Omega fgdx\right| = \sup_{\|g\|_{L^q(\Omega)}=1} \lim_{n\to\infty} \left|\int_\Omega f_n gdx\right| \\ &\leqslant \sup_{\|g\|_{L^q(\Omega)}=1} \liminf_{n\to\infty} \|f_n\|_{L^p(\Omega)}\|g\|_{L^q(\Omega)}\,.\end{aligned}$$

The inequality follows by noting that $\lim_{k\to\infty}\left|\int_\Omega f_k gdx\right| = \liminf_{k\to\infty}\left|\int_\Omega f_k gdx\right|$. □

Theorem 1.67. *If $f_k \rightharpoonup f$ in $L^p(\Omega)$, then f_k is bounded in $L^p(\Omega)$.*

Proof. This is an immediate consequence of the uniform boundedness principle[4], by identifying f_k with an element ϕ_k of $L^q(\Omega)'$, and using Theorem 1.51 to conclude that $\|\phi_k\|_{L^q(\Omega)'} = \|f_k\|_{L^p(\Omega)}$. □

An important result in analysis, known as Egoroff's theorem, is useful in answering a variety of questions about convergence of sequences of functions.

Theorem 1.68 (Egoroff's Theorem)**.** *Suppose that $|\Omega| < \infty$ and $f_k(x) \to f(x)$ for all $x \in \Omega$. Then for each $\epsilon > 0$, there exists $E \subseteq \Omega$ with $|E| < \epsilon$ such that $f_k \to f$ uniformly on $E^{\complement}$.*

We use the notation $E^{\complement}$ to denote the complement of the set E in Ω.

Proof. For each $\delta > 0$ and each $k \in \mathbb{N}$, we define the subsets

$$E_k^\delta = \left\{ x \in \Omega \,\middle|\, |f_j(x) - f(x)| \geqslant \delta \text{ for some } j \geqslant k \right\}.$$

Since $f_j(x) \to f(x)$ for all $x \in \Omega$, it follows that $\bigcap_{k=1}^{\infty} E_k^\delta = \varnothing$ for each $\delta > 0$, so that $|E_k^\delta| \to 0$ as $k \to \infty$.

If for each $\epsilon > 0$, we set $\delta = 2^{-k}$, then there exists a positive integer N_k such that

$$\left| E_{N_k}^{2^{-k}} \right| \leqslant 2^{-k}\epsilon\,.$$

We define the set

$$E = \bigcup_{k=1}^{\infty} E_{N_k}^{2^{-k}}\,.$$

Then $|E| \leqslant \epsilon$ and if $x \in U^{\complement}$ and $j \geqslant N_k$, then $|f_j(x) - f(x)| < 2^{-k}$, which provides the uniform convergence on the complement of E. □

[4]The uniform boundedness principle is a fundamental theorem of functional analysis: Suppose X and Y are Banach spaces and $\mathcal{F}$ is a collection of continuous linear operators from X to Y. If

$$\sup_{L \in \mathcal{F}} \|L(x)\|_Y < \infty \text{ for all } x \in X\,,$$

then

$$\sup_{L \in \mathcal{F}} \|L\|_{\mathcal{B}(X,Y)} := \sup_{L \in \mathcal{F}, \|x\|=1} \|L(x)\|_Y < \infty\,.$$

Theorem 1.69. *Suppose that $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ is a bounded domain, and*

$$\sup_k \|f_k\|_{L^p(\Omega)} \leqslant M < \infty \quad \text{and} \quad f_k \to f \quad a.e.$$

If $1 < p < \infty$, then $f_k \rightharpoonup f$ in $L^p(\Omega)$.

Proof. Egoroff's theorem states that for all $\epsilon > 0$, there exists $E \subseteq \Omega$ such that $\mu(E) < \epsilon$ and $f_k \to f$ uniformly on $E^{\complement}$. By definition, $f_k \rightharpoonup f$ in $L^p(\Omega)$ for $p \in (1, \infty)$ if $\int_\Omega (f_k - f) g dx \to 0$ for all $g \in L^q(\Omega)$, $q = \dfrac{p}{p-1}$. We have the inequality

$$\left| \int_\Omega (f_k - f) g dx \right| \leqslant \int_E |f_k - f|\, |g|\, dx + \int_{E^{\complement}} |f_k - f|\, |g|\, dx\,.$$

Choose $n \in \mathbb{N}$ sufficiently large, so that $|f_k(x) - f(x)| \leqslant \delta$ for all $x \in E^{\complement}$. By Hölder's inequality,

$$\int_{E^{\complement}} |f_k - f|\, |g|\, dx \leqslant \|f_k - f\|_{L^p(E^{\complement})} \|g\|_{L^q(E^{\complement})} \leqslant \delta \mu(E^{\complement}) \|g\|_{L^q(\Omega)} \leqslant C\delta$$

for a constant $C < \infty$.

By the Dominated Convergence Theorem, $\|f_k - f\|_{L^p(\Omega)} \leqslant 2M$ so by Hölder's inequality, the integral over E is bounded by $2M\|g\|_{L^q(E)}$. Next, we use the fact that the integral is continuous with respect to the measure of the set over which the integral is taken. In particular, if $0 \leqslant h$ is integrable, then for all $\delta > 0$, there exists $\epsilon > 0$ such that if the set E_ϵ has measure $\mu(E_\epsilon) < \epsilon$, then $\int_{E_\epsilon} h dx \leqslant \delta$. To see this, either approximate h by simple functions, or use the Dominated Convergence theorem for the integral $\int_\Omega \mathbf{1}_{E_\epsilon}(x) h(x) dx$. □

Remark 1.70. The proof of Theorem 1.69 does not work in the case that $p = 1$, as Hölder's inequality gives

$$\int_E |f_k - f|\, |g|\, dx \leqslant \|f_k - f\|_{L^1(\Omega)} \|g\|_{L^\infty(E)}\,,$$

so we lose the smallness of the right-hand side.

Remark 1.71. Suppose that $E \subseteq \Omega$ is bounded and measurable, and let $g = \mathbf{1}_E$. If $f_n \rightharpoonup f$ in $L^p(\Omega)$, then

$$\int_E f_k(x) dx \to \int_E f(x) dx;$$

hence, if $f_k \rightharpoonup f$, then the average of f_n converges to the average of f pointwise.

REMARK 1.72. If $u_k \rightharpoonup u$ in $L^p(\Omega)$ and $v_k \to v$ in $L^q(\Omega)$, then $\int_\Omega u_k v_k dx \to \int_\Omega uv dx$.

REMARK 1.73. For $1 < p < \infty$, if $u_k \rightharpoonup u$ in $L^p(\Omega)$ and $\|u\|_{L^p(\Omega)} = \lim_{k\to\infty} \|u_k\|_{L^p(\Omega)}$, then $u_k \to u$ in $L^p(\Omega)$ strongly.

1.6 Exercises

PROBLEM 1.1. Use the Monotone Convergence Theorem to prove Fatou's Lemma.

PROBLEM 1.2. Use Fatou's Lemma to prove the Dominated Convergence Theorem.

PROBLEM 1.3. Let Ω denote an open subset of $\mathbb{R}^n$. If $f \in L^1(\Omega) \cap L^\infty(\Omega)$, show that $f \in L^p(\Omega)$ for $1 < p < \infty$. If $|\Omega| < \infty$, then show that $\lim_{p\nearrow\infty} \|f\|_{L^p} = \|f\|_{L^\infty}$. (**Hint**: For $\epsilon > 0$, you can prove that the set $E = \{x \in \Omega \ : \ |f(x)| > \|f\|_{L^\infty} - \epsilon\}$ has positive Lebesgue measure, and the inequality $[\|f\|_{L^\infty} - \epsilon]\, \mathbf{1}_E \leqslant |f|$ holds.) Can you remove the assumption that $|\Omega| < \infty$?

PROBLEM 1.4. Theorem 1.30 states that if $1 \leqslant p < \infty$, $f \in L^p$, $\{f_n\} \subseteq L^p$, $f_n \to f$ a.e., and $\lim_{n\to\infty} \|f_n\|_{L^p} = \|f\|_{L^p}$, then $\lim_{n\to\infty} \|f_n - f\|_{L^p} \to 0$. Show by an example that this theorem is false when $p = \infty$.

PROBLEM 1.5. Show that equality holds in the inequality

$$a^\lambda b^{1-\lambda} \leqslant \lambda a + (1-\lambda)b, \quad \lambda \in (0,1), a, b \geqslant 0$$

if and only if $a = b$. Use this to show that if $f \in L^p$ and $g \in L^q$ for $1 < p, q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\int_\Omega |fg| dx = \|f\|_{L^p} \ \|g\|_{L^q}$$

holds if and only if there exists two constants C_1 and C_2 (not both zero) such that $C_1|f|^p = C_2|g|^q$ holds.

PROBLEM 1.6. Use the result of Problem 1.5 to prove that if $f, g \in L^3(\Omega)$ satisfy

$$\|f\|_{L^3} = \|g\|_{L^3} = \int_\Omega f^2\, g\, dx = 1\,,$$

then $g = |f|$ a.e.

PROBLEM 1.7. If for $j = 1, 2$ and $p_j \in [1, \infty]$ and $u_j \in L^{p_j}$, show that $u_1 u_2 \in L^r$ provided that $\frac{1}{r} = \frac{1}{p_1} + \frac{1}{p_2}$ and

$$\|u_1 u_2\|_{L^r} \leqslant \|u_1\|_{L^{p_1}} \|u_2\|_{L^{p_2}}\,.$$

Show that this implies that the generalized Hölder's inequality, which states that if for $j = 1, ..., m$ and $p_j \in [1, \infty]$ with $\sum_{j=1}^{m} \frac{1}{p_j} = 1$, then

$$\int_{\mathbb{R}^n} |u_1 \cdots u_m|\, dx \leqslant \|u_1\|_{L^{p_1}} \cdots \|u_m\|_{L^{p_m}}\,.$$

PROBLEM 1.8. (a) Let f_n and g_n denote two sequences in $L^p(\Omega)$ with $1 \leqslant p \leqslant \infty$ such that $f_n \to f$ in $L^p(\Omega)$ $g_n \to g$ in $L^p(\Omega)$. Set $h_n = \max\{f_n, g_n\}$ and prove that $h_n \to h$ in $L^p(\Omega)$, where $h = \max\{f, g\}$.

(b) Let f_n be a sequence in $L^p(\Omega)$ with $1 \leqslant p < \infty$ and let g_n be a bounded sequence in $L^\infty(\Omega)$. Assume that $f_n \to f$ in $L^p(\Omega)$ and that $g_n \to g$ a.e. Prove that $f_n g_n \to fg$ in $L^p(\Omega)$.

PROBLEM 1.9. Let $1 \leqslant p < \infty$ and $1 \leqslant q \leqslant \infty$.

(a) Prove that $L^1(\Omega) \cap L^\infty(\Omega)$ is a dense subset of $L^p(\Omega)$.

(b) Prove that the set $\left\{f \in L^p(\Omega) \cap L^q(\Omega) \,\middle|\, \|f\|_{L^q(\Omega)} \leqslant 1\right\}$ is closed in $L^p(\Omega)$.

(c) Let f_n be a sequence in $L^p(\Omega) \cap L^q(\Omega)$ and let $f \in L^p(\Omega)$. Assume that

$$f_n \to f \ \text{ in } L^p(\Omega) \ \text{ and } \ \|f_n\|_{L^q(\Omega)} \leqslant C\,.$$

Prove that $f \in L^r(\Omega)$ and that $f_n \to f$ in $L^r(\Omega)$ for every r between p and q, $r \neq q$.

PROBLEM 1.10. Assume $|\Omega| < \infty$.

(a) Let $f \in \bigcap_{1 \leqslant p < \infty} L^p(\Omega)$ and assume that there is a constant C such that

$$\|f\|_{L^p(\Omega)} \leqslant C \quad \forall\, 1 \leqslant p < \infty\,.$$

Prove that $f \in L^\infty(\Omega)$.

(b) Construct an example of a function $f \in \bigcap_{1 \leqslant p < \infty} L^p(\Omega)$ such that $f \notin L^\infty(\Omega)$ with $\Omega = (0,1)$.

PROBLEM 1.11. Given $f \in L^1(\mathbb{S}^1)$, $0 < r < 1$, define

$$P_r f(\theta) = \sum_{n=-\infty}^{\infty} \hat{f}_n r^{|n|} e^{in\theta}, \quad \widehat{f}_n = \frac{1}{2\pi} \int_0^{2\pi} f(\theta) e^{-in\theta} d\theta \,.$$

Show that

$$P_r f(\theta) = p_r * f(\theta) = \frac{1}{2\pi} \int_0^{2\pi} p_r(\theta - \phi) f(\phi) d\phi \,,$$

where

$$p_r(\theta) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{in\theta} = \frac{1 - r^2}{1 - 2r\cos\theta + r^2} \,.$$

Show that $\frac{1}{2\pi} \int_0^{2\pi} p_r(\theta) d\theta = 1$.

PROBLEM 1.12. If $f \in L^p(\mathbb{S}^1)$, $1 \leqslant p < \infty$, show that

$$P_r f \to f \ \text{ in } \ L^p(\mathbb{S}^1) \ \text{ as } \ r \nearrow 1 \,.$$

PROBLEM 1.13. Suppose that $Y = [0,1]^{\mathrm{n}}$ is the unit square in $\mathbb{R}^{\mathrm{n}}$ and let $a(y)$ denote a Y-periodic function in $L^\infty(\mathbb{R}^{\mathrm{n}})$. For $\epsilon > 0$, let $a_\epsilon(x) = a\left(\frac{x}{\epsilon}\right)$, and let $\bar{a} = \int_Y a(y)dy$ denote the average value of a. Prove that $a_\epsilon \overset{*}{\rightharpoonup} \bar{a}$ as $\epsilon \to 0$. Prove the same results for $L^\infty(\mathbb{R}^{\mathrm{n}})$ replaced by $L^p(\mathbb{R}^{\mathrm{n}})$, $p \geqslant 1$, and weak-$*$ convergence replaced by weak convergence.

PROBLEM 1.14. Let $f_n = \sqrt{n}\mathbf{1}_{(0,\frac{1}{n})}$. Prove that $f_n \rightharpoonup 0$ in $L^2(0,1)$, that $f_n \to 0$ in $L^1(0,1)$, but that f_n does not converge strongly in $L^2(0,1)$.

PROBLEM 1.15. Let $X \subseteq L^1(\Omega)$ denote a closed vector space in $L^1(\Omega)$, and suppose that $X \subseteq \bigcup_{1 < q \leqslant \infty} L^q(\Omega)$. Use the Baire category theorem (Theorem B.11) and the sets

$$X_n = \left\{ f \in X \cap L^{1+1/n}(\Omega) \,\middle|\, \|f\|_{L^{1+1/n}(\Omega)} \leqslant n \right\}, \quad n \in \mathbb{N} \,,$$

to prove that there exists some $p > 1$ for which $X \subseteq L^p(\Omega)$.

Problem 1.16. Let $v : \Omega \to \mathbb{R}$ denote a measurable function and suppose that for $1 \leqslant q \leqslant p < \infty$,

$$uv \in L^q(\Omega) \quad \forall\, u \in L^p(\Omega)\,.$$

Use the closed graph theorem (Theorem B.16) to prove that $v \in L^{\frac{pq}{p-q}}(\Omega)$.

Problem 1.17. Prove that the space $\mathscr{C}^0_c(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for any $p \in [1, \infty)$. (We use the notation $\mathscr{C}^0_c(\mathbb{R}^n)$ to denote the space of continuous functions on $\mathbb{R}^n$ with compact support.)

Problem 1.18. For $u \in \mathscr{C}^0(\mathbb{R}^n; \mathbb{R})$, $\operatorname{spt}(u)$ is the closure of the set $\{x \in \mathbb{R}^n \;:\; u(x) \neq 0\}$, but this definition may not make any sense for functions $u \in L^p(\Omega)$; for example, what is the support $\mathbf{1}_{\mathbb{Q}}$, the indicator over the rational numbers?

Let $u : \mathbb{R}^n \to \mathbb{R}$, and let $\{\Omega_\alpha\}_{\alpha \in A}$ denote the collection of all open sets on $\mathbb{R}^n$ such that for each $\alpha \in A$, $u = 0$ a.e. on Ω_α. Define $\Omega = \bigcup\limits_{\alpha \in A} \Omega_\alpha$. Prove that $u = 0$ a.e. on Ω.

The support of u, $\operatorname{spt}(u)$ is $\Omega^\complement$, the complement of Ω. Notice that if $v = w$ a.e. on $\mathbb{R}^n$, then $\operatorname{spt}(v) = \operatorname{spt}(w)$; furthermore, if $u \in \mathscr{C}^0(\mathbb{R}^n)$, then $\Omega^\complement = \overline{\big\{x \in \mathbb{R}^n \,\big|\, u(x) \neq 0\big\}}$.

(Hint. Since A is not necessary countable, it is not clear that $f = 0$ a.e. on Ω, so find a countable family U_n of open sets in $\mathbb{R}^n$ such that every open set on $\mathbb{R}^n$ is the union of some of the sets from $\{U_n\}$.)

Problem 1.19. Prove that if $u \in L^1(\mathbb{R}^n)$ and $v \in L^p(\mathbb{R}^n)$ for $1 \leqslant p \leqslant \infty$, then

$$\operatorname{spt}(u * v) \subseteq \overline{\operatorname{spt}(u) + \operatorname{spt}(v)}\,.$$

Problem 1.20. (a) Let $u \in \mathscr{C}^0_c(\mathbb{R}^n)$ and $v \in L^1_{\text{loc}}(\mathbb{R}^n)$. Show that $u * v$ is well-defined for all $x \in \mathbb{R}^n$ and that $u * v \in \mathscr{C}(\mathbb{R}^n)$.

(b) If for $k \in \mathbb{N}$, $u \in \mathscr{C}^k_c(\mathbb{R}^n)$, then show that $u * v \in \mathscr{C}^k(\mathbb{R}^n)$ and that $D^\alpha(u * v) = (D^\alpha u) * v$ for all $\alpha \in \mathbb{Z}^n_+$ with $|\alpha| \leqslant k$.

Problem 1.21. (a) If $u \in L^p(\mathbb{R}^n)$ for $1 \leqslant p < \infty$, and $u^\epsilon = \eta_\epsilon * u$, show that $u^\epsilon \to u$ in $L^p(\mathbb{R}^n)$ as $\epsilon \to 0$.

(b) Let Ω denote an open and smooth subset of $\mathbb{R}^n$. Prove that $\mathscr{C}^\infty_c(\Omega)$ is dense in $L^p(\Omega)$ for $1 \leqslant p < \infty$.

Problem 1.22. Prove that if $u \in L^1_{\text{loc}}(\Omega)$ satisfies $\displaystyle\int_\Omega u(x)v(x)dx = 0$ for all $v \in \mathscr{C}^\infty_c(\Omega)$, then $u = 0$ a.e. in Ω.

PROBLEM 1.23. For $w : \mathbb{R} \to \mathbb{R}$, define the sequence $u_n(x) = w(x+n)$.

(a) Suppose that $w \in L^p(\mathbb{R})$ for $1 < p < \infty$. Prove that $u_n \rightharpoonup 0$ in $L^p(\mathbb{R})$.

(b) Suppose $w \in L^\infty(\mathbb{R})$. For $\delta > 0$, define

$$E_\delta = \left\{ x \in \mathbb{R} \,\middle|\, |w(x)| > \delta \right\}.$$

Suppose that $w(x) \to 0$ as $|x| \to 0$ in the following weak sense: $|E_\delta| < \infty$ for all $\delta > 0$. Prove that $u_n \overset{*}{\rightharpoonup} 0$ in $L^\infty(\mathbb{R})$.

(c) For $w = \mathbf{1}_{(0,1)}$, prove that there does not exist a subsequence u_{n_k} that converges weakly in $L^1(\mathbb{R})$. (Hint. Argue by contradiction, and use a piecewise constant test function that alternates sign on each adjacent interval.)

PROBLEM 1.24. Let $u \in L^\infty(\mathbb{R}^{\mathrm{n}})$ and let η_ϵ be the mollifiers from Definition 1.38. For $\epsilon > 0$ consider the sequence $\psi_\epsilon \in L^\infty(\mathbb{R}^{\mathrm{n}})$ such that

$$\|\psi_\epsilon\|_{L^\infty(\mathbb{R}^{\mathrm{n}})} \leqslant 1 \quad \forall\, \epsilon > 0 \quad \text{and} \quad \psi_\epsilon \to \psi \text{ a.e. in } \mathbb{R}^{\mathrm{n}},$$

and define

$$v^\epsilon = \eta_\epsilon * (\psi_\epsilon\, u) \quad \text{and} \quad v = \psi\, u\,.$$

(a) Prove that $v^\epsilon \overset{*}{\rightharpoonup} v$ in $L^\infty(\mathbb{R}^{\mathrm{n}})$.

(b) Prove that $v^\epsilon \to v$ in $L^1(B)$ for every ball $B \subseteq \mathbb{R}^{\mathrm{n}}$.

PROBLEM 1.25.

(a) For $u \in L^\infty(\Omega)$, $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$, prove that there exists a sequence $u_n \in \mathscr{C}_c^\infty(\Omega)$ such that

1. $\|u_n\|_{L^\infty(\Omega)} \leqslant \|u\|_{L^\infty(\Omega)}$ for all $n \in \mathbb{N}$;
2. $u_n \to u$ a.e. on Ω;
3. $u_n \overset{*}{\rightharpoonup} u$ in $L^\infty(\Omega)$.

(b) If $u \geqslant 0$ a.e. in Ω, show that the sequence u_n constructed above can be chosen to satisfy

4. $u_n \geqslant 0$ a.e. in Ω.

(c) Show that $\mathscr{C}_c^\infty(\Omega)$ is dense in $L^\infty(\Omega)$ with respect to the weak-$*$ topology.

Chapter 2

Introduction to Sobolev Spaces

2.1 Integration Formulas in Multiple Dimensions

The divergence theorem and its corollaries are fundamental to analysis in multiple space dimensions.

Theorem 2.1 (Divergence Theorem). *Let $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ be a Lipschitz domain; that is, $\partial\Omega$ locally is the graph of a Lipschitz function, and $w = (w_1, \cdots, w_{\mathrm{n}}) \in \mathscr{C}^1(\bar{\Omega})$ with outward pointing normal N. Then*

$$\int_\Omega \mathrm{div} w \, dx = \int_{\partial\Omega} w \cdot \mathrm{N} dS \, .$$

Now suppose that f is a scalar $\mathscr{C}^1$-function, and $\mathrm{N} = (\mathrm{N}_1, \cdots, \mathrm{N}_{\mathrm{n}})$. By setting $w = fe_i$, where e_i is the unit vector pointing to the positive x_i-axis, then the divergence theorem implies

$$\int_\Omega f_{x_i} \, dx = \int_{\partial\Omega} f \mathrm{N}^i dS \, .$$

Suppose further that f is the product of two $\mathscr{C}^1$-functions h and g; then, the equality above shows that

$$\int_\Omega g h_{x_i} \, dx = \int_{\partial\Omega} g h \mathrm{N}^i dS - \int_\Omega g_{x_i} h \, dx \, .$$

This is the multi-dimensional version of integration-by-parts.

Let Ω be a domain for which the divergence theorem holds and let $u \in \mathscr{C}^2(\bar{\Omega})$ and $v \in \mathscr{C}^1(\bar{\Omega})$-functions. Then we have **Green's first identity**:

$$\int_\Omega \nabla v \cdot \nabla u \, dx + \int_\Omega v \Delta u \, dx = \int_\Omega \mathrm{div}(v \nabla u) \, dx = \int_{\partial\Omega} v \frac{\partial u}{\partial \mathrm{N}} dS \, , \tag{2.1}$$

where $\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_{\mathrm{n}}^2}$. Suppose $v \in \mathscr{C}^2(\bar{\Omega})$ as well. Interchanging u and v in (2.1) and forming the difference of the two equalities, we obtain **Green's second identity**:

$$\int_\Omega (v\Delta u - u\Delta v)\, dx = \int_{\partial\Omega} \left[v\frac{\partial u}{\partial \mathrm{N}} - u\frac{\partial v}{\partial \mathrm{N}}\right] dS \tag{2.2}$$

2.2 Weak Derivatives

DEFINITION 2.2 (Test functions)**.** For $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$, set

$$\mathscr{C}_c^\infty(\Omega) = \left\{u \in \mathscr{C}^\infty(\Omega) \,\middle|\, \mathrm{spt}(u) \subseteq \mathcal{V} \subset\subset \Omega\right\},$$

the collection of smooth functions with compact support. Traditionally $\mathscr{D}(\Omega)$ is often used to denote $\mathscr{C}_c^\infty(\Omega)$, and $\mathscr{D}(\Omega)$ is often referred to as the *space of test functions*.

For $u \in \mathscr{C}^1(\mathbb{R})$, we can define $\frac{du}{dx}$ by the integration-by-parts formula; namely,

$$\int_{\mathbb{R}} \frac{du}{dx}(x)\varphi(x)\, dx = -\int_{\mathbb{R}} u(x)\frac{d\varphi}{dx}(x)\, dx \qquad \forall\, \varphi \in \mathscr{C}_c^\infty(\mathbb{R})\,. \tag{2.3}$$

Notice, however, that the right-hand side is well-defined, whenever $u \in L^1_{\mathrm{loc}}(\mathbb{R})$

DEFINITION 2.3. An element $\alpha \in \mathbb{N}^{\mathrm{n}}$ (nonnegative integers) is called a multi-index. For such an $\alpha = (\alpha_1, ..., \alpha_n)$, we write $D^\alpha = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_{\mathrm{n}}^{\alpha_n}}$ and $|\alpha| = \alpha_1 + \cdots + \alpha_n$.

EXAMPLE 2.4. Let $\mathrm{n} = 2$. If $|\alpha| = 0$, then $\alpha = (0,0)$; if $|\alpha| = 1$, then $\alpha = (1,0)$ or $\alpha = (0,1)$. If $|\alpha| = 2$, then $\alpha = (2,0), (1,1)$ or $(0,2)$.

DEFINITION 2.5 (Weak derivative)**.** Suppose that $u \in L^1_{\mathrm{loc}}(\Omega)$. Then $v^\alpha \in L^1_{\mathrm{loc}}(\Omega)$ is called the α^{th} weak derivative of u, written $v^\alpha = D^\alpha u$, if

$$\int_\Omega u(x)\, D^\alpha\varphi(x)\, dx = (-1)^{|\alpha|} \int_\Omega v^\alpha(x)\varphi(x)\, dx \qquad \forall\, \varphi \in \mathscr{C}_c^\infty(\Omega)\,.$$

EXAMPLE 2.6. Let $\mathrm{n} = 1$ and set $\Omega = (0,2)$. Define the function

$$u(x) = \begin{cases} x & 0 \leqslant x < 1\,, \\ 1 & 1 \leqslant x \leqslant 2\,. \end{cases}$$

Then the function

$$v(x) = \begin{cases} 1 & 0 \leqslant x < 1\,, \\ 0 & 1 \leqslant x \leqslant 2\,, \end{cases}$$

is the weak derivative of u. To see this, note that for $\varphi \in \mathscr{C}_c^\infty\big((0,2)\big)$,

$$\begin{aligned}\int_0^2 u(x)\frac{d\varphi}{dx}(x)\,dx &= \int_0^1 x\frac{d\varphi}{dx}(x)\,dx + \int_1^2 \frac{d\varphi}{dx}(x)\,dx \\ &= -\int_0^1 \varphi(x)\,dx + x\varphi(x)\Big|_{x=0}^{x=1} + \varphi(x)\Big|_{x=1}^{x=2} = -\int_0^1 \varphi(x)\,dx \\ &= -\int_0^2 v(x)\varphi(x)\,dx\,.\end{aligned}$$

Example 2.7. Let n $= 1$ and set $\Omega = (0,2)$. Define the function

$$u(x) = \begin{cases} x & 0 \leqslant x < 1\,, \\ 2 & 1 \leqslant x \leqslant 2\,. \end{cases}$$

Then the weak derivative <u>does not</u> exist!

To prove this, assume for the sake of contradiction that there exists $v \in L^1_{\rm loc}(\Omega)$ such that for all $\varphi \in \mathscr{C}_c^\infty\big((0,2)\big)$,

$$\int_0^2 v(x)\varphi(x)\,dx = -\int_0^2 u(x)\frac{d\varphi}{dx}(x)\,dx\,.$$

Then

$$\begin{aligned}\int_0^2 v(x)\varphi(x)\,dx &= -\int_0^1 x\frac{d\varphi}{dx}(x)\,dx - 2\int_1^2 \frac{d\varphi}{dx}(x)\,dx \\ &= \int_0^1 \varphi(x)\,dx - \varphi(1) + 2\varphi(1) = \int_0^1 \varphi(x)\,dx + \varphi(1)\,.\end{aligned}$$

Suppose that $\{\varphi_j\}_{j=1}^\infty$ is a sequence in $\mathscr{C}_c^\infty(0,2)$ such that $\varphi_j(1) = 1$ and $\varphi_j(x) \to 0$ for $x \neq 1$. Then

$$1 = \varphi_j(1) = \int_0^2 v(x)\varphi_j(x)\,dx - \int_0^1 \varphi_j(x)\,dx \to 0\,,$$

which provides the contradiction.

Definition 2.8. For $p \in [1,\infty]$, define

$$W^{1,p}(\Omega) = \big\{u \in L^p(\Omega)\,\big|\,\text{weak derivative of } u \text{ exists, and } Du \in L^p(\Omega)\big\}\,,$$

where Du is the weak derivative of u.

EXAMPLE 2.9. Let n $= 1$ and set $\Omega = (0,1)$. Define the function $u(x) = \sin\frac{1}{x}$. Then $u \in L^1(0,1)$ and $\frac{du}{dx} = -\frac{\cos(1/x)}{x^2} \in L^1_{\mathrm{loc}}(0,1)$, but $u \notin W^{1,p}(\Omega)$ for any p.

DEFINITION 2.10. In the case $p = 2$, we set $H^1(\Omega) = W^{1,p}(\Omega)$.

EXAMPLE 2.11. Let $\Omega = B(0,1) \subseteq \mathbb{R}^2$ and set $u(x) = |x|^{-\alpha}$. We want to determine the values of α for which $u \in H^1(\Omega)$.

Since $|x|^{-\alpha} = \sum_{j=1}^{3}(x_j x_j)^{-\alpha/2}$, then $\frac{\partial}{\partial x_i}|x|^{-\alpha} = -\alpha|x|^{-\alpha-2}x_i$ is well-defined away from $x = 0$.

Step 1. We show that $u \in L^1_{\mathrm{loc}}(\Omega)$. To see this, note that $\int_\Omega |x|^{-\alpha}\,dx = \int_0^{2\pi}\int_0^1 r^{-\alpha} r\,dr\,d\theta < \infty$ whenever $\alpha < 2$.

Step 2. Set the vector $v(x) = -\alpha|x|^{-\alpha-2}x$ (so that each component is given by $v_i(x) = -\alpha|x|^{-\alpha-2}x_i$). We show that

$$\int_{B(0,1)} u(x)D\varphi(x)\,dx = -\int_{B(0,1)} v(x)\varphi(x)\,dx \quad \forall\, \varphi \in \mathscr{C}_c^\infty(B(0,1))\,.$$

To see this, let $\Omega_\delta = B(0,1) - B(0,\delta)$, let n denote the unit normal to $\partial\Omega_\delta$ (pointing toward the origin). Integration by parts yields

$$\int_{\Omega_\delta} |x|^{-\alpha}D\varphi(x)\,dx = \int_0^{2\pi} \delta^{-\alpha}\varphi(x)n(x)\delta d\theta + \alpha\int_{\Omega_\delta} |x|^{-\alpha-2}x\,\varphi(x)\,dx\,.$$

Since $\lim_{\delta\to0}\delta^{1-\alpha}\int_0^{2\pi}\varphi(x)n(x)d\theta = 0$ if $\alpha < 1$, we see that

$$\lim_{\delta\to0}\int_{\Omega_\delta} |x|^{-\alpha}D\varphi(x)\,dx = \lim_{\delta\to0}\alpha\int_{\Omega_\delta} |x|^{-\alpha-2}x\,\varphi(x)\,dx$$

Since $\int_0^{2\pi}\int_0^1 r^{-\alpha-1} r\,dr\,d\theta < \infty$ if $\alpha < 1$, the Dominated Convergence Theorem shows that v is the weak derivative of u.

Step 3. $v \in L^2(\Omega)$, whenever $\int_0^{2\pi}\int_0^1 r^{-2\alpha-2} r\,dr\,d\theta < \infty$ which holds if $\alpha < 0$.

REMARK 2.12. Note that if the weak derivative exists, it is unique. To see this, suppose that both v_1 and v_2 are the weak derivative of u on Ω. Then $\int_\Omega (v_1 - v_2)\varphi\,dx = 0$ for all $\varphi \in \mathscr{C}_c^\infty(\Omega)$, so that $v_1 = v_2$ a.e.

Theorem 2.13 (Product rule). *For $u \in W^{k,p}(\Omega)$ and $\zeta \in \mathscr{C}_c^\infty(\Omega)$, the product $\zeta u \in W^{k,p}(\Omega)$ and*

$$D^\alpha(\zeta u) = \sum_{|\beta| \leqslant |\alpha|} \binom{\alpha}{\beta} D^\beta \zeta \, D^{\alpha-\beta} u \,, \tag{2.4}$$

where $\displaystyle\binom{\alpha}{\beta} = \frac{|\alpha|!}{|\beta|! \, |\alpha - \beta|!}$.

Proof. We begin with the case that $|\alpha| = 1$. We suppose that v^α is the αth weak derivative of u. Then, for all test funtions $\varphi \in \mathscr{C}_c^\infty(\Omega)$,

$$\int_\Omega \zeta u \, D^\alpha \varphi \, dx = \int_\Omega [u D^\alpha(\zeta\varphi) - u(D^\alpha \zeta)\varphi] \, dx = \int_\Omega [-\zeta v^\alpha - u D^\alpha \zeta] \, \varphi \, dx \,,$$

where we have used the fact that $\zeta\varphi \in \mathscr{C}_c^\infty(\Omega)$.

Having established (2.4) for $|\alpha| = 1$, we now use an induction argument. Assume that (2.4) holds for all $|\alpha| \leqslant \ell$ and all functions $\zeta \in \mathscr{C}^\infty(\Omega)$. Choose a multi-index α with $|\alpha| = \ell + 1$. Then $\alpha = \beta + \gamma$ for some $|\beta| = \ell$, $|\gamma| = 1$. Then for φ as above,

$$\begin{aligned}
\int_\Omega \zeta u D^\alpha \varphi \, dx = \int_\Omega \zeta u D^\beta (D^\gamma \varphi) \, dx &= (-1)^{|\beta|} \int_\Omega \sum_{|\sigma| \leqslant |\beta|} \binom{\beta}{\sigma} D^\sigma \zeta D^{\beta-\sigma} u D^\gamma \varphi \, dx \\
&= (-1)^{|\beta|+|\gamma|} \int_\Omega \sum_{|\sigma| \leqslant |\beta|} \binom{\beta}{\sigma} D^\gamma (D^\sigma \zeta D^{\beta-\sigma} u) \varphi \, dx \\
&= (-1)^{|\alpha|} \int_\Omega \sum_{|\sigma| \leqslant |\beta|} \binom{\beta}{\sigma} \big[D^\rho \zeta D^{\alpha-\rho} u + D^\sigma \zeta D^{\alpha-\sigma} u \big] \varphi \, dx \\
&= (-1)^{|\alpha|} \int_\Omega \Big[\sum_{|\sigma| \leqslant |\alpha|} \binom{\alpha}{\sigma} D^\sigma \zeta D^{\alpha-\sigma} u \Big] \varphi \, dx \,,
\end{aligned}$$

where $\rho = \sigma + \gamma$ in the fourth equality, and the fifth equality follows since

$$\binom{\beta}{\sigma - \gamma} + \binom{\beta}{\sigma} = \binom{\alpha}{\sigma} .$$

□

2.3 Definition of Sobolev Spaces

Definition 2.14. For integers $k \geqslant 0$ and $1 \leqslant p \leqslant \infty$,

$$W^{k,p}(\Omega) = \big\{ u \in L^1_{\text{loc}}(\Omega) \,\big|\, D^\alpha u \text{ exists and is in } L^p(\Omega) \text{ for } |\alpha| \leqslant k \big\}.$$

Definition 2.15. For $u \in W^{k,p}(\Omega)$ define

$$\|u\|_{W^{k,p}(\Omega)} = \Big(\sum_{|\alpha|\leqslant k} \|D^\alpha u\|^p_{L^p(\Omega)}\Big)^{\frac{1}{p}} \text{ for } 1 \leqslant p < \infty\,,$$

and

$$\|u\|_{W^{k,\infty}(\Omega)} = \sum_{|\alpha|\leqslant k} \|D^\alpha u\|_{L^\infty(\Omega)}\,.$$

The function $\|\cdot\|_{W^{k,p}(\Omega)}$ is clearly a norm since it is a finite sum of L^p norms.

Definition 2.16. A sequence $u_j \to u$ in $W^{k,p}(\Omega)$ if $\lim_{j\to\infty} \|u_j - u\|_{W^{k,p}(\Omega)} = 0$.

Theorem 2.17. *$W^{k,p}(\Omega)$ is a Banach space.*

Proof. Let u_j denote a Cauchy sequence in $W^{k,p}(\Omega)$. It follows that for all $|\alpha| \leqslant k$, $D^\alpha u_j$ is a Cauchy sequence in $L^p(\Omega)$. Since $L^p(\Omega)$ is a Banach space (see Theorem 1.32), for each α there exists $u^\alpha \in L^p(\Omega)$ such that

$$D^\alpha u_j \to u^\alpha \quad \text{in} \quad L^p(\Omega)\,.$$

When $\alpha = (0, ..., 0)$ we set $u := u^{(0,\dots,0)}$ so that $u_j \to u$ in $L^p(\Omega)$. We must show that $u^\alpha = D^\alpha u$.

For each $\varphi \in \mathscr{C}^\infty_c(\Omega)$,

$$\begin{aligned}\int_\Omega u D^\alpha \varphi\, dx &= \lim_{j\to\infty} \int_\Omega u_j D^\alpha \varphi\, dx = (-1)^{|\alpha|} \lim_{j\to\infty} \int_\Omega D^\alpha u_j \varphi\, dx \\ &= (-1)^{|\alpha|} \int_\Omega u^\alpha \varphi\, dx\,;\end{aligned}$$

thus, $u^\alpha = D^\alpha u$ and hence $D^\alpha u_j \to D^\alpha u$ in $L^p(\Omega)$ for each $|\alpha| \leqslant k$, which shows that $u_j \to u$ in $W^{k,p}(\Omega)$. □

Definition 2.18. For integers $k \geqslant 0$ and $p = 2$, we define

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

$H^k(\Omega)$ is a Hilbert space with inner-product given by

$$(u, v)_{H^k(\Omega)} = \sum_{|\alpha|\leqslant k} (D^\alpha u, D^\alpha v)_{L^2(\Omega)}\,.$$

2.4 A Simple Version of the Sobolev Embedding Theorem

For two Banach spaces $\mathbb{B}_1$ and $\mathbb{B}_2$, we say that $\mathbb{B}_1$ is continuously embedded in $\mathbb{B}_2$, denoted by $\mathbb{B}_1 \hookrightarrow \mathbb{B}_2$, if $\|u\|_{\mathbb{B}_2} \leqslant C\|u\|_{\mathbb{B}_1}$ for some constant C and for $u \in \mathbb{B}_1$. We wish to determine which Sobolev spaces $W^{k,p}(\Omega)$ can be continuously embedded in the space of continuous functions. To motivate the type of analysis that is to be employed, we study a special case.

Theorem 2.19 (Sobolev embedding in 2-D)**.** *For $kp > 2$,*

$$\max_{x \in \mathbb{R}^2} |u(x)| \leqslant C\|u\|_{W^{k,p}(\mathbb{R}^2)} \qquad \forall\, u \in \mathscr{C}_c^\infty(\Omega)\,. \tag{2.5}$$

Proof. Given $u \in \mathscr{C}_c^\infty(\Omega)$, we prove that for all $x \in \operatorname{spt}(u)$,

$$|u(x)| \leqslant C\|D^\alpha u(x)\|_{L^p(\Omega)} \quad \forall\, |\alpha| \leqslant k\,.$$

By choosing a coordinate system centered about x, we can assume that $x = 0$; thus, it suffices to prove that

$$|u(0)| \leqslant C\|D^\alpha u(x)\|_{L^p(\Omega)} \quad \forall\, |\alpha| \leqslant k\,.$$

Let $g \in \mathscr{C}^\infty([0,\infty))$ with $0 \leqslant g \leqslant 1$, such that $g(x) = 1$ for $x \in \left[0, \frac{1}{2}\right]$ and $g(x) = 0$ for $x \in \left[\frac{3}{4}, \infty\right)$.

By the fundamental theorem of calculus,

$$\begin{aligned} u(0) &= -\int_0^1 \partial_r[g(r)u(r,\theta)]dr = -\int_0^1 \partial_r r\, \partial_r[g(r)u(r,\theta)]dr \\ &= \int_0^1 r\, \partial_r^2[g(r)u(r,\theta)]dr = \frac{(-1)^k}{(k-1)!}\int_0^1 r^{k-1}\, \partial_r^k[g(r)u(r,\theta)]dr \\ &= \frac{(-1)^k}{(k-1)!}\int_0^1 r^{k-2}\, \partial_r^k[g(r)u(r,\theta)]r dr\,. \end{aligned}$$

Integrating both sides from 0 to 2π, we see that

$$u(0) = \frac{(-1)^k}{2\pi(k-1)!}\int_0^{2\pi}\int_0^1 r^{k-2}\, \partial_r^k[g(r)u(r,\theta)]r dr d\theta\,.$$

The change of variables from Cartesian to polar coordinates is given by

$$x(r,\theta) = r\cos\theta\,, \quad y(r,\theta) = r\sin\theta\,.$$

By the chain rule,

$$
\begin{aligned}
\partial_r u(x(r,\theta), y(r,\theta)) &= \partial_x u \cos\theta + \partial_y u \sin\theta\,, \\
\partial_r^2 u(x(r,\theta), y(r,\theta)) &= \partial_x^2 u \cos^2\theta + 2\partial_{xy}^2 u \cos\theta\sin\theta + \partial_y^2 u \sin^2\theta \\
&\ \ \vdots
\end{aligned}
$$

It follows that $\partial_r^k = \sum\limits_{|\alpha|\leqslant k} a_\alpha(\theta) D^\alpha$, where a_α consists of trigonometric polynomials of θ, so that

$$
\begin{aligned}
u(0) &= \frac{(-1)^k}{2\pi(k-1)!}\int_{B(0,1)} r^{k-2} \sum_{|\alpha|\leqslant k} a_\alpha(\theta) D^\alpha[g(r)u(x)]\,dx \\
&\leqslant C\|r^{k-2}\|_{L^q(B(0,1))} \sum_{|\alpha|\leqslant k} \|D^\alpha(gu)\|_{L^p(B(0,1))} \\
&\leqslant C\Big(\int_0^1 r^{\frac{p(k-2)}{p-1}} r dr\Big)^{\frac{p-1}{p}} \|u\|_{W^{k,p}(\mathbb{R}^2)}\,.
\end{aligned}
$$

Hence, we require $\dfrac{p(k-2)}{p-1} + 1 > -1$ or $kp > 2$. □

2.5 Approximation of $W^{k,p}(\Omega)$ by Smooth Functions

Define $\Omega_\epsilon = \big\{x \in \Omega \,\big|\, \operatorname{dist}(x, \partial\Omega) > \epsilon\big\}$.

Definition 2.20. A sequence $u_j \to u$ in $W^{k,p}_{\rm loc}(\Omega)$ if $u_j \to u$ in $W^{k,p}(\widetilde{\Omega})$ for each $\widetilde{\Omega} \subset\subset \Omega$.

Theorem 2.21 (local approximation)**.** *For integers $k \geqslant 0$ and $1 \leqslant p < \infty$, let*

$$
u^\epsilon = \eta_\epsilon * u \quad in \quad \Omega_\epsilon\,,
$$

where η_ϵ is the standard mollifier defined in Definition 1.38*. Then*

(A) *$u^\epsilon \in \mathscr{C}^\infty(\Omega_\epsilon)$ for each $\epsilon > 0$, and*

(B) *$u^\epsilon \to u$ in $W^{k,p}_{\rm loc}(\Omega)$ as $\epsilon \to 0$.*

Proof. Theorem 1.42 proves part (A). Next, let v^α denote the the α-th weak partial derivative of u. To prove part (B), we show that $D^\alpha u^\epsilon = \eta_\epsilon * v^\alpha$ in Ω_ϵ. For $x \in \Omega_\epsilon$,

$$\begin{aligned} D^\alpha u^\epsilon(x) = D^\alpha \int_\Omega \eta_\epsilon(x-y)u(y)dy &= \int_\Omega D_x^\alpha \eta_\epsilon(x-y)u(y)dy \\ &= (-1)^{|\alpha|} \int_\Omega D_y^\alpha \eta_\epsilon(x-y)u(y)dy \\ &= \int_\Omega \eta_\epsilon(x-y)v^\alpha(y)dy = (\eta_\epsilon * v^\alpha)(x)\,. \end{aligned}$$

By part (D) of Theorem 1.42, $D^\alpha u^\epsilon \to v^\alpha$ in $L^p_{\rm loc}(\Omega)$. □

We next consider the case that Ω is bounded, and some improvements of the above *local* approximation result.

Theorem 2.22 (Global approximation on Ω)**.** *For $\Omega \subseteq \mathbb{R}^{\rm n}$ open and bounded, and for $u \in W^{k,p}(\Omega)$, $1 \leqslant p < \infty$, there exists functions $u_j \in \mathscr{C}^\infty(\Omega) \cap W^{k,p}(\Omega)$ such that $u_j \to u$ in $W^{k,p}(\Omega)$.*

Proof. For $k = 1, 2, 3, ...,$ we define the open set

$$\Omega_k = \left\{ x \in \Omega \,\middle|\, \operatorname{dist}(x, \partial\Omega) > \frac{1}{k} \right\},$$

so that $\Omega = \bigcup_{k=1}^\infty \Omega_k$. Next, we define the "annular" regions $\omega_k = \Omega_{k+3} \backslash \overline{\Omega_{k+1}}$. We choose an additional open set $\omega_0 \subset\subset \Omega$ such that $\Omega = \bigcup_{k=0}^\infty \omega_k$.

Let ζ_k denote a smooth partition of unity subordinate to the cover ω_k. By Theorem 2.13, $\zeta_k u \in W^{k,p}(\Omega)$, and $\operatorname{spt}(\zeta_k u) \subseteq \omega_k$. By Theorem 2.21, for each $\delta > 0$, we can choose ϵ_k sufficiently small so that

$$u^{\epsilon_k} = \eta_{\epsilon_k} * (\zeta_k u)$$

is smooth and satisfies

$$\|u^{\epsilon_k} - \zeta_k u\|_{W^{k,p}(\Omega)} \leqslant \frac{\delta}{2^{k+1}} \quad \text{for } k = 0, 1, 2, ...\,,$$

with $\operatorname{spt}(u^{\epsilon_k}) \subseteq \Omega_{k+4} \backslash \overline{\Omega_k}$.

We let $v = \sum_{k=0}^{\infty} u^{\epsilon_k}$. Since for each open set $\widetilde{\Omega} \subset\subset \Omega$, there are only finitely many nonzero terms in the sum, we see that $v \in \mathscr{C}^\infty(\Omega)$, and since $u = \sum_{k=0}^{\infty} \zeta_k u$, for each $\widetilde{\Omega} \subset\subset \Omega$

$$\|v - u\|_{W^{k,p}(\Omega)} \leqslant \sum_{k=0}^{\infty} \|u^{\epsilon_k} - \zeta_k u\|_{W^{k,p}(\Omega)} \leqslant \delta \sum_{k=0}^{\infty} \frac{1}{2^{k+1}} = \delta\,.$$

By taking the supremum over open sets $\widetilde{\Omega} \subset\subset \Omega$, we conclude that $\|v - u\|_{W^{k,p}(\Omega)} \leqslant \delta$. □

Theorem 2.23 (Global approximation on $\overline{\Omega}$)**.** *Suppose that $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ is a smooth, open, bounded subset, and that $u \in W^{k,p}(\Omega)$ for some $1 \leqslant p < \infty$ and $k \in \mathbb{N}$. Then there exists a sequence $\{u_j\}_{j=1}^{\infty} \subseteq \mathscr{C}^\infty(\bar{\Omega})$ such that*

$$u_j \to u \qquad \textit{in} \quad W^{k,p}(\Omega)\,.$$

Proof. We employ Theorem 2.43 (which will be proven below) to obtain an extension Eu of u such that

$$Eu = u \quad \text{in} \;\; \Omega\,, \quad \text{and} \quad \|Eu\|_{W^{k,p}(\mathbb{R}^{\mathrm{n}})} \leqslant C\|u\|_{W^{k,p}(\Omega)}\,.$$

Choose $v_j \in \mathscr{C}_c^\infty(\mathbb{R}^{\mathrm{n}})$ so that $v_j \to Eu$ in $W^{k,p}(\mathbb{R}^{\mathrm{n}})$, and define $u_j = v_j|_{\bar{\Omega}}$; that is, u_j is the restriction of v_j to $\bar{\Omega}$. Then clearly $u_j \in \mathscr{C}^\infty(\bar{\Omega})$, and

$$\|u_j - u\|_{W^{k,p}(\Omega)} \leqslant \|v_j - Eu\|_{W^{k,p}(\mathbb{R}^{\mathrm{n}})} \to 0 \quad \text{as} \;\; j \to \infty\,.$$ □

Remark 2.24. Using these global approximation theorems, it follows that the inequality (2.5) holds for all $u \in W^{k,p}(\mathbb{R}^2)$.

2.6 Hölder Spaces

Recall that for $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ open and smooth, the class of Lipschitz functions $u : \Omega \to \mathbb{R}$ satisfies the estimate

$$|u(x) - u(y)| \leqslant C|x - y| \;\; \forall\, x, y \in \Omega$$

for some constant C.

Definition 2.25 (Classical derivative)**.** A function $u : \Omega \to \mathbb{R}$ is differentiable at $x \in \Omega$ if there exists $f : \Omega \to \mathscr{L}(\mathbb{R}^n; \mathbb{R})$ such that

$$\frac{|u(x) - u(y) - f(x) \cdot (x - y)|}{|x - y|} \to 0\,.$$

We call $f(x)$ the classical derivative (or gradient) of $u(x)$, and denote it by $Du(x)$.

Definition 2.26. If $u : \Omega \to \mathbb{R}$ is bounded and continuous, then

$$\|u\|_{\mathscr{C}^0(\bar{\Omega})} = \max_{x \in \Omega} |u(x)|\,.$$

If in addition u has a continuous and bounded derivative, then

$$\|u\|_{\mathscr{C}^1(\bar{\Omega})} = \|u\|_{\mathscr{C}^0(\bar{\Omega})} + \|Du\|_{\mathscr{C}^0(\bar{\Omega})}\,.$$

The Hölder spaces *interpolate* between $\mathscr{C}^0(\bar{\Omega})$ and $\mathscr{C}^1(\bar{\Omega})$.

Definition 2.27. For $0 < \gamma \leqslant 1$, the space $\mathscr{C}^{0,\gamma}(\bar{\Omega})$ consists of those functions for which

$$\|u\|_{\mathscr{C}^{0,\gamma}(\bar{\Omega})} := \|u\|_{\mathscr{C}^0(\bar{\Omega})} + [u]_{\mathscr{C}^{0,\gamma}(\bar{\Omega})} < \infty\,,$$

where the γth Hölder semi-norm $[u]_{\mathscr{C}^{0,\gamma}(\bar{\Omega})}$ is defined as

$$[u]_{\mathscr{C}^{0,\gamma}(\bar{\Omega})} = \max_{\substack{x,y \in \Omega \\ x \neq y}} \left(\frac{|u(x) - u(y)|}{|x - y|^\gamma} \right).$$

The space $\mathscr{C}^{0,\gamma}(\bar{\Omega})$ is a Banach space.

2.7 Morrey's Inequality

We can now offer a refinement and extension of the simple version of the Sobolev Embedding Theorem 2.19.

Theorem 2.28 (Morrey's inequality)**.** *Let $B_r \subseteq \mathbb{R}^n$ denote a ball of radius r, and let $n < p \leqslant \infty$. For $x, y \in B_r$*

$$|u(x) - u(y)| \leqslant C|x - y|^{1 - \frac{n}{p}} \|Du\|_{L^p(B_r)} \qquad \forall\, u \in W^{1,p}(B_r)\,. \tag{2.6}$$

NOTATION 2.29 (Averaging)**.** Let $B(0,1) \subseteq \mathbb{R}^{\mathrm{n}}$. The volume of $B(0,1)$ is given by $\alpha_n = \dfrac{\pi^{\frac{\mathrm{n}}{2}}}{\Gamma(\frac{\mathrm{n}}{2}+1)}$ and the surface area is $|\mathbb{S}^{\mathrm{n}-1}| = n\alpha_n$. We define

$$\fint_{B(x,r)} f(y)dy = \frac{1}{\alpha_n r^{\mathrm{n}}} \int_{B(x,r)} f(y)dy$$
$$\fint_{\partial B(x,r)} f(y)dS = \frac{1}{n\alpha_n r^{\mathrm{n}-1}} \int_{\partial B(x,r)} f(y)dS\,.$$

LEMMA 2.30. *Let $B_r \subseteq \mathbb{R}^{\mathrm{n}}$ denote a ball of radius r and let $u \in \mathscr{C}^1(\overline{B_r}) \cap W^{1,p}(B_r)$ for $p > n$. Then, with $\bar{u} = \fint_{B_r} u(y)dy$, for all $x \in B_r$,*

$$|\bar{u} - u(x)| \leqslant Cr^{1-n/p}\|Du\|_{L^p(B_r)}\,. \tag{2.7}$$

Proof. By the fundamental theorem of calculus, for $y \in B_r$,

$$u(y) - u(x) = \int_0^1 \frac{d}{dt}u(x + t(y-x))dt = \int_0^1 Du(x + t(y-x)) \cdot (y-x)dt\,,$$

so that

$$|u(y) - u(x)| \leqslant 2r \int_0^1 |Du(x + t(y-x))|dt\,,$$

and hence

$$\fint_{B_r} |u(y) - u(x)|dy \leqslant \frac{2r}{|B_r|} \int_{B_r} \int_0^1 |Du(x + t(y-x))|dtdy\,.$$

It follows that

$$\begin{aligned}|\bar{u} - u(x)| &\leqslant Cr^{1-n} \int_{B_r} \int_0^1 |Du(x + t(y-x))|dtdy \\ &\leqslant Cr^{1-n} \int_0^1 \int_{B_r} |Du(x + t(y-x))|dydt\,.\end{aligned}$$

We define the change of variable $z(y) = x + t(y-x)$ so that $|\det D_z y| = 1/t^n$. Then by the change-of-variables formula,

$$|\bar{u} - u(x)| \leqslant Cr^{1-n} \int_0^1 \int_{B_{tr}} |Du(z)|dz\; t^{-n}dt\,.$$

By Hölder's inequality,

$$\int_{B_{tr}} |Du(z)|dy \leqslant \|Du\|_{L^p(B_{tr})}|B_{tr}|^{1/q} \leqslant C\|Du\|_{L^p(B_r)}(tr)^{n/q}\,,$$

where $\frac{1}{q} = 1 - \frac{1}{p}$ is the conjugate exponent to p. Hence,

$$|\bar{u} - u(x)| \leqslant Cr^{1-n/p} \|Du\|_{L^p(B_r)} \int_0^1 t^{-n/p} dt \leqslant Cr^{1-n/p} \|Du\|_{L^p(B_r)} ,$$

the last inequality following when $p > n$. □

Proof of Theorem 2.28. Suppose that $u \in \mathscr{C}^1(\overline{B_r})$. By Lemma 2.30,

$$|\bar{u} - u(x)|dy \leqslant Cr^{1-n/p} \|Du\|_{L^p(B_r)} \qquad \forall\, x \in B_r ,$$
$$|\bar{u} - u(y)|dy \leqslant Cr^{1-n/p} \|Du\|_{L^p(B_r)} \qquad \forall\, y \in B_r .$$

It follows from the triangle inequality that

$$|u(x) - u(y)|dy \leqslant Cr^{1-n/p} \|Du\|_{L^p(B_r)} \qquad \forall\, x, y \in B_r . \tag{2.8}$$

Given any two points $x, y \in \mathbb{R}^{\mathrm{n}}$, there exists a ball B_r of radius $r = |x - y|$ containing x and y, which proves (2.6) for $u \in \mathscr{C}^1(\overline{B_r})$. For $u \in W^{1,p}(B_r)$, we use a Theorem 2.23, which provides a sequence $u^\epsilon \in \mathscr{C}^\infty(\overline{B_r})$ such that $u^\epsilon \to u$ in $W^{1,p}(B_r)$. □

Morrey's inequality implies the following embedding theorem.

Theorem 2.31 (Sobolev embedding theorem for $k = 1$). *There exists a constant $C = C(p, \mathrm{n})$ such that*

$$\|u\|_{\mathscr{C}^{0,1-\frac{\mathrm{n}}{p}}(\mathbb{R}^{\mathrm{n}})} \leqslant C\|u\|_{W^{1,p}(\mathbb{R}^{\mathrm{n}})} \qquad \forall\, u \in W^{1,p}(\mathbb{R}^{\mathrm{n}}) .$$

Proof. First assume that $u \in \mathscr{C}_c^1(\mathbb{R}^{\mathrm{n}})$. Given Morrey's inequality, it suffices to show that $\max |u| \leqslant C\|u\|_{W^{1,p}(\mathbb{R}^{\mathrm{n}})}$. Using Lemma 2.30, for all $x \in \mathbb{R}^{\mathrm{n}}$,

$$\begin{aligned} |u(x)| &\leqslant \left| u(x) - \fint_{B(x,1)} u(y)dy \right| + \fint_{B(x,1)} |u(y)|dy \\ &\leqslant C\|Du\|_{L^p(\mathbb{R}^{\mathrm{n}})} + C\|u\|_{L^p(\mathbb{R}^{\mathrm{n}})} \leqslant C\|u\|_{W^{1,p}(\mathbb{R}^{\mathrm{n}})} , \end{aligned}$$

the first inequality following whenever $p > \mathrm{n}$. Thus,

$$\|u\|_{\mathscr{C}^{0,1-\frac{\mathrm{n}}{p}}(\mathbb{R}^{\mathrm{n}})} \leqslant C\|u\|_{W^{1,p}(\mathbb{R}^{\mathrm{n}})} \qquad \forall\, u \in \mathscr{C}_c^1(\mathbb{R}^{\mathrm{n}}) . \tag{2.9}$$

By the density of $\mathscr{C}_c^\infty(\mathbb{R}^{\mathrm{n}})$ in $W^{1,p}(\mathbb{R}^{\mathrm{n}})$, there is a sequence $\{u_j\}_{j=1}^\infty \subseteq \mathscr{C}_c^\infty(\mathbb{R}^{\mathrm{n}})$ such that

$$u_j \to u \in W^{1,p}(\mathbb{R}^{\mathrm{n}}) .$$

By (2.9), for $j,k \in \mathbb{N}$,

$$\|u_j - u_k\|_{\mathscr{C}^{0,1-\frac{\mathrm{n}}{p}}(\mathbb{R}^\mathrm{n})} \leqslant C\|u_j - u_k\|_{W^{1,p}(\mathbb{R}^\mathrm{n})}\,.$$

Since $\mathscr{C}^{0,1-\frac{\mathrm{n}}{p}}(\mathbb{R}^\mathrm{n})$ is a Banach space, there exists a $U \in \mathscr{C}^{0,1-\frac{\mathrm{n}}{p}}(\mathbb{R}^\mathrm{n})$ such that

$$u_j \to U \quad \text{in} \quad \mathscr{C}^{0,1-\frac{\mathrm{n}}{p}}(\mathbb{R}^\mathrm{n})\,.$$

It follows that $U = u$ a.e. in Ω. By the continuity of norms with respect to strong convergence, we see that

$$\|U\|_{\mathscr{C}^{0,1-\frac{\mathrm{n}}{p}}(\mathbb{R}^\mathrm{n})} \leqslant C\|u\|_{W^{1,p}(\mathbb{R}^\mathrm{n})}$$

which completes the proof. □

In proving the above embedding theorem, we established that for $p > \mathrm{n}$, we have the inequality

$$\|u\|_{L^\infty(\mathbb{R}^\mathrm{n})} \leqslant C\|u\|_{W^{1,p}(\mathbb{R}^\mathrm{n})}\,. \tag{2.10}$$

We will see later that (2.10), via a scaling argument, leads to the following important *interpolation inequality*: for $p > \mathrm{n}$,

$$\|u\|_{L^\infty(\mathbb{R}^\mathrm{n})} \leqslant C(\mathrm{n},p)\|Du\|^{\frac{\mathrm{n}}{p}}_{L^p(\mathbb{R}^\mathrm{n})}\|u\|^{\frac{p-n}{p}}_{L^p(\mathbb{R}^\mathrm{n})}\,.$$

Corollary 2.32 (Sobolev embedding theorem $kp > n$)**.** *There exists a constant* $C = C(k,p,n)$ *such that*

$$\|u\|_{\mathscr{C}^{k-[\frac{n}{p}]-1,\gamma}(\mathbb{R}^\mathrm{n})} \leqslant C\|u\|_{W^{k,p}(\mathbb{R}^\mathrm{n})} \qquad \forall\, u \in W^{k,p}(\mathbb{R}^\mathrm{n})\,,$$

where

$$\gamma = \begin{cases} \big[\frac{n}{p}\big] + 1 - \frac{n}{p} & \text{if } \frac{n}{p} \notin \mathbb{N}\,, \\ \text{any } \alpha \in \mathbb{R} \cap (0,1) & \text{if } \frac{n}{p} \in \mathbb{N}\,. \end{cases}$$

Proof. The proof follows immediately as a consequence of Theorem 2.31 applied to weak derivatives of u. □

Another important consequence of Morrey's inequality is the relationship between the weak and classical derivative of a function. We begin by recalling the definition of classical differentiability. A function $u : \mathbb{R}^\mathrm{n} \to \mathbb{R}^m$ is differentiable at a point x if there exists a linear operator $L : \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^m$ such that for each $\epsilon > 0$, there exists $\delta > 0$ with $|y - x| < \delta$ implying that

$$\|u(y) - u(x) - L(y - x)\| \leqslant \epsilon \|y - x\| \,.$$

When such an L exists, we write $Du(x) = L$ and call it the classical derivative.

As a consequence of Morrey's inequality, we extract information about the classical differentiability properties of weak derivatives.

Theorem 2.33 (Differentiability a.e.). *If $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$, $\mathrm{n} < p \leqslant \infty$ and $u \in W^{1,p}_{\mathrm{loc}}(\Omega)$, then u is differentiable a.e. in Ω, and its gradient equals its weak gradient almost everywhere.*

Proof. We first restrict $\mathrm{n} < p < \infty$. By a version Lebesgue's differentiation theorem, for almost every $x \in \Omega$,

$$\lim_{r \to 0} ⨍_{B(x,r)} |Du(x) - Du(z)|^p dz = 0 \,, \tag{2.11}$$

where Du denotes the weak derivative of u. Thus, for $r > 0$ sufficiently small, we see that

$$⨍_{B(x,r)} |Du(x) - Du(z)|^p dz < \epsilon \,.$$

Fix a point $x \in \Omega$ for which (2.11) holds, and define the function

$$w_x(y) = u(y) - u(x) - Du(x) \cdot (y - x) \,.$$

Notice that $w_x(x) = 0$ and that

$$D_y w_x(y) = Du(y) - Du(x) \,.$$

Set $r = |x - y|$. Since $|u(y) - u(x) - Du(x) \cdot (y - x)| = |w_x(y) - w_x(x)|$, an application of the inequality (2.8) that we obtained in the proof of Morrey's inequality then yields the estimate

$$\begin{aligned} \big|u(y) - u(x) - Du(x) \cdot (y - x)\big| &\leqslant C r^{1 - \frac{\mathrm{n}}{p}} \|Dw_x\|_{L^p(B(x,r))} \\ &\leqslant C r ⨍_{B(x,r)} |Du(y) - Du(x)|^p dz \leqslant C|x - y|\epsilon \end{aligned}$$

from which it follows that $Du(x)$ is the classical derivative of u at the point x.

The case that $p = \infty$ follows from the inclusion $W^{1,\infty}_{\mathrm{loc}}(\Omega) \subseteq W^{1,p}_{\mathrm{loc}}(\Omega)$ for all $1 \leqslant p < \infty$. □

Theorem 2.34. *Let Ω denote an open, bounded, and smooth domain of $\mathbb{R}^{\rm n}$, and let $u \in H^1(\Omega)$. Then u is absolutely continuous on almost all straight lines parallel to the coordinate axes. Moreover, the weak derivatives of u coincides with the classical derivative of u almost everywhere.*

Proof. It suffices to assume that $\Omega = \big\{x \in \mathbb{R}^{\rm n} \,\big|\, 0 < x_i < 1\,, 1 \leqslant i \leqslant {\rm n}\big\}$, and show that

$$u(x) = \underbrace{\int_0^{x_{\rm n}} \frac{\partial u}{\partial x_{\rm n}}(x',t)dt}_{\equiv v(x)} + \text{const}\,,$$

where the integrand $\dfrac{\partial u}{\partial x_{\rm n}}$ is the weak derivative of u with respect to $x_{\rm n}$.

Let $\omega = \big\{x \in \mathbb{R}^{{\rm n}-1} \,\big|\, 0 < x_i < 1\,, 1 \leqslant i \leqslant {\rm n}-1\big\}$ so that $\Omega = \omega \times (0,1)$, and let $\zeta \in \mathscr{C}_c^\infty(\omega)$ and $\varphi \in \mathscr{C}_c^\infty(0,1)$ be test functions. Since v is absolutely continuous in $x_{\rm n}$, integration by parts implies

$$\int_0^1 v(x',t)\varphi'(t)dt = -\int_0^1 v_{x_{\rm n}}(x',t)\varphi(t)dt\,,$$

where $v_{x_{\rm n}}$ denotes the classical derivative of v with respect to $x_{\rm n}$. Multiplying both sides by $\zeta(x')$ and integrating over ω, we find that

$$\int_\Omega v(x)\zeta(x')\varphi'(x_{\rm n})\,dx = -\int_\Omega v_{x_{\rm n}}(x)\zeta(x')\varphi(x_{\rm n})\,dx\,.$$

By the definition of weak derivative,

$$\int_\Omega u(x)\zeta(x')\varphi'(x_{\rm n})\,dx = -\int_\Omega \frac{\partial u}{\partial x_{\rm n}}(x)\zeta(x')\varphi(x_{\rm n})\,dx\,.$$

Since the classical derivative $v_{x_{\rm n}}$ is the same as $\dfrac{\partial u}{\partial x_{\rm n}}$, the right-hand side of the two equalities above are the same; hence due to the fact that the test function $\zeta \in \mathscr{C}_c^\infty(\omega)$ is arbitrary,

$$\int_0^1 \big(u(x',x_{\rm n}) - v(x',x_{\rm n})\big)\varphi'(x_{\rm n})\,dx_{\rm n} = 0$$

for almost every $x' \in \omega$. As a consequence, by Problem 2.4, we find that

$$u(x',x_{\rm n}) - v(x',x_{\rm n}) = \text{a constant independent of } x_{\rm n}$$

which shows that u is absolutely continuous on almost all straight lines parallel the $x_{\rm n}$-axis. $\square$

2.8 The Gagliardo-Nirenberg-Sobolev Inequality

In the previous section, we considered the embedding for the case that $p > \mathrm{n}$.

Theorem 2.35 (Gagliardo-Nirenberg-Sobolev inequality)**.** *For* $1 \leqslant p < \mathrm{n}$, *set* $p^* = \dfrac{\mathrm{n}p}{\mathrm{n}-p}$. *Then*

$$\|u\|_{L^{p^*}(\mathbb{R}^{\mathrm{n}})} \leqslant C(p,\mathrm{n})\|Du\|_{L^p(\mathbb{R}^{\mathrm{n}})} \qquad \forall\, u \in W^{1,p}(\mathbb{R}^{\mathrm{n}})\,.$$

Proof for the case $\mathrm{n}=2$. Suppose first that $p = 1$ in which case $p^* = 2$, and we must prove that

$$\|u\|_{L^2(\mathbb{R}^2)} \leqslant C\|Du\|_{L^1(\mathbb{R}^2)} \qquad \forall\, u \in \mathscr{C}^1_c(\mathbb{R}^2)\,. \tag{2.12}$$

Since u has compact support, by the fundamental theorem of calculus,

$$u(x_1,x_2) = \int_{-\infty}^{x_1} \partial_1 u(y_1,x_2)dy_1 = \int_{-\infty}^{x_2} \partial_2 u(x_1,y_2)dy_2$$

so that

$$|u(x_1,x_2)| \leqslant \int_{-\infty}^{\infty} |\partial_1 u(y_1,x_2)|dy_1 \leqslant \int_{-\infty}^{\infty} |Du(y_1,x_2)|dy_1$$

and

$$|u(x_1,x_2)| \leqslant \int_{-\infty}^{\infty} |\partial_2 u(x_1,y_2)|dy_2 \leqslant \int_{-\infty}^{\infty} |Du(x_1,y_2)|dy_2\,.$$

Hence, it follows that

$$|u(x_1,x_2)|^2 \leqslant \int_{-\infty}^{\infty} |Du(y_1,x_2)|dy_1 \int_{-\infty}^{\infty} |Du(x_1,y_2)|dy_2$$

Integrating over $\mathbb{R}^2$, we find that

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |u(x_1,x_2)|^2\, dx_1\, dx_2 \\
&\qquad \leqslant \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Big(\int_{-\infty}^{\infty} |Du(y_1,x_2)|dy_1 \int_{-\infty}^{\infty} |Du(x_1,y_2)|dy_2\Big)\, dx_1 dx_2 \\
&\qquad \leqslant \Big(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |Du(x_1,x_2)|\, dx_1 dx_2\Big)^2
\end{aligned}$$

which is (2.12).

Next, if $1 \leqslant p < 2$, substitute $|u|^\gamma$ for u in (2.12) to find that

$$\Big(\int_{\mathbb{R}^2} |u|^{2\gamma}\,dx\Big)^{\frac{1}{2}} \leqslant C\gamma \int_{\mathbb{R}^2} |u|^{\gamma-1}|Du|\,dx$$
$$\leqslant C\gamma \|Du\|_{L^p(\mathbb{R}^2)} \Big(\int_{\mathbb{R}^2} |u|^{\frac{p(\gamma-1)}{p-1}}\,dx\Big)^{\frac{p-1}{p}}$$

Choose γ so that $2\gamma = \dfrac{p(\gamma-1)}{p-1}$; hence, $\gamma = \dfrac{p}{2-p}$, and

$$\Big(\int_{\mathbb{R}^2} |u|^{\frac{2p}{2-p}}\,dx\Big)^{\frac{2-p}{2p}} \leqslant C\gamma \|Du\|_{L^p(\mathbb{R}^2)}\,,$$

so that

$$\|u\|_{L^{\frac{2p}{2-p}}(\mathbb{R}^\mathrm{n})} \leqslant C_{p,\mathrm{n}} \|Du\|_{L^p(\mathbb{R}^\mathrm{n})} \tag{2.13}$$

for all $u \in \mathscr{C}_c^1(\mathbb{R}^2)$.

Since $\mathscr{C}_c^\infty(\mathbb{R}^2)$ is dense in $W^{1,p}(\mathbb{R}^2)$, there exists a sequence $\{u_j\}_{j=1}^\infty \subseteq \mathscr{C}_c^\infty(\mathbb{R}^2)$ such that

$$u_j \to u \qquad \text{in} \quad W^{1,p}(\mathbb{R}^2)\,.$$

Hence, by (2.13), for all $j, k \in \mathbb{N}$,

$$\|u_j - u_k\|_{L^{\frac{2p}{2-p}}(\mathbb{R}^\mathrm{n})} \leqslant C_{p,\mathrm{n}} \|Du_j - Du_k\|_{L^p(\mathbb{R}^\mathrm{n})}$$

so there exists $U \in L^{\frac{2p}{2-p}}(\mathbb{R}^\mathrm{n})$ such that

$$u_j \to U \qquad \text{in} \quad L^{\frac{2p}{2-p}}(\mathbb{R}^\mathrm{n})\,.$$

Hence $U = u$ a.e. in $\mathbb{R}^2$, and by continuity of the norms, (2.13) holds for all $u \in W^{1,p}(\mathbb{R}^2)$. □

Proof for the general case of dimension n. Following the proof for $\mathrm{n} = 2$, we see that

$$|u(x)|^{\frac{\mathrm{n}}{\mathrm{n}-1}} \leqslant \prod_{i=1}^{\mathrm{n}} \Big(\int_{-\infty}^{\infty} |Du(x_1, ..., y_i, ..., x_\mathrm{n})|dy_i\Big)^{\frac{1}{\mathrm{n}-1}}$$

so that

$$\int_{-\infty}^{\infty} |u(x)|^{\frac{\mathrm{n}}{\mathrm{n}-1}}\,dx_1 \leqslant \int_{-\infty}^{\infty} \prod_{i=1}^{\mathrm{n}} \Big(\int_{-\infty}^{\infty} |Du(x_1, ..., y_i, ..., x_\mathrm{n})|dy_i\Big)^{\frac{1}{\mathrm{n}-1}}\,dx_1$$
$$= \Big(\int_{-\infty}^{\infty} |Du|dy_1\Big)^{\frac{1}{\mathrm{n}-1}} \int_{-\infty}^{\infty} \prod_{i=2}^{\mathrm{n}} \Big(\int_{-\infty}^{\infty} |Du|dy_i\Big)^{\frac{1}{\mathrm{n}-1}}\,dx_1$$
$$\leqslant \Big(\int_{-\infty}^{\infty} |Du|dy_1\Big)^{\frac{1}{\mathrm{n}-1}} \prod_{i=2}^{\mathrm{n}} \Big(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |Du|\,dx_1 dy_i\Big)^{\frac{1}{\mathrm{n}-1}}\,,$$

where the last inequality follows from Hölder's inequality.

Integrating the last inequality with respect to x_2, we find that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |u(x)|^{\frac{\mathrm{n}}{\mathrm{n}-1}}\, dx_1\, dx_2 \leqslant \Big(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |Du|\, dx_1 dy_2\Big)^{\frac{1}{\mathrm{n}-1}} \int_{-\infty}^{\infty} \prod_{\substack{i=1\\ i\neq 2}}^{\mathrm{n}} I_i^{\frac{1}{\mathrm{n}-1}}\, dx_2\,,$$

where

$$I_1 = \int_{-\infty}^{\infty} |Du| dy_1\,, \quad I_i = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |Du| dx_1 dy_i \text{ for } i = 3, \cdots, \mathrm{n}\,.$$

Applying Hölder's inequality, we find that

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |u(x)|^{\frac{\mathrm{n}}{\mathrm{n}-1}}\, dx_1 dx_2 \\
&\qquad \leqslant \Big(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |Du|\, dx_1 dy_2\Big)^{\frac{1}{\mathrm{n}-1}} \Big(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |Du| dy_1\, dx_2\Big)^{\frac{1}{\mathrm{n}-1}} \times \\
&\qquad\qquad \times \prod_{i=3}^{\mathrm{n}} \Big(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |Du|\, dx_1\, dx_2 dy_i\Big)^{\frac{1}{\mathrm{n}-1}}\,.
\end{aligned}$$

Next, continue to integrate with respect to $x_3, ..., x_{\mathrm{n}}$ to find that

$$\begin{aligned}
\int_{\mathbb{R}^{\mathrm{n}}} |u|^{\frac{\mathrm{n}}{\mathrm{n}-1}}\, dx &\leqslant \prod_{i=1}^{\mathrm{n}} \Big(\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} |Du|\, dx_1 \cdots dy_i \cdots dx_{\mathrm{n}}\Big)^{\frac{1}{\mathrm{n}-1}} \\
&= \Big(\int_{\mathbb{R}^{\mathrm{n}}} |Du| dx\Big)^{\frac{\mathrm{n}}{\mathrm{n}-1}}\,.
\end{aligned}$$

This proves the case that $p = 1$. The case that $1 < p < \mathrm{n}$ follows identically as in the proof of $\mathrm{n} = 2$. □

It is common to employ the Sobolev embedding theorems for the case that $p = n$ and of particular interest is the case that $p = 2$ in dimension $n = 2$; as stated, neither Morrey's inequality or the Galiardo-Nirenberg inequality can be applied in this setting, but in fact, we have the following

Theorem 2.36. *Suppose that $u \in H^1(\mathbb{R}^2)$. Then for all $2 \leqslant q < \infty$,*

$$\|u\|_{L^q(\mathbb{R}^2)} \leqslant C\sqrt{q}\, \|u\|_{H^1(\mathbb{R}^2)}\,.$$

Proof. We first consider the case that $u \in \mathscr{C}_c^\infty(\mathbb{R}^2)$. Let x and y be points in $\mathbb{R}^2$, and write $r = |x - y|$. Let $\theta \in \mathbb{S}^1$. Introduce spherical coordinates (r, θ) with origin at x, and let g be the same cut-off function that was used in the proof of Theorem 2.19. We define $U(r,\theta) := g(r)u(x+re^{i\theta})$ or equivalently, $U(y-x) = g(|x-y|)u(y)$, were $y = x + re^{i\theta}$. Then

$$U(0,\theta) = -\int_0^1 \frac{\partial U}{\partial r}(r,\theta)dr\,;$$

thus

$$|U(0,\theta)| \leqslant \int_0^1 |DU(r,\theta)|dr\,.$$

Using the fact that $u(x) = \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} U(0,\theta)d\theta$, we have that

$$\begin{aligned}|u(x)| &\leqslant \frac{1}{2\pi}\int_0^{2\pi}\int_0^1 r^{-1}|DU(r,\theta)|rdrd\theta \\ &\leqslant \frac{1}{2\pi}\int_{\mathbb{R}^2} \mathbf{1}_{B(x,1)}(y)|x-y|^{-1}|DU(y)|dy := K * |DU|\,,\end{aligned}$$

where the integral kernel $K(x) = \dfrac{1}{2\pi}\mathbf{1}_{B(0,1)}|x|^{-1}$.

Using Young's inequality from Theorem 1.48, we obtain the estimate

$$\|K * f\|_{L^q(\mathbb{R}^2)} \leqslant \|K\|_{L^k(\mathbb{R}^2)}\|f\|_{L^2(\mathbb{R}^2)} \quad \text{for} \quad \frac{1}{k} = \frac{1}{q} - \frac{1}{2} + 1\,. \tag{2.14}$$

Using the inequality (2.14) with $f = |DU|$, we see that

$$\begin{aligned}\|u\|_{L^q(\mathbb{R}^2)} &\leqslant C\|DU\|_{L^2(\mathbb{R}^2)}\Big(\int_{B(0,1)} |y|^{-k}dy\Big)^{\frac{1}{k}} \\ &\leqslant C\|DU\|_{L^2(\mathbb{R}^2)}\Big[\int_0^1 r^{1-k}dr\Big]^{\frac{1}{k}} \leqslant C\|u\|_{H^1(\mathbb{R}^2)}\Big[\frac{q+2}{4}\Big]^{\frac{1}{k}}\,.\end{aligned}$$

When $q \to \infty$, $\dfrac{1}{k} \to \dfrac{1}{2}$ and $(q+2)^{\frac{1}{k}} \leqslant C\sqrt{q}$ for some $C > 0$ independent of k, so

$$\|u\|_{L^q(\mathbb{R}^2)} \leqslant C\sqrt{q}\|u\|_{H^1(\mathbb{R}^2)}\,.$$

Using the density of $\mathscr{C}_c^\infty(\mathbb{R}^2)$ in $H^1(\mathbb{R}^2)$ completes the proof. $\square$

In fact, the above theorem holds more generally for $u \in W^{1,\mathrm{n}}(\mathbb{R}^\mathrm{n})$. Then for all $\mathrm{n} \leqslant q < \infty$,

$$\|u\|_{L^q(\mathbb{R}^\mathrm{n})} \leqslant C\sqrt{q}\|u\|_{W^{1,\mathrm{n}}(\mathbb{R}^\mathrm{n})}\,.$$

Remark 2.37. For functions $u \in \mathscr{C}_c^\infty(\mathbb{R}^2)$ such that the support of u is contained in a set Ω with finite measure, the inequality $\|u\|_{L^q(\mathbb{R}^2)} \leqslant C\sqrt{q}\,\|u\|_{H^1(\mathbb{R}^2)}$ holds for all $1 \leqslant q < \infty$, but the constant depends on $|\Omega|$.

Evidently, it is not possible to obtain the estimate $\|u\|_{L^\infty(\mathbb{R}^n)} \leqslant C\|u\|_{W^{1,n}(\mathbb{R}^n)}$ with a constant $C < \infty$. The following provides an example of a function in this borderline situation.

Example 2.38. Let $\Omega \subseteq \mathbb{R}^n$ denote the open unit ball in $\mathbb{R}^2$. The unbounded function $u = \log\log\left(1 + \frac{1}{|x|}\right)$ belongs to $W^{1,n}(B(0,1))$. We show this for the case that $n = 2$.

First, note that

$$\int_\Omega |u(x)|^2\,dx = \int_0^{2\pi}\int_0^1 \left[\log\log\left(1+\frac{1}{r}\right)\right]^2 r dr d\theta\,.$$

The only potential singularity of the integrand occurs at $r = 0$, but according to L'Hospital's rule,

$$\lim_{r\to 0} r\left[\log\log\left(1+\frac{1}{r}\right)\right]^2 = 0, \tag{2.15}$$

so the integrand is continuous and hence $u \in L^2(\Omega)$.

In order to compute the partial derivatives of u, note that

$$\frac{\partial}{\partial x_j}|x| = \frac{x_j}{|x|}\,, \text{ and } \frac{d}{dz}|f(z)| = \frac{f(z)}{|f(z)|}\frac{df}{dz}\,,$$

where $f : \mathbb{R} \to \mathbb{R}$ is differentiable. It follows that for x away from the origin,

$$Du(x) = \frac{-x}{\log(1+1/|x|)(|x|+1)|x|^2}\,, \quad (x \neq 0)\,.$$

Let $\varphi \in \mathscr{C}_c^\infty(\Omega)$ and fix $\epsilon > 0$. Then

$$\int_{\Omega\setminus B_\epsilon(0)} u(x)\frac{\partial\varphi}{\partial x_i}(x)dx = -\int_{\Omega\setminus B(0,\epsilon)}\frac{\partial u}{\partial x_i}(x)\varphi(x)\,dx + \int_{\partial B(0,\epsilon)} u\varphi \mathrm{N}_i dS\,,$$

where $\mathrm{N} = (\mathrm{N}_1, ..., \mathrm{N}_n)$ denotes the inward-pointing unit normal on the curve $\partial B(0,\epsilon)$, so that $\mathrm{N}dS = \epsilon(\cos\theta, \sin\theta)d\theta$. It follows that

$$\begin{aligned}\int_{\Omega - B_\epsilon(0)} u(x)D\varphi(x)\,dx = &-\int_{\Omega - B_\epsilon(0)} Du(x)\varphi(x)\,dx \\ &- \int_0^{2\pi}\epsilon(\cos\theta,\sin\theta)\,\log\log\left(1+\frac{1}{\epsilon}\right)\varphi(\epsilon,\theta)d\theta\,.\end{aligned} \tag{2.16}$$

We claim that $Du \in L^2(\Omega)$ (and hence also in $L^1(\Omega)$), for

$$\begin{aligned}\int_\Omega |Du(x)|^2\,dx &= \int_0^{2\pi}\int_0^1 \frac{1}{r(r+1)^2\Big[\log\Big(1+\frac{1}{r}\Big)\Big]^2}dr d\theta \\ &\leqslant \pi\int_0^{1/2}\frac{1}{r(\log r)^2}dr + \pi\int_{1/2}^1 \frac{1}{r(r+1)^2\big[\log\big(1+1/r\big)\big]^2}dr\,,\end{aligned}$$

where we use the inequality $\log\big(1+\dfrac{1}{r}\big) \geqslant \log\dfrac{1}{r} = -\log r \geqslant 0$ for $0 \leqslant r \leqslant 1$. The second integral on the right-hand side is clearly bounded, while

$$\int_0^{1/2}\frac{1}{r(\log r)^2}dr = \int_{-\infty}^{-\log 2}\frac{1}{t^2e^t}e^t dt = \int_{-\infty}^{-\log 2}\frac{1}{t^2}dt < \infty\,,$$

so that $Du \in L^2(\Omega)$. Letting $\epsilon \to 0$ in (2.16) and using (2.15) for the boundary integral, by the Dominated Convergence Theorem, we conclude that

$$\int_\Omega u(x)D\varphi(x)\,dx = -\int_\Omega Du(x)\varphi(x)\,dx \qquad \forall\,\varphi \in \mathscr{C}_c^\infty(\Omega)\,.$$

We conclude this section by stating the following theorem which can be proved by induction.

Theorem 2.39 (Gagliardo-Nirenberg-Sobolev inequality for $W^{k,p}(\mathbb{R}^{\mathrm{n}})$)**.** *Suppose that* $1 \leqslant kp < \mathrm{n}$*, and* $D^k u \in L^p(\mathbb{R}^{\mathrm{n}})$*. Then* $u \in L^{\frac{\mathrm{n}p}{\mathrm{n}-kp}}(\mathbb{R}^{\mathrm{n}})$*, and*

$$\|u\|_{L^{\frac{\mathrm{n}p}{\mathrm{n}-kp}}(\mathbb{R}^{\mathrm{n}})} \leqslant C\|D^k u\|_{L^p(\mathbb{R}^{\mathrm{n}})} \text{ for a constant } C = C(k,p,\mathrm{n})\,. \tag{2.17}$$

Remark 2.40. For $0 < s < \dfrac{\mathrm{n}}{2}$, there exists a constant $C = C(\mathrm{n}, s)$ such that

$$\|u\|_{L^{\frac{2\mathrm{n}}{\mathrm{n}-2s}}(\mathbb{R}^{\mathrm{n}})} \leqslant C\|u\|_{H^s(\mathbb{R}^{\mathrm{n}})} \qquad \forall\, u \in H^s(\mathbb{R}^{\mathrm{n}})\,. \tag{2.18}$$

In other words, (2.17) holds for the case $p = 2$ and any real number $k \in \big(0, \dfrac{\mathrm{n}}{2}\big)$.

Moreover, with the help of the Morrey inequality (2.8), we can establish the following

Theorem 2.41 (Morrey's inequality for $W^{k,p}(\mathbb{R}^{\mathrm{n}})$)**.** *Suppose that* $\mathrm{n} < kp$*, and* $u \in W^{k,p}(\mathbb{R}^{\mathrm{n}})$*. Then* $u \in \mathscr{C}^{k-1-[\frac{\mathrm{n}}{p}],1+[\frac{\mathrm{n}}{p}]-\frac{\mathrm{n}}{p}}(\mathbb{R}^{\mathrm{n}})$*, and*

$$\|u\|_{\mathscr{C}^{k-1-[\frac{\mathrm{n}}{p}],1+[\frac{\mathrm{n}}{p}]-\frac{\mathrm{n}}{p}}(\mathbb{R}^{\mathrm{n}})} \leqslant C\|u\|_{W^{k,p}(\mathbb{R}^{\mathrm{n}})} \text{ for a constant } C = C(k,p,\mathrm{n})\,. \tag{2.19}$$

In the rest of this section, a more general version of Sobolev inequality is introduced.

Theorem 2.42 (Interpolation inequality for $W^{k,p}(\mathbb{R}^{\mathrm{n}})$)**.** *Let* n *be a given positive integer. For $j \leqslant k \lneqq \ell$, $0 < \theta \leqslant 1$ and $1 \leqslant q, r \leqslant p \leqslant \infty$ satisfying*

$$\frac{1}{p} - \frac{k}{\mathrm{n}} = \theta\Big(\frac{1}{q} - \frac{\ell}{\mathrm{n}}\Big) + (1-\theta)\Big(\frac{1}{r} - \frac{j}{\mathrm{n}}\Big)\,, \tag{2.20a}$$
$$\frac{1}{r} - \frac{j}{\mathrm{n}} > \frac{1}{p} - \frac{k}{\mathrm{n}} \geqslant \frac{1}{q} - \frac{\ell}{\mathrm{n}}\,, \tag{2.20b}$$

there exists a generic constant $C = C(p, q, r, j, k, \ell)$ such that

$$\|D^k u\|_{L^p(\mathbb{R}^{\mathrm{n}})} \leqslant C \|D^\ell u\|^{\theta}_{L^q(\mathbb{R}^{\mathrm{n}})} \|D^j u\|^{1-\theta}_{L^r(\mathbb{R}^{\mathrm{n}})} \qquad \forall\, u \in W^{\ell,q}(\mathbb{R}^{\mathrm{n}}) \cap W^{j,r}(\mathbb{R}^{\mathrm{n}})\,. \tag{2.21}$$

Proof. We prove the case of $\ell = k+1$, and the general case can be obtained using the established case and Theorem 2.39. Moreover, we also note that when $\dfrac{1}{p} - \dfrac{k}{n} = \dfrac{1}{q} - \dfrac{k+1}{n}$, we must have $\theta = 1$ and in this case (2.21) is a special case of Theorem 2.39. Without loss of generality, we assume that $\dfrac{1}{p} - \dfrac{k}{n} > \dfrac{1}{q} - \dfrac{k+1}{n}$.

Let Φ be the fundamental solution of $-\Delta$, $\chi \in \mathscr{C}^\infty_c(\mathbb{R}^{\mathrm{n}})$ such that $\chi = 1$ in a small ball centered at the origin, and define $F = \chi\Phi$. Then (C.10) implies that

$$\Phi * \Delta u = \mathrm{div}(\Phi * Du) = (D\Phi) * (Du) \qquad \forall\, u \in \mathscr{C}^\infty_c(\mathbb{R}^{\mathrm{n}})\,;$$

thus by the fact that $\Delta\Phi = 0$ on $\mathbb{R}^{\mathrm{n}}\backslash\{0\}$, using (C.6) we further obtain that if $u \in \mathscr{C}^\infty_c(\mathbb{R}^{\mathrm{n}})$,

$$\begin{aligned} u(x) &= -(\Phi * \Delta u)(x) = -((D\Phi) * (Du))(x) \\ &= -(DF * Du)(x) - (D((1-\chi)\Phi) * Du)(x) \\ &= -(DF * Du)(x) + \Delta((1-\chi)\Phi) * u = -(DF * Du)(x) + (\psi * u)(x) \end{aligned}$$

for some $\psi \in \mathscr{C}^\infty_c(\mathbb{R}^{\mathrm{n}})$ given by $\psi = \Delta((1-\chi)\Phi) = -\Phi\Delta\chi + 2\mathrm{div}(\Phi\nabla\chi)$. In other words, for some $\psi \in \mathscr{C}^\infty_c(\mathbb{R}^{\mathrm{n}})$,

$$u = -(DF) * (Du) + \psi * u \qquad \forall\, u \in \mathscr{C}^\infty_c(\mathbb{R}^{\mathrm{n}})\,. \tag{2.22}$$

Therefore, by the fact that $DF \in L^1(\mathbb{R}^{\mathrm{n}})$, (2.22) shows that for $j \leqslant k$,

$$D^k u = -(DF) * (D^{k+1} u) + D^{k-j}\psi * D^j u \qquad \forall\, u \in \mathscr{C}^\infty_c(\mathbb{R}^{\mathrm{n}})$$

and Young's inequality further provides that for all $u \in \mathscr{C}_c^\infty(\mathbb{R}^n)$,

$$\|D^k u\|_{L^p(\mathbb{R}^n)} \leqslant \|DF\|_{L^s(\mathbb{R}^n)} \|D^{k+1}u\|_{L^q(\mathbb{R}^n)} + \|D^{k-j}\psi\|_{L^t(\mathbb{R}^n)} \|D^j u\|_{L^r(\mathbb{R}^n)} ,$$

where $1 + \dfrac{1}{p} = \dfrac{1}{q} + \dfrac{1}{s} = \dfrac{1}{r} + \dfrac{1}{t}$ which is equivalent to that $1 \leqslant q, r \leqslant p \leqslant \infty$. We note that $|(DF)(x)| = \mathcal{O}(|x|^{1-n})$ as $|x| \to 0$ and DF has compact support, $\|DF\|_{L^s(\mathbb{R}^n)} < \infty$ if and only if $1 \leqslant s < \dfrac{n}{n-1}$. Therefore, if $1 \leqslant s < \dfrac{n}{n-1}$,

$$\|D^k u\|_{L^p(\mathbb{R}^n)} \leqslant C_s \|D^{k+1}u\|_{L^q(\mathbb{R}^n)} + C\|D^j u\|_{L^r(\mathbb{R}^n)} \qquad \forall\, u \in \mathscr{C}_c^\infty(\mathbb{R}^n)\,.$$

Now we initiate the scaling argument. For each $u \in \mathscr{C}_c^\infty(\mathbb{R}^n)$ and $\lambda > 0$, let $v(x) = u(\lambda^{-1}x)$. Then $v \in \mathscr{C}_c^\infty(\mathbb{R}^n)$ so that

$$\|D^k v\|_{L^p(\mathbb{R}^n)} \leqslant C_s \|D^{k+1}v\|_{L^q(\mathbb{R}^n)} + C\|D^j v\|_{L^r(\mathbb{R}^n)}\,.$$

The change of variable $y = \lambda^{-1}x$ then implies that for all $\lambda > 0$,

$$\|D^k u\|_{L^p(\mathbb{R}^n)} \leqslant C_s \lambda^{-1+\frac{n}{q}-\frac{n}{p}} \|D^{k+1}u\|_{L^q(\mathbb{R}^n)} + C\lambda^{k-j+\frac{n}{r}-\frac{n}{p}} \|D^j u\|_{L^r(\mathbb{R}^n)}\,, \tag{2.23}$$

Since the minimization of the right-hand side of (2.23) over $\lambda > 0$ cannot be trivial,

$$\Big(-1 + \frac{n}{q} - \frac{n}{p}\Big)\Big(k - j + \frac{n}{r} - \frac{n}{p}\Big) < 0\,.$$

By the fact that $1 + \dfrac{1}{p} = \dfrac{1}{q} + \dfrac{1}{s}$ and $1 \leqslant s < \dfrac{n}{n-1}$, we conclude that $-1 + \dfrac{n}{q} - \dfrac{n}{p} < 0$; thus

$$k - j + \frac{n}{r} - \frac{n}{p} > 0 \quad \text{or equivalently,} \quad \frac{1}{r} - \frac{j}{n} > \frac{1}{p} - \frac{k}{n}\,.$$

For $A, B, \alpha, \beta > 0$,

$$\min_{\lambda > 0}(A\lambda^\alpha + B\lambda^{-\beta}) = C(\alpha,\beta) A^{\frac{\beta}{\alpha+\beta}} B^{\frac{\alpha}{\alpha+\beta}}\,; \tag{2.24}$$

thus with the assignments $A = C\|D^j u\|_{L^r(\mathbb{R}^n)}$, $B = C_s\|D^{k+1}u\|_{L^q(\mathbb{R}^n)}$, $\alpha = k - j + \dfrac{n}{r} - \dfrac{n}{p}$ and $\beta = \dfrac{n}{p} - \dfrac{n}{q} + 1$ in (2.24), we obtain that

$$\|D^k u\|_{L^p(\mathbb{R}^n)} \leqslant C\|D^{k+1}u\|^\theta_{L^q(\mathbb{R}^n)} \|D^j u\|^{1-\theta}_{L^r(\mathbb{R}^n)}$$

for $\theta = \dfrac{\alpha}{\alpha+\beta} = \dfrac{k - j + \frac{n}{r} - \frac{n}{p}}{k - j + 1 - \frac{n}{q} + \frac{n}{r}}$.

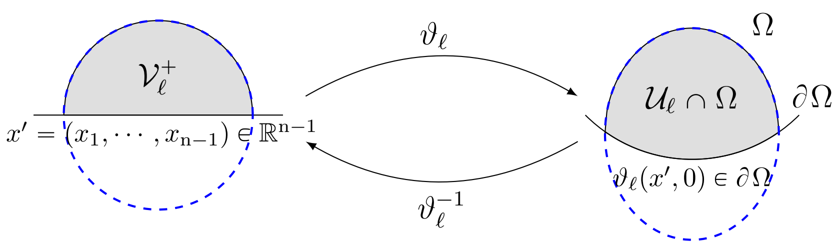
Letting $\dfrac{1}{r} - \dfrac{j}{n} = \sigma$, $\dfrac{1}{q} - \dfrac{k+1}{n} = \tau$ and $\dfrac{1}{p} - \dfrac{k}{n} = \kappa$, then $\sigma > \kappa > \tau$ (which validate (2.20b)) $\theta = \dfrac{\sigma - \kappa}{\sigma - \tau}$; thus $0 < \theta < 1$ and

$$\theta\Big(\frac{1}{q} - \frac{k+1}{n}\Big) + (1-\theta)\Big(\frac{1}{r} - \frac{j}{n}\Big) = \frac{\sigma-\kappa}{\sigma-\tau}\tau + \frac{\kappa-\tau}{\sigma-\tau}\sigma = \frac{\kappa(\sigma-\tau)}{\sigma-\tau} = \kappa = \frac{1}{p} - \frac{k}{n}\,. \qquad \square$$

2.9 Local Coordinates near $\partial\Omega$

Let $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ denote an open, bounded subset with $\mathscr{C}^1$-boundary, and let $\{\mathcal{U}_\ell\}_{\ell=1}^K$ denote an open covering of $\partial\Omega$, such that for each $\ell \in \{1, 2, ..., K\}$, with $\mathcal{V}_\ell = B(0, r_\ell)$ denoting the open ball of radius r_ℓ centered at the origin, $\mathcal{V}_\ell^+ = \mathcal{V}_\ell \cap \{x_{\mathrm{n}} > 0\}$ and $\mathcal{V}_\ell^- = \mathcal{V}_\ell \cap \{x_{\mathrm{n}} < 0\}$ denoting the upper and lower half of $\mathcal{V}_\ell$, respectively, there exist $\mathscr{C}^1$-class *charts* ϑ_ℓ which satisfy

$$
\begin{aligned}
\vartheta_\ell : \mathcal{V}_\ell \to \mathcal{U}_\ell \ \text{ is a } \mathscr{C}^1 \text{ diffeomorphism}\,, \\
\vartheta_\ell(\mathcal{V}_\ell^+) = \mathcal{U}_\ell \cap \Omega\,, \\
\vartheta_\ell(\mathcal{V}_\ell \cap \{x_{\mathrm{n}} = 0\}) = \mathcal{U}_\ell \cap \partial\Omega\,.
\end{aligned}
\tag{2.25}
$$



2.10 Sobolev Extension and Trace Theorems

Let $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ denote an open, bounded domain with $\mathscr{C}^1$-boundary.

Theorem 2.43. *Suppose that $\widetilde{\Omega} \subseteq \mathbb{R}^{\mathrm{n}}$ is a bounded and open domain such that $\Omega \subset\subset \widetilde{\Omega}$. Then for $1 \leqslant p \leqslant \infty$, there exists a bounded linear operator*

$$E : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^{\mathrm{n}})$$

such that for all $u \in W^{1,p}(\Omega)$,

1. *$Eu = u$ a.e. in Ω;*

2. *$\operatorname{spt}(Eu) \subseteq \widetilde{\Omega}$;*

3. *$\|Eu\|_{W^{1,p}(\mathbb{R}^{\mathrm{n}})} \leqslant C\|u\|_{W^{1,p}(\Omega)}$ for a constant $C = C(p, \Omega, \widetilde{\Omega})$.*

Theorem 2.44. *For $1 \leqslant p < \infty$, there exists a bounded linear operator*

$$\tau : W^{1,p}(\Omega) \to L^p(\partial\Omega)$$

such that for all $u \in W^{1,p}(\Omega)$

1. $\tau u = u|_{\partial\Omega}$ *for all* $u \in W^{1,p}(\Omega) \cap \mathscr{C}^0(\bar{\Omega})$;

2. $\|\tau u\|_{L^p(\partial\Omega)} \leqslant C\|u\|_{W^{1,p}(\Omega)}$ *for a constant* $C = C(p, \Omega)$.

Proof. Suppose that $u \in \mathscr{C}^1(\bar{\Omega})$, $z \in \partial\Omega$, and that $\partial\Omega$ is locally flat near z. In particular, for $r > 0$ sufficiently small, $B(z,r) \cap \partial\Omega \subseteq \{x_{\rm n} = 0\}$. Let $0 \leqslant \xi \in \mathscr{C}_c^\infty(B(z,r))$ such that $\xi = 1$ on $B(z, r/2)$. Set $\Gamma = \partial\Omega \cap B(z, r/2)$, $B^+(z,r) = B(z,r) \cap \Omega$, and let $dx_h = dx_1 \cdots dx_{\rm n-1}$. Then

$$\begin{aligned}
\int_\Gamma |u|^p \, dx_h &\leqslant \int_{\{x_{\rm n}=0\}} \xi |u|^p \, dx_h = -\int_{B^+(z,r)} \frac{\partial}{\partial x_{\rm n}} (\xi |u|^p) \, dx \\
&\leqslant -\int_{B^+(z,r)} \frac{\partial \xi}{\partial x_{\rm n}} |u|^p \, dx - p \int_{B^+(z,2\delta)} \xi |u|^{p-2} u \frac{\partial u}{\partial x_{\rm n}} dx \\
&\leqslant C \int_{B^+(z,r)} |u|^p \, dx + C \| |u|^{p-1} \|_{L^{\frac{p}{p-1}}(B^+(z,r))} \Big\| \frac{\partial u}{\partial x_{\rm n}} \Big\|_{L^p(B^+(z,r))} \\
&\leqslant C \int_{B^+(z,r)} (|u|^p + |Du|^p) \, dx \, . \qquad (2.26)
\end{aligned}$$

On the other hand, if the boundary is not locally flat near $z \in \partial\Omega$, then we use a $\mathscr{C}^1$-diffeomorphism to locally *straighten the boundary.* More specifically, suppose that $z \in \partial\Omega \cap U_\ell$ for some $\ell \in \{1, ..., K\}$ and consider the $\mathscr{C}^1$-chart ϑ_ℓ defined in (2.25). Define the function $U = u \circ \vartheta_\ell$; then $U : \mathcal{V}_\ell^+ \to \mathbb{R}$. Setting $\Gamma = \mathcal{V}_\ell \cap \{x_{\rm n} = 0\}$, we see from the inequality (2.26) that

$$\int_\Gamma |U|^p \, dx_h \leqslant C_\ell \int_{\mathcal{V}_\ell^+} (|U|^p + |DU|^p) \, dx \, .$$

Using the fact that $D\vartheta_\ell$ is bounded and continuous on $\mathcal{V}_\ell^+$, the change of variables formula shows that

$$\int_{\mathcal{U}_\ell \cup \partial\Omega} |u|^p dS \leqslant C_\ell \int_{\mathcal{U}_\ell^+} (|u|^p + |Du|^p) \, dx \, .$$

Summing over all $\ell \in \{1, ..., K\}$ shows that

$$\int_{\partial\Omega} |u|^p dS \leqslant C \int_\Omega (|u|^p + |Du|^p) dx \, . \qquad (2.27)$$

The inequality (2.27) holds for all $u \in \mathscr{C}^1(\bar{\Omega})$. According to Theorem 2.23, for $u \in W^{1,p}(\Omega)$ there exists a sequence $\{u_j\}_{j=1}^\infty \subseteq \mathscr{C}^\infty(\bar{\Omega})$ such that $u_j \to u$ in $W^{1,p}(\Omega)$. By inequality (2.27),

$$\|\tau u_k - \tau u_j\|_{L^p(\partial\Omega)} \leqslant C \|u_k - u_j\|_{W^{1,p}(\Omega)} \, ,$$

so that τu_j is Cauchy in $L^p(\partial\Omega)$, and hence a limit exists in $L^p(\partial\Omega)$. We define the trace operator τ as this limit:

$$\lim_{j\to 0} \|\tau u - \tau u_j\|_{L^p(\partial\Omega)} = 0\,.$$

Since the sequence u_j converges uniformly to u if $u \in \mathscr{C}^0(\bar\Omega)$, we see that $\tau u = u|_{\partial\Omega}$ for all $u \in W^{1,p}(\Omega) \cup \mathscr{C}^0(\bar\Omega)$. □

Sketch of the proof of Theorem 2.43. Just as in the proof of the trace theorem, first suppose that $u \in \mathscr{C}^1(\bar\Omega)$ and that near $z \in \partial\Omega$, $\partial\Omega$ is *locally flat*, so that for some $r > 0$, $\partial\Omega \cup B(z,r) \subseteq \{x_{\rm n} = 0\}$. Letting $B^+ = B(z,r)\cup\{x_{\rm n} \geqslant 0\}$ and $B^- = B(z,r)\cup\{x_{\rm n} \leqslant 0\}$, we define the extension of u by

$$\bar u(x) = \begin{cases} u(x) & \text{if } x \in B^+\,, \\ -3u(x_1,...,x_{\rm n-1},-x_{\rm n}) + 4u(x_1,...,x_{\rm n-1},-x_{\rm n}/2) & \text{if } x \in B^-\,. \end{cases}$$

Define $u^+ = \bar u|_{B^+}$ and $u^- = \bar u|_{B^-}$.

It is clear that $u^+ = u^-$ on $\{x_{\rm n} = 0\}$, and by the chain rule, it follows that

$$\frac{\partial u^-}{\partial x_{\rm n}}(x) = 3\frac{\partial u^-}{\partial x_{\rm n}}(x_1,...,-x_{\rm n}) - 2\frac{\partial u^-}{\partial x_{\rm n}}(x_1,...,-\frac{x_{\rm n}}{2})\,,$$

so that $\dfrac{\partial u^+}{\partial x_{\rm n}} = \dfrac{\partial u^-}{\partial x_{\rm n}}$ on $\{x_{\rm n} = 0\}$. This shows that $\bar u \in \mathscr{C}^1(B(z,r))$. using the charts ϑ_ℓ to locally straighten the boundary, and the density of the $\mathscr{C}^\infty(\bar\Omega)$ in $W^{1,p}(\Omega)$, the theorem is proved. □

Later, we will provide a proof for higher-order Sobolev extensions of H^k-type functions.

2.11 Integration by parts for functions in $H^1(\Omega)$

We can now state the following theorem which is a generalization of (2.3) and the divergence theorem.

Theorem 2.45. *Suppose that $\Omega \subseteq \mathbb{R}^{\rm n}$ is a bounded domain with $\mathscr{C}^1$-boundary. Then for each $i \in \{1,\cdots,{\rm n}\}$,*

$$\int_\Omega \frac{\partial u}{\partial x_i} v\,dx = \int_{\partial\Omega} uv{\rm N}_i\,dS - \int_\Omega u\frac{\partial v}{\partial x_i}\,dx \qquad \forall\, u,v \in H^1(\Omega)\,.$$

Proof. By Theorem 2.22, there exists $\{u_k\}_{k=1}^\infty, \{v_k\}_{k=1}^\infty \subseteq \mathscr{C}^\infty(\Omega) \cap H^1(\Omega)$ such that $u_k \to u$ and $v_k \to v$ in $H^1(\Omega)$. Moreover, Theorem 2.44 implies that $u_k \to u$ and $v_k \to v$ in $L^2(\partial\Omega)$. Therefore, the divergence theorem implies that

$$\begin{aligned}\int_\Omega \frac{\partial u}{\partial x_i} v\, dx &= \lim_{k\to\infty} \int_\Omega \frac{\partial u_k}{\partial x_i} v_k\, dx = \lim_{k\to\infty} \Big[\int_{\partial\Omega} u_k v_k \mathrm{N}_i\, dS - \int_\Omega \frac{\partial u_k}{\partial x_i} v_k\, dx \Big] \\ &= \int_{\partial\Omega} uv\mathrm{N}_i\, dS - \int_\Omega \frac{\partial u}{\partial x_i} v\, dx\,. \end{aligned}$$ □

2.12 The subspace $W_0^{1,p}(\Omega)$

Definition 2.46. We let $W_0^{1,p}(\Omega)$ denote the closure of $\mathscr{C}_c^\infty(\Omega)$ in $W^{1,p}(\Omega)$.

Theorem 2.47. *Suppose that $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ is bounded with $\mathscr{C}^1$-boundary, and that $u \in W^{1,p}(\Omega)$. Then*

$$u \in W_0^{1,p}(\Omega) \quad \textit{if and only if} \quad \tau u = 0 \quad \textit{on} \quad \partial\Omega\,.$$

Proof. We first assume that $u \in W_0^{1,p}(\Omega)$ and prove that $\tau u = 0$ on $\partial\Omega$. Since $u \in W_0^{1,p}(\Omega)$, there exists $\{u_k\}_{k=1}^\infty \subseteq \mathscr{C}_c^\infty(\Omega)$ such that $u_k \to u$ in $W^{1,p}(\Omega)$. Since $\tau : W^{1,p}(\Omega) \to L^p(\partial\Omega)$ is bounded,

$$\|\tau u\|_{L^p(\partial\Omega)} = \lim_{k\to\infty} \|\tau u_k\|_{L^p(\partial\Omega)} = 0\,.$$

Next, we establish that $u \in W_0^{1,p}(\Omega)$ provided that $\tau u = 0$ on $\partial\Omega$. Let $\{\mathcal{U}_\ell\}_{\ell=1}^K$ denote an open covering of $\partial\Omega$ such that for each $\ell \in \{1, 2, ..., K\}$, there exist $\mathscr{C}^1$-class charts ϑ_ℓ given by (5.4) which satisfy that

$$\vartheta_\ell : B(0, r_\ell) \subseteq \mathbb{R}^{\mathrm{n}} \to \mathcal{U}_\ell \cap \Omega \ \text{ is a } \mathscr{C}^1\text{-diffeomorphism}\,.$$

Let $\mathcal{U}_0 \subset\subset \Omega$ be such that $\{\mathcal{U}_\ell\}_{\ell=0}^K$ forms an open cover of Ω, and let $\{\xi_\ell\}_{\ell=0}^K$ denote a partition of unity subordinate to this open cover; that is, for each $\ell \in \{0, 1, \cdots, K\}$, $0 \leqslant \xi_\ell \leqslant 1$ and $\mathrm{spt}(\xi_\ell) \subseteq \mathcal{U}_\ell$, as well as $\sum\limits_{\ell=0}^K \xi_\ell = 1$. We then construct a new partition of unity $\{\zeta_\ell\}_{\ell=0}^K$ subordinate to $\{\mathcal{U}_m\}_{m=0}^K$ by

$$\zeta_\ell = \frac{\xi_\ell^2}{\sum_{m=0}^K \xi_m^2}$$

so that $\sqrt{\zeta_\ell} \in \mathscr{C}_c^\infty(\mathbb{R}^{\mathrm{n}})$ for all $\ell \in \{0, 1, \cdots, K\}$. For a given $u \in W_0^{1,p}(\Omega)$, we define $u^{(\ell)} = \sqrt{\zeta_\ell}(u \circ \vartheta_\ell)$. Then $u^{(\ell)} \in W^{1,p}(\mathbb{R}^{\mathrm{n}}_+)$ and $\tau u^{(\ell)} = 0$ for all $\ell \in \{1, \cdots, K\}$. By definition of the trace, for each ℓ there exists a sequence $\{u_k^{(\ell)}\}_{k=1}^\infty \subseteq \mathscr{C}^\infty(\overline{\mathbb{R}^{\mathrm{n}}_+})$ such that $u_k^{(\ell)} \to u^{(\ell)}$ in $W^{1,p}(\mathbb{R}^{\mathrm{n}}_+)$ and $\tau u_k^{(\ell)} = u_k^{(\ell)}|_{\mathbb{R}^{\mathrm{n}-1}} \to 0$ in $L^p(\mathbb{R}^{\mathrm{n}-1})$ as $k \to \infty$. Note that for $x_h \in \mathbb{R}^{\mathrm{n}-1}$ and $x_{\mathrm{n}} \geqslant 0$,

$$u_k^{(\ell)}(x_h, x_{\mathrm{n}}) = u_k^{(\ell)}(x_h, 0) + \int_0^{x_{\mathrm{n}}} u_{k,\mathrm{n}}^{(\ell)}(x_h, t)\, dt\,;$$

thus Hölder's inequality implies that

$$\begin{aligned}
&\int_{\mathbb{R}^{\mathrm{n}-1}} \big|u_k^{(\ell)}(x_h, x_{\mathrm{n}})\big|^p\, dx_h \\
&\qquad \leqslant C\Big[\int_{\mathbb{R}^{\mathrm{n}-1}} \big|u_k^{(\ell)}(x_h, 0)\big|^p\, dx_h + x_{\mathrm{n}}^{p-1} \int_0^{x_{\mathrm{n}}} \int_{\mathbb{R}^{\mathrm{n}-1}} \big|Du_k^{(\ell)}(x_h, t)\big|^p\, dx_h dt\Big].
\end{aligned}$$

Passing to the limit as $k \to \infty$, by the fact that $u_k^{(\ell)} \to u^{(\ell)}$ in $W^{1,p}(\mathbb{R}^{\mathrm{n}}_+)$ and $\tau u_k^{(\ell)} \to 0$ in $L^p(\mathbb{R}^{\mathrm{n}-1})$ we find that

$$\int_{\mathbb{R}^{\mathrm{n}-1}} \big|u^{(\ell)}(x_h, x_{\mathrm{n}})\big|^p\, dx_h \leqslant C x_{\mathrm{n}}^{p-1} \int_0^{x_{\mathrm{n}}} \int_{\mathbb{R}^{\mathrm{n}-1}} \big|Du^{(\ell)}(x_h, t)\big|^p\, dx_h dt\,. \tag{2.28}$$

Let $\chi \in \mathscr{C}^\infty(\mathbb{R}_+)$ satisfy

$$\chi = 1 \ \text{ on } \ [0, 1]\,, \qquad \chi = 0 \ \text{ on } \ [2, \infty)\,, \qquad \text{and} \qquad 0 \leqslant \chi \leqslant 1\,.$$

Define $\chi_k(x) = \chi(k x_{\mathrm{n}})$ for $x \in \mathbb{R}^{\mathrm{n}}_+$ as well as $v_k^{(\ell)} = (1 - \chi_k) u^{(\ell)}$. Then using (2.28),

$$\begin{aligned}
&\int_{\mathbb{R}^{\mathrm{n}}_+} \big|Dv_k^{(\ell)}(x) - Du^{(\ell)}(x)\big|^p\, dx \\
&\qquad \leqslant C\Big[\int_{\mathbb{R}^{\mathrm{n}}_+} \big|\chi_k(x)\big|^p \big|Du^{(\ell)}(x)\big|^p\, dx + \int_{\mathbb{R}^{\mathrm{n}}_+} \big|D\chi_k(x)\big|^p \big|u^{(\ell)}(x)\big|^p\, dx\Big] \\
&\qquad \leqslant C\Big[\int_{\mathbb{R}^{\mathrm{n}}_+} \big|\chi_k(x)\big|^p \big|Du^{(\ell)}(x)\big|^p\, dx + k^p \int_0^{\frac{2}{k}} \int_{\mathbb{R}^{\mathrm{n}-1}} \big|u^{(\ell)}(x_h, t)\big|^p\, dx_h dt\Big] \\
&\qquad \leqslant C\Big[\int_{\mathbb{R}^{\mathrm{n}}_+} \big|\chi_k(x)\big|^p \big|Du^{(\ell)}(x)\big|^p\, dx + \int_0^{\frac{2}{k}} \int_{\mathbb{R}^{\mathrm{n}-1}} \big|u^{(\ell)}(x_h, t)\big|^p\, dx_h dt\Big]
\end{aligned}$$

which converges to 0 as $k \to \infty$. In other words, $\big\{Dv_k^{(\ell)}\big\}_{k=1}^\infty$ converges to $Du^{(\ell)}$ in $L^p(\mathbb{R}^{\mathrm{n}}_+)$. It is also clear that $\big\{v_k^{(\ell)}\big\}_{k=1}^\infty$ converges to $u^{(\ell)}$ in $L^p(\mathbb{R}^{\mathrm{n}}_+)$ since

$$\big\|v_k^{(\ell)} - u^{(\ell)}\big\|_{L^p(\mathbb{R}^{\mathrm{n}}_+)} = \big\|\chi_k u^{(\ell)}\big\|_{L^p(\mathbb{R}^{\mathrm{n}}_+)} \leqslant \big\|u^{(\ell)}\big\|_{L^p(\mathbb{R}^{\mathrm{n}-1} \times [0, \frac{2}{k}))}\,.$$

Define $u_k = \zeta_0 u + \sum\limits_{\ell=1}^{K} \sqrt{\zeta_\ell}(v_k^{(\ell)} \circ \vartheta_\ell^{-1})$. Then $u_k \in \mathscr{C}_c^\infty(\Omega)$ for $k \gg 1$. Moreover,

$$\begin{aligned}
\|u_k - u\|_{W^{1,p}(\Omega)} &= \Big\|u_k - \sum_{\ell=0}^{K} \zeta_\ell u\Big\|_{W^{1,p}(\Omega)} \\
&\leqslant \sum_{\ell=1}^{K} \Big\|\sqrt{\zeta_\ell}(v_k^{(\ell)} \circ \vartheta_\ell^{-1}) - \sqrt{\zeta_\ell}(u^{(\ell)} \circ \vartheta_\ell^{-1})\Big\|_{W^{1,p}(\Omega)} \\
&\leqslant C \sum_{\ell=1}^{K} \|v_k^{(\ell)} - u^{(\ell)}\|_{W^{1,p}(\Omega)}
\end{aligned}$$

which implies that $\{u_k\}_{k=1}^\infty$ converges to u in $W^{1,p}(\Omega)$. As a consequence, $u \in W_0^{1,p}(\Omega)$. ▫

We can now state some embedding theorems for bounded domains Ω.

Theorem 2.48 (Gagliardo-Nirenberg inequality for $W^{1,p}(\Omega)$)**.** *Suppose that $\Omega \subseteq \mathbb{R}^{\rm n}$ is open and bounded with $\mathscr{C}^1$-boundary, and $1 \leqslant p < {\rm n}$. Then there exists a generic constant $C = C(p, {\rm n}, \Omega)$ such that*

$$\|u\|_{L^{\frac{{\rm n}p}{{\rm n}-p}}(\Omega)} \leqslant C\|u\|_{W^{1,p}(\Omega)} \qquad \forall\, u \in W^{1,p}(\Omega)\,.$$

Proof. Choose $\widetilde{\Omega} \subseteq \mathbb{R}^{\rm n}$ bounded such that $\Omega \subset\subset \widetilde{\Omega}$, and let Eu denote the Sobolev extension of u to $\mathbb{R}^{\rm n}$ such that $Eu = u$ a.e., $\operatorname{spt}(Eu) \subseteq \widetilde{\Omega}$, and $\|Eu\|_{W^{1,p}(\mathbb{R}^{\rm n})} \leqslant C\|u\|_{W^{1,p}(\Omega)}$. Then by the Gagliardo-Nirenberg inequality,

$$\|u\|_{L^{\frac{{\rm n}p}{{\rm n}-p}}(\Omega)} \leqslant \|Eu\|_{L^{\frac{{\rm n}p}{{\rm n}-p}}(\mathbb{R}^{\rm n})} \leqslant C\|D(Eu)\|_{L^p(\mathbb{R}^{\rm n})} \leqslant C\|Eu\|_{W^{1,p}(\mathbb{R}^{\rm n})} \leqslant C\|u\|_{W^{1,p}(\Omega)}\,. \quad ▫$$

By following the proof of Theorem 2.35, we have the following generalization for integers $k \geqslant 1$:

Theorem 2.49 (Gagliardo-Nirenberg-Sobolev inequality for $W^{k,p}(\Omega)$)**.** *Suppose that $\Omega \subseteq \mathbb{R}^{\rm n}$ is open and bounded with $\mathscr{C}^1$-boundary, and $1 \leqslant kp < {\rm n}$. Then there exists a generic constant $C = C(k, p, {\rm n}, \Omega)$ such that*

$$\|u\|_{L^{\frac{{\rm n}p}{{\rm n}-kp}}(\Omega)} \leqslant C\|u\|_{W^{k,p}(\Omega)} \qquad \forall\, u \in W^{k,p}(\Omega)\,. \tag{2.29}$$

In fact, as mentioned in Remark 2.40, the theorem is true for real numbers $s > 0$ replacing integers $k \geqslant 1$, and follows from linear interpolation and the theory of fractional-order Sobolev spaces defined later in Section 5.2. In the important case that $p = 2$, we are then able to answer the question of which H^s spaces embed in L^q spaces. For example, when $\mathrm{n} = 2$ and $s = \frac{1}{2}$, we see that $\|u\|_{L^4(\Omega)} \leqslant C\|u\|_{H^{\frac{1}{2}}(\Omega)}$, and when $\mathrm{n} = 3$ and $s = \frac{1}{2}$, $\|u\|_{L^{\frac{12}{5}}(\Omega)} \leqslant C\|u\|_{H^{\frac{1}{2}}(\Omega)}$.

Theorem 2.50 (Gagliardo-Nirenberg inequality for $W_0^{1,p}(\Omega)$)**.** *Suppose that $\Omega \subseteq \mathbb{R}^\mathrm{n}$ is open and bounded with $\mathscr{C}^1$-boundary, and $1 \leqslant p < \mathrm{n}$. Then there exists a generic constant $C = C(p, \mathrm{n}, \Omega)$ such that for all $1 \leqslant q \leqslant \frac{\mathrm{n}p}{\mathrm{n}-p}$,*

$$\|u\|_{L^q(\Omega)} \leqslant C\|Du\|_{L^p(\Omega)} \qquad \forall\, u \in W_0^{1,p}(\Omega)\,. \tag{2.30}$$

Proof. By definition there exists a sequence $\{u_j\}_{j=1}^\infty \subseteq \mathscr{C}_c^\infty(\Omega)$ such that $u_j \to u$ in $W^{1,p}(\Omega)$. Extend each u_j by 0 on $\Omega^\complement$. Applying Theorem 2.35 to this extension, and using the continuity of the norms, we obtain

$$\|u\|_{L^{\frac{\mathrm{n}p}{n-p}}(\Omega)} \leqslant C\|Du\|_{L^p(\Omega)}\,.$$

Since Ω is bounded, the assertion follows by Hölder's inequality. □

Theorem 2.51. *Suppose that $\Omega \subseteq \mathbb{R}^2$ is open and bounded with $\mathscr{C}^1$-boundary. Then there exists a generic constant $C = C(\Omega)$ such that for all $1 \leqslant q < \infty$,*

$$\|u\|_{L^q(\Omega)} \leqslant C\sqrt{q}\|u\|_{H^1(\Omega)} \qquad \forall\, u \in H_0^1(\Omega)\,. \tag{2.31}$$

Proof. The proof follows that of Theorem 2.36. Instead of introducing the cut-off function g, we employ a partition of unity subordinate to the finite covering of the bounded domain Ω, in which case it suffices to assume that $\mathrm{spt}(u) \subseteq \mathrm{spt}(U)$ with U also defined in the proof Theorem 2.36. □

Remark 2.52. Inequality (2.30) is commonly referred to as the *Poincaré inequality*; it is invaluable in the study of the *Dirichlet problem* for Poisson's equation, since the right-hand side provides an $H^1(\Omega)$-equivalent norm for all $u \in H_0^1(\Omega)$. We will show that (2.30) holds for all dimensions n. In particular, there exists constants C_1, C_2 such that

$$C_1\|Du\|_{L^2(\Omega)} \leqslant \|u\|_{H^1(\Omega)} \leqslant C_2\|Du\|_{L^2(\Omega)}\,.$$

It follows that (2.31) can be written as

$$\|u\|_{L^q(\Omega)} \leqslant C\sqrt{q}\|Du\|_{L^2(\Omega)} \qquad \forall\, u \in H_0^1(\Omega)\,. \tag{2.32}$$

A more general form of the Poincaré inequality is given as follows:

Lemma 2.53 (Poincaré inequality). *Let $\Omega \subseteq \mathbb{R}^n$ denote an open, bounded, connected, and smooth domain. Then there exists a generic constant $C = C(\Omega)$ such that*

$$\|u - \bar{u}\|_{L^2(\Omega)} \leqslant C\|Du\|_{L^2(\Omega)} \quad \forall\, u \in H^1(\Omega)\,, \tag{2.33}$$

where $\bar{u} := ⨍_\Omega u(y)dy$ denotes the average value of u over Ω.

Proof. Suppose for the sake of contradiction that (2.33) does not hold. Then there is a sequence $\{u_j\}_{j=1}^\infty \subseteq H^1(\Omega)$ satisfying

$$\|u_j - \bar{u}_j\|_{L^2(\Omega)} > j\|Du_j\|_{L^2(\Omega)}\,, \tag{2.34}$$

with an associated sequence on the unit ball of $H^1(\Omega)$ given by

$$w_j = \frac{u_j - \bar{u}_j}{\|u_j - \bar{u}_j\|_{L^2(\Omega)}} \text{ with } \|w_j\|_{L^2(\Omega)} = 1 \text{ and } \bar{w}_j = 0\,.$$

According to (2.34), $\|Dw_j\|_{L^2(\Omega)} < j^{-1}$, so that $\|w_j\|^2_{H^1(\Omega)} < 1 + j^{-2} < \infty$. Strong compactness, given by Theorem 2.67 (see also Theorem 6.7) provides a subsequence $\{w_{j_k}\}_{k=1}^\infty$ and a limit $w \in L^2(\Omega)$ such that $w_{j_k} \to w$ in $L^2(\Omega)$ as $k \to \infty$. The limit w satisfies $\bar{w} = 0$ and $\|w\|_{L^2(\Omega)} = 1$.

Letting $\varphi \in \mathscr{C}_c^\infty(\Omega)$. We see that

$$\begin{aligned}\int_\Omega w(x)D\varphi(x)\,dx &= \lim_{k\to\infty}\int_\Omega w_{j_k}(x)D\varphi(x)\,dx \\ &= -\lim_{k\to\infty}\int_\Omega Dw_{j_k}(x)\,\varphi(x)\,dx \leqslant \lim_{k\to\infty} j_k^{-1}\|\varphi\|_{L^2(\Omega)} = 0\,.\end{aligned}$$

This shows that the weak derivative of w exists and is equal to zero almost everywhere; that is, $w \in H^1(\Omega)$ and $Dw = 0$ a.e. As Ω is connected, we see that w is a constant, and since $\bar{w} = 0$, we see that $w = 0$, contradicting the fact that $\|w\|_{L^2(\Omega)} = 1$. $\square$

Corollary 2.54. *Whenever $\bar{u} = ⨍_\Omega u(y)dy = 0$, $\|Du\|_{L^2(\Omega)}$ is an equivalent norm on $H^1(\Omega)$. In particular, there exists constants C_1, C_2 such that*

$$C_1\|Du\|_{L^2(\Omega)} \leqslant \|u\|_{H^1(\Omega)} \leqslant C_2\|Du\|_{L^2(\Omega)} \qquad \forall\, u \in H^1(\Omega)/\mathbb{R}\,.$$

The identical proof also shows that the validity of the following two results:

Lemma 2.55 (Poincaré inequality for $H_0^1(\Omega)$)**.** *Let $\Omega \subseteq \mathbb{R}^n$ denote an open, bounded, connected, and smooth domain. Then*

$$\|u\|_{L^2(\Omega)} \leqslant C\|Du\|_{L^2(\Omega)} \qquad \forall\, u \in H_0^1(\Omega)\,, \tag{2.35}$$

where the constant C depends on Ω.

Lemma 2.56 (Another Poincaré inequality)**.** *Let $\Omega \subseteq \mathbb{R}^n$ denote an open, bounded, connected, and smooth domain. Then for $k \in L^\infty(\partial\Omega)$ and $k \geqslant 0$ on $\partial\Omega$ and $k > 0$ on a set of surface measure greater than zero. Then*

$$\|u\|_{L^2(\Omega)} \leqslant C\big(\|\sqrt{k}\,u\|_{L^2(\partial\Omega)} + \|Du\|_{L^2(\Omega)}\big) \qquad \forall\, u \in H^1(\Omega)\,, \tag{2.36}$$

where the constant C depends on Ω.

Integration by parts for functions in $W_0^{1,p}(\Omega)$

Having established Theorem 2.47, using the density argument we can conclude the following

Theorem 2.57. *Let $\Omega \subseteq \mathbb{R}^n$ be a bounded domain with $\mathscr{C}^1$-boundary. Then for $1 < p < \infty$,*

$$\int_\Omega u\varphi,_j\, dx = -\int_\Omega u,_j\, \varphi\, dx \qquad \forall\, u \in W_0^{1,p}(\Omega) \text{ and } \varphi \in W^{1,p'}(\Omega)\,,$$

where $p' = \dfrac{p}{p-1}$ is the conjugate of p.

The proof is simple and is left as an exercise.

2.13 Weak Solutions to the Dirichlet Problem

Suppose that $\Omega \subseteq \mathbb{R}^n$ is an open, bounded domain with $\mathscr{C}^1$-boundary. A classical problem in the linear theory of partial differential equations consists of finding solutions to the *Dirichlet problem*:

$$-\Delta u = f \qquad \text{in} \quad \Omega\,, \tag{2.37a}$$

$$u = 0 \qquad \text{on} \quad \partial\Omega\,, \tag{2.37b}$$

where $\Delta = \sum_{i=1}^{\mathrm{n}} \frac{\partial^2}{\partial x_i^2}$ denotes the Laplace operator or *Laplacian*. As written, (2.37) is the so-called *strong form* of the Dirichlet problem, as it requires that u to possess certain weak second-order partial derivatives. A major turning-point in the modern theory of linear partial differential equations was the realization that *weak solutions* of (2.37) could be defined, which only require weak first-order derivatives of u to exist. (We will see more of this idea later when we discuss the theory of distributions.)

DEFINITION 2.58. The dual space of $H_0^1(\Omega)$ is denoted by $H^{-1}(\Omega)$. For $f \in H^{-1}(\Omega)$,

$$\|f\|_{H^{-1}(\Omega)} = \sup_{\|\psi\|_{H_0^1(\Omega)}=1} \langle f, \psi \rangle ,$$

where $\langle f, \psi \rangle$ denotes the duality pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$.

THEOREM 2.59 (The distributional space $H^{-1}(\Omega)$)**.** *For any $f \in H^{-1}(\Omega)$, there exist $n+1$ functions $f_j \in L^2(\Omega)$, $j = 0, 1, 2, ..., \mathrm{n}$ such that for all $v \in H_0^1(\Omega)$,*

$$\langle f, v \rangle = \int_\Omega \Big[f_0(x) v(x) + \sum_{i=1}^n f_i(x) \frac{\partial v}{\partial x_i}(x) \Big]\, dx \,, \tag{2.38}$$

and

$$\|f\|_{H^{-1}(\Omega)} = \inf \Big\{ \Big(\int_\Omega \sum_{j=0}^n |f_j(x)|^2 \, dx \Big)^{\frac{1}{2}} \,\Big|\, f \text{ satisfying } (2.38) \Big\} . \tag{2.39}$$

Proof. By the Riesz Representation Theorem, for every $f \in H^{-1}(\Omega)$ there exists $u \in H_0^1(\Omega)$ satisfying

$$(u, v)_{L^2(\Omega)} + (Du, Dv)_{L^2(\Omega)} = \langle f, v \rangle \qquad \forall\, v \in H_0^1(\Omega) \,. \tag{2.40}$$

Letting $f_0 = u$ and $f_i = \partial u / \partial x_i$ for $i = 1, ..., n$ gives the relation (2.38).

Then for $f \in H^{-1}(\Omega)$, we may write

$$\langle f, v \rangle = \int_\Omega \Big[g_0(x) v(x) + \sum_{i=1}^n g_i(x) \frac{\partial v}{\partial x_i}(x) \Big]\, dx \,, \tag{2.41}$$

for all $v \in H_0^1(\Omega)$ and $g_j \in L^2(\Omega)$ for $j = 0, 1, 2, ..., n$. Setting $u = v$ in (2.40) yields

$$\|u\|_{H^1(\Omega)}^2 \leqslant \int_\Omega \sum_{j=0}^n |g_0(x)|^2 \, dx \,.$$

Hence, since $f_0 = u$ and $f_i = \partial u/\partial x_i$, we see that

$$\int_\Omega \sum_{j=0}^n |f_j(x)|^2\, dx \leqslant \int_\Omega \sum_{j=0}^n |g_j(x)|^2\, dx\,. \tag{2.42}$$

From (2.38), we infer that

$$\|f\|_{H^{-1}(\Omega)} \leqslant \Big(\int_\Omega \sum_{j=0}^n |f_j(x)|^2\, dx\Big)^{\frac{1}{2}} \text{ if } \|v\|_{H_0^1(\Omega)} \leqslant 1\,.$$

Thus, with $v = u\|u\|^{-1}_{H_0^1(\Omega)}$ in (2.40), we have that

$$\|f\|^2_{H^{-1}(\Omega)} = \int_\Omega \sum_{j=0}^n |f_j(x)|^2\, dx\,. \tag{2.43}$$

Then, (2.39) follows from (2.41)-(2.43). □

Definition 2.60. A function $u \in H_0^1(\Omega)$ is a weak solution of (2.37) if

$$\int_\Omega Du \cdot Dv\, dx = \langle f, v\rangle \qquad \forall\, v \in H_0^1(\Omega)\,.$$

Remark 2.61. Note that f can be taken in $H^{-1}(\Omega)$. According to the Sobolev embedding theorem, this implies that when n = 1, the forcing function f can be taken to be the Dirac Delta distribution.

Remark 2.62. The motivation for Definition 2.60 is as follows. Since $\mathscr{C}_c^\infty(\Omega)$ is dense in $H_0^1(\Omega)$, multiply equation (2.37a) by $\varphi \in \mathscr{C}_c^\infty(\Omega)$, integrate over Ω, and employ the integration-by-parts formula to obtain $\int_\Omega Du \cdot D\varphi\, dx = \int_\Omega f\varphi\, dx$; the boundary terms vanish because φ is compactly supported.

Theorem 2.63 (Existence and uniqueness of weak solutions)**.** *For any $f \in H^{-1}(\Omega)$, there exists a unique weak solution to* (2.37)*.*

Proof. Using the Poincaré inequality, $\|Du\|_{L^2(\Omega)}$ is an H^1-equivalent norm for all $u \in H_0^1(\Omega)$, and $(Du, Dv)_{L^2(\Omega)}$ defines the inner-product on $H_0^1(\Omega)$. As such, according to the definition of weak solutions to (2.37), we are seeking $u \in H_0^1(\Omega)$ such that

$$(u, v)_{H_0^1(\Omega)} = \langle f, v\rangle \ \ \forall\, v \in H_0^1(\Omega)\,. \tag{2.44}$$

The existence of a unique $u \in H_0^1(\Omega)$ satisfying (2.44) is provided by the Riesz representation theorem for Hilbert spaces. □

REMARK 2.64. Note that the Riesz representation theorem shows that there exists a distribution, denoted by $-\Delta u \in H^{-1}(\Omega)$ such that

$$\langle -\Delta u, v\rangle = \langle f, v\rangle \qquad \forall\, v \in H_0^1(\Omega)\,.$$

The operator $-\Delta : H_0^1(\Omega) \to H^{-1}(\Omega)$ is thus an isomorphism.

A fundamental question in the theory of linear partial differential equations is commonly referred to as *elliptic regularity*, and can be explained as follows: in order to develop an existence and uniqueness theorem for the Dirichlet problem, we have significantly generalized the notion of solution to the class of weak solutions, which permitted very weak forcing functions in $H^{-1}(\Omega)$. Now suppose that the forcing function is smooth; is the weak solution smooth as well? Furthermore, does the weak solution agree with the classical solution? The answer is yes, and we will develop this regularity theory in Section 7, where it will be shown that for integers $k \geqslant 2$, $-\Delta : H^k(\Omega) \cap H_0^1(\Omega) \to H^{k-2}(\Omega)$ is also an isomorphism. An important consequence of this result is that $(-\Delta)^{-1} : H^{k-2}(\Omega) \to H^k(\Omega) \cap H_0^1(\Omega)$ is a *compact* linear operator, and as such has a countable set of eigenvalues, a fact that is eminently useful in the construction of solutions for heat- and wave-type equations.

For this reason, as well as the consideration of weak limits of nonlinear combinations of sequences, we must develop a compactness theorem, which generalizes the well-known Arzela-Ascoli theorem to Sobolev spaces.

2.14 Strong Compactness

In Section 1.5.3, we defined the notion of weak converence and weak compactness for L^p-spaces. Recall that for $1 \leqslant p < \infty$, a sequence $\{u_j\}_{j=1}^\infty \subseteq L^p(\Omega)$ converges *weakly* to $u \in L^p(\Omega)$, denoted $u_j \rightharpoonup u$ in $L^p(\Omega)$, if $\int_\Omega u_j v\, dx \to \int_\Omega uv\, dx$ for all $v \in L^q(\Omega)$, with $q = \frac{p}{p-1}$. We can extend this definition to Sobolev spaces.

DEFINITION 2.65. For $1 \leqslant p < \infty$, $u_j \rightharpoonup u$ in $W^{1,p}(\Omega)$ provided that $u_j \rightharpoonup u$ in $L^p(\Omega)$ and $Du_j \rightharpoonup Du$ in $L^p(\Omega)$.

Alaoglu's Lemma (Theorem 1.57) then implies the following theorem.

Theorem 2.66 (Weak compactness in $W^{1,p}(\Omega)$)**.** *Let $\Omega \subseteq \mathbb{R}^{\rm n}$ and $1 < p < \infty$. Suppose that*

$$\sup_j \|u_j\|_{W^{1,p}(\Omega)} \leqslant M < \infty$$

for some constant M independent of j. Then there exists a subsequence $u_{j_k} \rightharpoonup u$ in $W^{1,p}(\Omega)$.

It turns out that weak compactness often does not suffice for limit processes involving nonlinearities, and that the Gagliardo-Nirenberg inequality can be used to obtain the following strong compactness theorem.

Theorem 2.67 (Rellich's theorem on a bounded domain Ω)**.** *Let $\Omega \subseteq \mathbb{R}^{\rm n}$ be an open, bounded domain with $\mathscr{C}^1$-boundary, and $1 \leqslant p < {\rm n}$. Then $W^{1,p}(\Omega)$ is compactly embedded in $L^q(\Omega)$ for all $1 \leqslant q < \frac{{\rm n}p}{{\rm n}-p}$; that is, if*

$$\sup_j \|u_j\|_{W^{1,p}(\Omega)} \leqslant M < \infty$$

for some constant M independent of j, then there exists a subsequence $u_{j_k} \to u$ in $L^q(\Omega)$. In the case that ${\rm n} = 2$ and $p = 2$, $H^1(\Omega)$ is compactly embedded in $L^q(\Omega)$ for $1 \leqslant q < \infty$.

In order to prove Rellich's theorem, we recall the following classical compactness theorem.

Theorem 2.68 (Arzelà-Ascoli Theorem)**.** *Let $\Omega \subseteq \mathbb{R}^{\rm n}$ be a bounded domain. Suppose that $\{u_j\}_{j=1}^{\infty} \subseteq \mathscr{C}^0(\bar{\Omega})$ is a sequence of equi-continuous functions and $\sup_j \|u_j\|_{\mathscr{C}^0(\bar{\Omega})} \leqslant M < \infty$. Then there exists a subsequence $\left\{u_{j_k}\right\}_{k=1}^{\infty}$ which converges uniformly on $\bar{\Omega}$.*

Proof of Rellich's theorem. The proof proceeds in four steps. First, we use Sobolev extension to extend our sequence of functions onto $\mathbb{R}^{\rm n}$. Second, we use mollification to produce a smooth sequence of functions which will satisfy the hypothesis of the Arzela-Ascoli theorem. Third, we show that our mollified sequence is very close in L^1 to our original extended sequence, and hence close in L^q for $1 \leqslant q < \frac{{\rm n}p}{{\rm n}-p}$. Finally, a classical diagonal argument provides convergence of a subsequence in L^q.

Step 1. Sobolev Extension. Let $\widetilde{\Omega} \subseteq \mathbb{R}^{\mathrm{n}}$ denote an open, bounded domain such that $\Omega \subset\subset \widetilde{\Omega}$. By the Sobolev extension theorem, the sequence $\{Eu_j\}_{j=1}^{\infty}$ satisfies $\mathrm{spt}(Eu_j) \subseteq \widetilde{\Omega}$, and

$$\sup_j \|Eu_j\|_{W^{1,p}(\mathbb{R}^{\mathrm{n}})} \leqslant CM \,.$$

Denote the sequence Eu_j by $\bar{u}_j$. By the Gagliardo-Nirenberg inequality, if $p^* = \dfrac{\mathrm{n}p}{\mathrm{n}-p}$,

$$\sup_j \|\bar{u}_j\|_{L^{p^*}(\mathbb{R}^{\mathrm{n}})} \leqslant C \sup_j \|\bar{u}_j\|_{W^{1,p}(\mathbb{R}^{\mathrm{n}})} \leqslant CM \,.$$

Step 2. Approximation by smooth functions. For $\epsilon > 0$, let η_ϵ denote the standard mollifiers and set $\bar{u}_j^\epsilon = \eta_\epsilon * Eu_j$. By choosing $\epsilon > 0$ sufficiently small, $\bar{u}_j^\epsilon \in \mathscr{C}_c^\infty(\widetilde{\Omega})$.

We compute that

$$\bar{u}_j^\epsilon = \int_{B(0,\epsilon)} \frac{1}{\epsilon^{\mathrm{n}}} \eta\Big(\frac{y}{\epsilon}\Big) \bar{u}_j(x-y)dy = \int_{B(0,1)} \eta(z)\bar{u}_j(x-\epsilon z)dz \,. \tag{2.45}$$

Applying the fundamental theorem of calculus to $\bar{u}_j$, we see that

$$\bar{u}_j(x-\epsilon z) - \bar{u}_j(x) = \int_0^1 \frac{d}{dt}\bar{u}_j(x-\epsilon tz)dt = -\epsilon \int_0^1 D\bar{u}_j(x-\epsilon tz)\cdot z\,dt \,. \tag{2.46}$$

Substitution of (2.46) into (2.45) shows that

$$\big|\bar{u}_j^\epsilon(x) - \bar{u}_j(x)\big| = \epsilon \int_{B(0,1)} \eta(z) \int_0^1 |D\bar{u}_j(x-\epsilon tz)|\,dzdt \,,$$

so that

$$\int_{\widetilde{\Omega}} |\bar{u}_j^\epsilon(x) - \bar{u}_j(x)|\,dx = \epsilon \int_{B(0,1)} \eta(z) \int_0^1 \int_{\widetilde{\Omega}} |D\bar{u}_j(x-\epsilon tz)|\,dxdzdt \,.$$

By the mean value theorem for integrals, there exists a $\vartheta \in (0,1)$ such that

$$\int_{\widetilde{\Omega}} |D\bar{u}_j(x-\epsilon\vartheta z)|dz = \int_0^1 \int_{\widetilde{\Omega}} |D\bar{u}_j(x-\epsilon tz)|\,dz\,dt \,.$$

Hence,

$$\begin{aligned}
\int_{\widetilde{\Omega}} |\bar{u}_j^\epsilon(x) - \bar{u}_j(x)|\,dx &= \epsilon \int_{\widetilde{\Omega}} \int_{B(0,1)} \eta(z)|D\bar{u}_j(x-\epsilon\vartheta z)|dzdx \\
&= \epsilon \int_{\widetilde{\Omega}} \int_{B(0,\epsilon\vartheta)} \eta_{\epsilon\vartheta}(w)|D\bar{u}_j(x-w)|dwdx \\
&= \epsilon \int_{\widetilde{\Omega}} \left(\eta_{\epsilon\vartheta} * |D\bar{u}_j|\right)(x)\,dx \,.
\end{aligned}$$

By Young's inequality for convolution,

$$\int_{\tilde{\Omega}} |\bar{u}_j^\epsilon(x) - \bar{u}_j(x)|\, dx \leqslant \epsilon \|D\bar{u}_j\|_{L^1(\tilde{\Omega})} \leqslant \epsilon \|D\bar{u}_j\|_{L^p(\tilde{\Omega})} < \epsilon CM\,.$$

Using the L^p-interpolation Lemma 1.21, for any $1 < q < \mathrm{n}p/(\mathrm{n}-p)$,

$$\begin{aligned} \|\bar{u}_j^\epsilon - \bar{u}_j\|_{L^q(\tilde{\Omega})} &\leqslant \|\bar{u}_j^\epsilon - \bar{u}_j\|^a_{L^1(\tilde{\Omega})} \|\bar{u}_j^\epsilon - \bar{u}_j\|^{1-a}_{L^{\frac{\mathrm{n}p}{\mathrm{n}-p}}(\tilde{\Omega})} \\ &\leqslant \epsilon^a C M^a \, \|D\bar{u}_j^\epsilon - D\bar{u}_j\|^{1-a}_{L^p(\tilde{\Omega})} \\ &\leqslant \epsilon^a CM\,. \end{aligned} \tag{2.47}$$

The inequality (2.47) shows that $\bar{u}_j^\epsilon$ is arbitrarily close to $\bar{u}_j$ in $L^q(\Omega)$ uniformly in $j \in \mathbb{N}$; as such, we attempt to use the smooth sequence $\bar{u}_j^\epsilon$ to construct a convergent subsequence $\bar{u}_{j_k}^\epsilon$.

Step 3. Extracting a convergent subsequence. Our goal is to employ the Arzela-Ascoli Theorem, so we show that for $\epsilon > 0$ fixed,

$$\sup_j \|\bar{u}_j^\epsilon\|_{\mathscr{C}^0(\bar{\tilde{\Omega}})} \leqslant \widetilde{M} < \infty \quad \text{and} \quad \{\bar{u}_j^\epsilon\}_{j=1}^\infty \text{ is equi-continous.}$$

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

$$\begin{aligned} \sup_j \|\bar{u}_j^\epsilon\|_{\mathscr{C}^0(\bar{\tilde{\Omega}})} &\leqslant \sup_j \sup_{x \in \bar{\tilde{\Omega}}} \int_{B(x,\epsilon)} \eta_\epsilon(x-y)|\bar{u}_j(y)|dy \\ &\leqslant \|\eta_\epsilon\|_{L^\infty(\mathbb{R}^\mathrm{n})} \sup_j \|\bar{u}_j\|_{L^1(\tilde{\Omega})} \leqslant C\epsilon^{-\mathrm{n}} < \infty\,, \end{aligned}$$

and similarly

$$\sup_j \|\bar{D}u_j^\epsilon\|_{\mathscr{C}^0(\bar{\tilde{\Omega}})} \leqslant \|D\eta_\epsilon\|_{L^\infty(\mathbb{R}^\mathrm{n})} \sup_j \|\bar{u}_j\|_{L^1(\tilde{\Omega})} \leqslant C\epsilon^{-\mathrm{n}-1} < \infty\,.$$

The latter inequality proves equicontinuity of the sequence $\{\bar{u}_j^\epsilon\}_{j=1}^\infty$, and hence there exists a subsequence $\{\bar{u}_{j_k}\}_{k=1}^\infty$ which converges uniformly on $\tilde{\Omega}$, so that

$$\limsup_{k,\ell \to \infty} \|\bar{u}_{j_k}^\epsilon - \bar{u}_{j_\ell}^\epsilon\|_{L^q(\tilde{\Omega})} = 0\,.$$

Step 4. Diagonal argument. Now, fix $\delta > 0$ and choose ϵ sufficiently small in (2.47) such that (with the triangle inequality)

$$\limsup_{k,\ell \to \infty} \|\bar{u}_{j_k} - \bar{u}_{j_\ell}\|_{L^q(\tilde{\Omega})} \leqslant \delta\,.$$

Letting $\delta = \dfrac{1}{2}, \dfrac{1}{3}$, etc., and using the diagonal argument to extract further subsequences, we can arrange to find a subsequence (again denoted by $\{\bar{u}_{j_k}\}_{k=1}^\infty$) of $\{\bar{u}_j\}_{j=1}^\infty$ such that

$$\limsup_{k,\ell\to\infty} \|\bar{u}_{j_k} - \bar{u}_{j_\ell}\|_{L^q(\tilde{\Omega})} = 0\,,$$

and hence

$$\limsup_{k,\ell\to\infty} \|u_{j_k} - u_{j_\ell}\|_{L^q(\Omega)} = 0\,,$$

The case that $\mathrm{n} = p = 2$ follows from Theorem 2.36. □

2.15 Weak convergence in $W^{1,p}(\Omega)$ for $1 < p < \infty$

If $W^{1,p}(\Omega)$ is reflexive, then the Alaoglu theorem (Theorem 1.57) can be used to study the weak-* convergence of a bounded sequence in $W^{1,p}(\Omega)$, which in turn is equivalent to the weak convergence of a bounded sequence in $W^{1,p}(\Omega)$. Our goal in this section is to establish the reflexivity of $W^{1,p}(\Omega)$ for $1 < p < \infty$.

Theorem 2.69 (Dual space of $W^{1,p}(\Omega)$)**.** *Let $1 < p < \infty$ with conjugate $p' = \dfrac{p}{p-1}$. For every $f \in W^{1,p}(\Omega)'$, there exists a unique vector-valued function $\boldsymbol{v} = (v_0, v_1, \cdots, v_\mathrm{n}) \in L^{p'}(\Omega)^{\mathrm{n}+1}$ such that*

$$\langle f, u\rangle = \int_\Omega u v_0\, dx + \sum_{\ell=1}^{\mathrm{n}} \int_\Omega \frac{\partial u}{\partial x_\ell} v_\ell\, dx\,,$$

where $\langle \cdot,\cdot\rangle$ denotes the duality pairing between distributions in $W^{1,p}(\Omega)'$ and functions in $W^{1,p}(\Omega)$. Moreover,

$$\|f\|_{W^{1,p}(\Omega)'} = \sum_{\ell=0}^{\mathrm{n}} \|v_\ell\|_{L^p(\Omega)}\,.$$

Proof. Define a bounded linear map $\mathrm{P} : W^{1,p}(\Omega) \to L^p(\Omega)^{\mathrm{n}+1}$ by

$$\mathrm{P}u = \Big(u, \frac{\partial u}{\partial x_1}, \cdots, \frac{\partial u}{\partial x_\mathrm{n}}\Big)\,,$$

and let W be the range of P. Since P is an isometry, W is a closed subspace of $L^p(\Omega)^{\mathrm{n}+1}$. For given $f \in W^{1,p}(\Omega)'$, define $L : W \to \mathbb{R}$ by

$$L(\mathrm{P}u) = \langle f, u\rangle\,.$$

Then $L \in W'$ since $\|L\|_{W'} \leqslant \|f\|_{W^{1,p}(\Omega)'}$. By the Hahn-Banach theorem, there exists an extension $\widetilde{L} : L^p(\Omega)^{\mathrm{n}+1} \to \mathbb{R}$ satisfying

1. $\widetilde{L}(\boldsymbol{w}) = L\boldsymbol{w}$ for all $\boldsymbol{w} \in W$; 2. $\|\widetilde{L}\|_{L^p(\Omega)^{\mathrm{n}+1\prime}} = \|L\|_{W'}$.

By the Risez representation theorem (Theorem 1.51), there exists a unique $\boldsymbol{v} = (v_0, v_1, \cdots, v_{\mathrm{n}}) \in L^{p'}(\Omega)^{\mathrm{n}+1}$ such that

$$\widetilde{L}(\boldsymbol{w}) = \sum_{\ell=0}^{\mathrm{n}} \int_\Omega w_\ell v_\ell \, dx \qquad \forall\, \boldsymbol{w} = (w_0, w_1, \cdots, w_{\mathrm{n}}) \in L^p(\Omega)^{\mathrm{n}+1}\,,$$

and $\|\widetilde{L}\|_{L^p(\Omega)^{\mathrm{n}+1\prime}} = \sum\limits_{\ell=0}^{\mathrm{n}} \|v_\ell\|_{L^p(\Omega)}$. In particular, we have

$$\langle f, u\rangle = L(\mathrm{P}u) = \widetilde{L}(\mathrm{P}u) = \int_\Omega u v_0 \, dx + \sum_{\ell=1}^{\mathrm{n}} \int_\Omega \frac{\partial u}{\partial x_\ell} v_\ell \, dx$$

which concludes the theorem. □

Remark 2.70. For the case $p = 2$, the existence of such a $\boldsymbol{v}$ in Theorem 2.69 is guaranteed by the Riesz representation theorem.

Let $\{u_k\}_{k=1}^\infty$ be a bounded sequence in $W^{1,p}(\Omega)$. Then $\{u_k\}_{k=1}^\infty$ and $\{\nabla u_k\}_{k=1}^\infty$ are both bounded sequences in $L^p(\Omega)$. Therefore, Theorem 1.59 implies that there exists a subsequence $\{u_{k_j}\}_{j=1}^\infty$ such that $u_{k_j} \rightharpoonup u$ in $L^p(\Omega)$ and $\nabla u_{k_j} \rightharpoonup v$ in $L^p(\Omega)$ for some functions $u, v \in L^p(\Omega)$. In other words,

$$\lim_{j\to\infty} \int_\Omega u_{k_j} \varphi \, dx = \int_\Omega u\varphi \, dx \quad \text{and} \quad \lim_{j\to\infty} \int_\Omega \nabla u_{k_j} \varphi \, dx = \int_\Omega v\varphi \, dx \quad \forall\, \varphi \in L^{p'}(\Omega)\,,$$

Let $\varphi \in \mathscr{C}_c^\infty(\Omega)$. Then $\varphi, \nabla\varphi \in L^{p'}(\Omega)$; thus by the definition of weak derivative,

$$\int_\Omega v\varphi \, dx = \lim_{j\to\infty} \int_\Omega \nabla u_{k_j} \varphi \, dx = -\lim_{j\to\infty} \int_\Omega u_{k_j} \nabla\varphi \, dx = -\int_\Omega u \nabla\varphi \, dx$$

which implies that $v = Du$ in the sense of distribution, or equivalently, v is the weak derivative of u. Therefore, we establish the following

Theorem 2.71. *Let $\{u_k\}_{k=1}^\infty$ be a bounded sequence in $W^{1,p}(\Omega)$ for $1 < p < \infty$. Then there exists a subsequence $\{u_{k_j}\}_{j=1}^\infty$ such that $\{u_{k_j}\}_{j=1}^\infty$ and $\{Du_{k_j}\}_{j=1}^\infty$ converges weakly to u and Du in $L^p(\Omega)$, respectively.*

To see that the convergence behavior in Theorem 2.71 is in fact the weak convergence in $W^{1,p}(\Omega)$, we make use of Theorem 2.69. Let $\{u_k\}_{k=1}^\infty$ be a bounded sequence in

$W^{1,p}(\Omega)$ and $f \in W^{1,p}(\Omega)'$. Then Theorem 2.69 provides a unique $\boldsymbol{v} = (v_0, v_1, \cdots, v_{\rm n}) \in L^{p'}(\Omega)^{{\rm n}+1}$ such that

$$\langle f, u\rangle = \int_\Omega u v_0\, dx + \sum_{\ell=1}^{\rm n} \int_\Omega \frac{\partial u}{\partial x_\ell} v_\ell\, dx\,.$$

Therefore, the subsequence $\{u_{k_j}\}_{j=1}^\infty$ provides by Theorem 2.71 satisfies

$$\begin{aligned}\lim_{j\to\infty} \langle f, u_{k_j}\rangle &= \lim_{j\to\infty} \Big[\int_\Omega u_{k_j} v_0\, dx + \sum_{\ell=1}^{\rm n} \int_\Omega \frac{\partial u_{k_j}}{\partial x_\ell} v_\ell\, dx\Big] = \int_\Omega u v_0\, dx + \sum_{\ell=1}^{\rm n} \int_\Omega \frac{\partial u}{\partial x_\ell} v_\ell\, dx \\ &= \langle f, u\rangle\,.\end{aligned}$$

The argument above establishes the following

Theorem 2.72. *Let $\{u_k\}_{k=1}^\infty$ be a bounded sequence in $W^{1,p}(\Omega)$ for $1 < p < \infty$. Then there exists a subsequence $\{u_{k_j}\}_{j=1}^\infty$ such that $\{u_{k_j}\}_{j=1}^\infty$ converges weakly in $W^{1,p}(\Omega)$.*

2.15.1 The Div-Curl Lemma

It is well-known that if $u_k \rightharpoonup u$ and $v_k \rightharpoonup v$ in $L^2(\Omega)$, $u_k v_k$ does not necessarily converge to uv weakly in $\mathcal{D}(\Omega)$, not even up to a subsequence. In this sub-section, the weak convergence of the product two weakly convergent sequences in $L^2(\Omega)$ is considered, and the goal is to show the weak convergence of the product of two weakly convergent sequence under certain additional constraints.

Theorem 2.73 (Div-Curl Lemma)**.** *Suppose that $\boldsymbol{u}_k \rightharpoonup \boldsymbol{u}$ and $\boldsymbol{v}_k \rightharpoonup \boldsymbol{v}$ both in $L^2(\mathbb{R}^{\rm n})$, and ${\rm div}\boldsymbol{u}_k$ and ${\rm curl}\boldsymbol{v}_k$ are compact in $H^{-1}(\mathbb{R}^{\rm n})$, where ${\rm n} = 2$ or 3. Then there exists a subsequence $\{k_j\}_{j=1}^\infty$ such that $\boldsymbol{u}_{k_j} \cdot \boldsymbol{v}_{k_j} \to \boldsymbol{u}\cdot\boldsymbol{v}$ in $\mathcal{D}'(\mathbb{R}^{\rm n})$.*

Proof. Let $\boldsymbol{w}_k \in H^2(\mathbb{R}^{\rm n})$ solve

$$\begin{aligned}\boldsymbol{w}_k - \Delta \boldsymbol{w}_k &= \boldsymbol{v}_k \qquad \text{in} \quad \mathbb{R}^{\rm n}\,,\\ \boldsymbol{w}_k &= \boldsymbol{0} \qquad \text{on} \quad \partial\mathbb{R}^{\rm n}\,,\end{aligned}$$

and $\boldsymbol{w}$ be the solution to the equation above with $\boldsymbol{v}$ replacing $\boldsymbol{v}_k$. Then

$$\|\boldsymbol{w}_k\|_{H^2(\mathbb{R}^{\rm n})} \leqslant C\|\boldsymbol{v}_k\|_{L^2(\mathbb{R}^{\rm n})}\,,$$

and the Rellich Theorem (with the help of diagonal process) implies that there exists a subsequence of $\{\boldsymbol{w}_k\}_{k=1}^{\infty}$, still denoted by $\{\boldsymbol{w}_k\}_{k=1}^{\infty}$, such that $\boldsymbol{w}_k \to \boldsymbol{w}$ in $H^1(B(0,R))$ for all $R > 0$. By the compactness of $\mathrm{div}\boldsymbol{u}_k$ and $\mathrm{curl}\boldsymbol{v}_k$ in $H^{-1}(\mathbb{R}^n)$, there exists a subsequence $\{k_j\}_{j=1}^{\infty}$ such that

$$\begin{aligned} \mathrm{div}\boldsymbol{u}_{k_j} &\to \mathrm{div}\boldsymbol{u} \qquad \text{in} \quad H^{-1}(\mathbb{R}^n)\,, \\ \mathrm{curl}\boldsymbol{v}_{k_j} &\to \mathrm{curl}\boldsymbol{v} \qquad \text{in} \quad H^{-1}(\mathbb{R}^n)\,. \end{aligned}$$

Let $\varphi \in \mathcal{D}(\mathbb{R}^n)$ be given. Then

$$-\Delta(\varphi\,\mathrm{curl}\boldsymbol{w}_k) = -\mathrm{curl}\boldsymbol{w}_k(\varphi + \Delta\varphi) - 2\nabla\varphi\cdot\nabla\mathrm{curl}\boldsymbol{w}_k + \varphi\,\mathrm{curl}\boldsymbol{v}_k\,. \tag{2.48}$$

Since $\varphi + \Delta\varphi$ is compactly supported,

$$\mathrm{curl}\boldsymbol{w}_k(\varphi + \Delta\varphi) \to \mathrm{curl}\boldsymbol{w}(\varphi + \Delta\varphi) \qquad \text{in} \quad L^2(\mathbb{R}^n)\,.$$

For the second term on the right-hand side of (2.48), by the definition of the dual space norm,

$$\begin{aligned} \|\nabla\varphi\cdot\nabla\mathrm{curl}(\boldsymbol{w}_{k_j} - \boldsymbol{w})\|_{H^{-1}(\mathbb{R}^n)} &= \sup_{\substack{\psi\in H_0^1(\mathbb{R}^n)\\ \|\psi\|_{H^1(\mathbb{R}^n)}=1}} \big\langle \nabla\varphi\cdot\nabla\mathrm{curl}(\boldsymbol{w}_{k_j} - \boldsymbol{w}), \psi\big\rangle \\ = \sup_{\substack{\psi\in H_0^1(\mathbb{R}^n)\\ \|\psi\|_{H^1(\mathbb{R}^n)}=1}} &\Big[\big\langle\Delta\varphi\,\mathrm{curl}(\boldsymbol{w}_{k_j} - \boldsymbol{w}), \psi\big\rangle + \big\langle\nabla\varphi\otimes\mathrm{curl}(\boldsymbol{w}_{k_j} - \boldsymbol{w}), \nabla\psi\big\rangle\Big] \\ \leqslant 2\|\nabla\varphi\|_{W^{1,\infty}(\mathbb{R}^n)}&\|\mathrm{curl}(\boldsymbol{w}_{k_j} - \boldsymbol{w})\|_{L^2(\mathrm{spt}(\varphi))} \to 0 \quad \text{as} \quad j\to\infty\,. \end{aligned}$$

As a consequence, the right-hand side of (2.48) converges strongly to $-\mathrm{curl}\boldsymbol{w}(\varphi + \Delta\varphi) - 2\nabla\varphi\cdot\nabla\mathrm{curl}\boldsymbol{w} + \varphi\,\mathrm{curl}\boldsymbol{v}$ in $H^{-1}(\mathbb{R}^n)$, and the elliptic estimate suggests that

$$\|\varphi\,\mathrm{curl}(\boldsymbol{w}_{k_j} - \boldsymbol{w})\|_{H^1(\mathbb{R}^n)} \to 0 \quad \text{as} \quad j\to\infty\,. \tag{2.49}$$

Finally, observing that

$$\begin{aligned} &\int_{\mathbb{R}^n} \boldsymbol{u}_{k_j}\cdot\boldsymbol{v}_{k_j}\varphi\,dx \\ &\quad= \int_{\mathbb{R}^n} \boldsymbol{u}_{k_j}\cdot\mathrm{curl}\mathrm{curl}\boldsymbol{w}_{k_j}\varphi\,dx - \int_{\mathbb{R}^n} \boldsymbol{u}_{k_j}\cdot\nabla\mathrm{div}\boldsymbol{w}_{k_j}\varphi\,dx + \int_{\mathbb{R}^n} \boldsymbol{u}_{k_j}\cdot\boldsymbol{w}_{k_j}\varphi\,dx \\ &\quad= \int_{\mathbb{R}^n} \boldsymbol{u}_{k_j}\cdot\mathrm{curl}(\varphi\,\mathrm{curl}\boldsymbol{w}_{k_j})\,dx - \int_{\mathbb{R}^n} \boldsymbol{u}_{k_j}\cdot(\nabla\varphi\times\mathrm{curl}\boldsymbol{w}_{k_j})\,dx \\ &\qquad+ \int_{\mathbb{R}^n} \mathrm{div}\boldsymbol{u}_{k_j}\mathrm{div}\boldsymbol{w}_{k_j}\varphi\,dx + \int_{\mathbb{R}^n} \boldsymbol{u}_{k_j}\cdot\nabla\varphi\,\mathrm{div}\boldsymbol{w}_{k_j}\,dx + \int_{\mathbb{R}^n} \boldsymbol{u}_{k_j}\cdot\boldsymbol{w}_{k_j}\varphi\,dx\,, \end{aligned}$$

we conclude that the right-hand side converges to corresponding terms without k_j; thus it is clear that

$$\lim_{j\to\infty}\int_{\mathbb{R}^n}\boldsymbol{u}_{k_j}\cdot\boldsymbol{v}_{k_j}\varphi\,dx=\int_{\mathbb{R}^n}\boldsymbol{u}\cdot\boldsymbol{v}\varphi\,dx\,. \qquad \square$$

2.16 Exercises

PROBLEM 2.1. Suppose that $1<p<\infty$. If $\tau_y f(x)=f(x-y)$, show that f belongs to $W^{1,p}(\mathbb{R}^n)$ if and only if $\tau_y f$ is a Lipschitz function of y with values in $L^p(\mathbb{R}^n)$; that is,

$$\|\tau_y f-\tau_z f\|_{L^p(\mathbb{R}^n)}\leqslant C|y-z|\,.$$

What happens in the case $p=1$?

PROBLEM 2.2 (Fundamental theorem of calculus for integrands in L^1)**.** Let $f\in L^1(\mathbb{R})$, and set

$$g(x)=\int_{-\infty}^{x}f(y)dy\,. \tag{$\star$}$$

Prove that g is continuous, and show that $\dfrac{dg}{dx}=f$, where $\dfrac{dg}{dx}$ denotes the weak derivative.
(**Hint**: Given $\varphi\in\mathscr{C}_c^\infty(\mathbb{R})$, use $(\star)$ to obtain

$$\int_{\mathbb{R}}\varphi'(x)g(x)\,dx=\int_{\mathbb{R}}\int_{-\infty}^{x}\varphi'(x)f(y)dy\,dx\,.$$

Then write this integral as

$$\lim_{h\to0}\frac{1}{h}\int_{\mathbb{R}}\left[\varphi(x+h)-\varphi(x)\right]g(x)\,dx=-\lim_{h\to0}\frac{1}{h}\int_{\mathbb{R}}\int_{x}^{x+h}f(y)\varphi(x)\,dy\,dx\,.)$$

PROBLEM 2.3 (Sobolev embedding for $W^{n,1}$)**.** Show that $W^{n,1}(\mathbb{R}^n)\subseteq C(\mathbb{R}^n)\cap L^\infty(\mathbb{R}^n)$.
(**Hint**: $u(x)=\displaystyle\int_{-\infty}^{0}\cdots\int_{-\infty}^{0}\frac{\partial^n}{\partial y_1\cdots\partial y_n}u(x+y)dy_1\cdots dy_n$.)

PROBLEM 2.4 (Absolute continuity of weakly differentiable functions)**.** If $u\in W^{1,p}(\mathbb{R}^n)$ for some $p\in[1,\infty)$ and $\dfrac{\partial u}{\partial x_j}=0$, $j=1,...,n$, on a connected open set $\Omega\subseteq\mathbb{R}^n$, show that u is equal a.e. to a constant on Ω. (**Hint**: *Approximate u using that $\eta_\epsilon*u\to u$ in $W^{1,p}(\mathbb{R}^n)$, where η_ϵ is a sequence of standard mollifiers. As we* *showed, given* $\delta > 0$, *we can choose* $\epsilon > 0$ *such that* $\|\eta_\epsilon \ast u - u\|_{W^{1,p}(\mathbb{R}^n)} < \delta$. *Show that* $\dfrac{\partial}{\partial x_j}(\eta_\epsilon * u) = 0$ *on* $\Omega_\epsilon \subset\subset \Omega$, *where* $\Omega_\epsilon \nearrow \Omega$ *as* $\epsilon \to 0$.*)*

More generally, if $\frac{\partial u}{\partial x_j} = f_j \in C(\Omega)$, $1 \leqslant j \leqslant n$, show that u is equal a.e. to a function in $\mathscr{C}^1(\Omega)$.

Problem 2.5. In case n = 1, deduce from Problems 2.2 and 2.4 that, if $u \in L^1_{\text{loc}}(\mathbb{R})$ and if $\dfrac{du}{dx} = f \in L^1(\mathbb{R})$, then

$$u(x) = c + \int_{-\infty}^{x} f(y)dy \qquad \text{a.e. } x \in \mathbb{R}\,,$$

for some constant c.

Problem 2.6 (Fundamental theorem of calculus and mean value theorem for integrals)**.** Show that for $u \in W^{1,1}(0,1)$,

$$u(1) - u(0) = \int_0^1 u'(y)dy\,,$$

and that there exists $\alpha \in (0,1)$ such that

$$u(\alpha) = \int_0^1 u(y)dy\,.$$

Problem 2.7. Let $\Omega := B(0, \frac{1}{2}) \subseteq \mathbb{R}^2$ denote the open ball of radius $\frac{1}{2}$. For $x = (x_1, x_2) \in \Omega$, let

$$u(x_1, x_2) = x_1 x_2 \log\left(|\log(|x|)|\right) \ \text{ where } |x| = \sqrt{x_1^2 + x_2^2}\,.$$

(a) Show that $u \in \mathscr{C}^1(\bar{\Omega})$;

(b) show that $\dfrac{\partial^2 u}{\partial x_j^2} \in C(\bar{\Omega})$ for $j = 1, 2$, but that $u \notin \mathscr{C}^2(\bar{\Omega})$;

(c) show that $u \in H^2(\Omega)$.

Problem 2.8. Prove that $\mathscr{C}_c^\infty(\mathbb{R}^n)$ is dense in $W^{k,p}(\mathbb{R}^n)$ for integers $k \geqslant 0$ and $1 \leqslant p < \infty$.

PROBLEM 2.9. Let η_ϵ denote the *standard mollifier*, and for $u \in H^3(\mathbb{R}^3)$, set $u^\epsilon = \eta_\epsilon * u$. Prove that

$$\|u^\epsilon - u\|_{L^\infty(\mathbb{R}^3)} \leqslant C\sqrt{\epsilon}\|u\|_{H^2(\mathbb{R}^3)}\,,$$

and that

$$\|u^\epsilon - u\|_{L^\infty(\mathbb{R}^3)} \leqslant C\epsilon\|u\|_{H^3(\mathbb{R}^3)}\,.$$

PROBLEM 2.10. Let $\Omega \subseteq \mathbb{R}^2$ denote an open, bounded, subset with smooth boundary. Prove the interpolation inequality:

$$\|Du\|^2_{L^2(\Omega)} \leqslant C\|u\|_{L^2(\Omega)}\|D^2u\|_{L^2(\Omega)} \qquad \forall\, u \in H^2(\Omega) \cap H^1_0(\Omega)\,,$$

where D^2u denotes the Hessian matrix of u, i.e., the matrix of second partial derivatives $\dfrac{\partial^2 u}{\partial x_i \partial x_j}$. Use the fact that $\mathscr{C}^\infty(\bar{\Omega}) \cap H^1_0(\Omega)$ is dense in $u \in H^2(\Omega) \cap H^1_0(\Omega)$.

PROBLEM 2.11. Let $D := B(0,1) \subseteq \mathbb{R}^2$ denote the unit disc, and let

$$u(x) = \big[-\log|x|\big]^\alpha.$$

Prove that the *weak derivative* of u exists for all $\alpha \geqslant 0$.

PROBLEM 2.12. Suppose that $\{f_n\}_{n=1}^\infty$ is a bounded sequence in $H^1(\Omega)$ for $\Omega \subseteq \mathbb{R}^2$ bounded. For which values of p does there exist an $f \in H^1(\Omega)$ such that for a subsequence f_{n_ℓ},

$$f_{n_\ell}\, Df_{n_\ell} \rightharpoonup f\, Df \quad \text{weakly in} \quad L^p(\Omega)\,?$$

PROBLEM 2.13. Suppose that $u_j \rightharpoonup u$ in $W^{1,1}(0,1)$. Show that $u_j \to u$ a.e.

We will use the notation $u'(x) = \dfrac{du}{dx}(x)$ in the following problems.

PROBLEM 2.14. Let $p > 1$ and set $\Omega = (0,1) \subseteq \mathbb{R}$.

(a) Suppose that X, Y, Z are Banach spaces, that X is compactly embedded in Y and that Y is continuously embedded in Z. Show that for all $\epsilon > 0$ there is a constant $C_\epsilon = C(\epsilon)$ such that

$$\|u\|_Y \leqslant \epsilon\|u\|_X + C_\epsilon\|u\|_Z \qquad \forall\, u \in X\,.$$

(**Hint**: *Argue by contradiction.*)

(b) Show that for all $\epsilon > 0$, there exists $C = C(\epsilon, p)$ such that

$$\|u\|_{L^\infty(\Omega)} \leqslant \epsilon \|u'\|_{L^p(\Omega)} + C\|u\|_{L^1(\Omega)} \quad \forall\, u \in W^{1,p}(\Omega)\,.$$

(c) Show that the inequality in (b) fails for $p = 1$ (**Hint**: *Consider the sequence* $u_n(x) = x^n$ *and let* $n \to \infty$.)

(d) For $1 \leqslant q < \infty$, show that there exists $C = C(\epsilon, q)$ such that

$$\|u\|_{L^q(\Omega)} \leqslant \epsilon \|u'\|_{L^1(\Omega)} + C\|u\|_{L^1(\Omega)} \quad \forall\, u \in W^{1,1}(\Omega)\,.$$

Problem 2.15. Let $\Omega = (0,1) \subseteq \mathbb{R}$.

(a) For $\bar{u} = \displaystyle\int_\Omega u(x)\,dx$, show that

$$\|u - \bar{u}\|_{L^\infty(\Omega)} \leqslant \|u'\|_{L^1(\Omega)} \quad \forall\, u \in W^{1,1}(\Omega)\,.$$

(**Hint**: *The average* $\bar{u} = u(x_0)$ *for some* $x_0 \in [0,1]$.)

(b) Show that the constant 1 in (a) is optimal. In particular, show that

$$\sup\left\{\|u - \bar{u}\|_{L^\infty(\Omega)} \,\middle|\, u \in W^{1,1}(\Omega) \text{ and } \|u'\|_{L^1(\Omega)} = 1\right\} = 1\,.$$

(**Hint**: *Consider a sequence* $u_n \in \mathscr{C}^\infty(\overline{\Omega})$ *such that* $u_n' \geqslant 0$ *on* $(0,1)$ *for all* $n \in \mathbb{N}$, $u_n(x) = 0$ *for all* $x \in \left[0, 1 - \frac{1}{n}\right]$ *for all* $n \in \mathbb{N}$.)

(c) Show that the supremum in (b) is not achieved, so that there exists no function $u \in W^{1,1}(\Omega)$ such that

$$\|u - \bar{u}\|_{L^\infty(\Omega)} = 1 \text{ and } \|u'\|_{L^1(\Omega)} = 1\,.$$

(d) Prove that

$$\|u\|_{L^\infty(\Omega)} \leqslant \frac{1}{2}\|u'\|_{L^1(\Omega)} \quad \forall\, u \in W_0^{1,1}(\Omega)\,.$$

(**Hint**: *Use that* $|u(x) - u(0)| \leqslant \displaystyle\int_0^x |u'(y)|dy$ *and* $|u(x) - u(1)| \leqslant \displaystyle\int_x^1 |u'(y)|dy$.)

(e) Show that $\frac{1}{2}$ is the best constant in (d). Is it achieved?

(**Hint**: *Fix* $\bar{x} \in \Omega$ *and consider a function* $u \in W_0^{1,1}(\Omega)$ *which is increasing on* $(0, \bar{x})$, *decreasing on* $(\bar{x}, 1)$, *with* $u(\bar{x}) = 1$.)

(f) Show that for $1 \leqslant q \leqslant \infty$ and $1 \leqslant p \leqslant \infty$,

$$\|u - \bar{u}\|_{L^q(\Omega)} \leqslant C\|u'\|_{L^p(\Omega)} \quad \forall\, u \in W^{1,p}(\Omega)\,,$$

and

$$\|u\|_{L^q(\Omega)} \leqslant C\|u'\|_{L^p(\Omega)} \quad \forall\, u \in W_0^{1,p}(\Omega)\,.$$

Prove that the best constants in these two inequalities are achieved when $1 \leqslant q \leqslant \infty$ and $1 < p \leqslant \infty$.

(**Hint**: *Minimize* $\|u'\|_{L^p(\Omega)}$ *in the class* $u \in W^{1,p}(\Omega)$ *such that* $\|u - \bar{u}\|_{L^q(\Omega)} = 1$ (*respectively,* $u \in W_0^{1,p}(\Omega)$ *such that* $\|u\|_{L^q(\Omega)} = 1$.))

Problem 2.16. Let $\Omega = (0, 1) \subseteq \mathbb{R}$.

(a) Suppose that $u \in W^{1,p}(\Omega)$ with $1 < p < \infty$. Show that if $u(0) = 0$, then $\dfrac{u(x)}{x} \in L^p(\Omega)$ and *Hardy's inequality* holds:

$$\left\| \frac{u}{x} \right\|_{L^p(\Omega)} \leqslant \frac{p}{p-1} \|u'\|_{L^p(\Omega)}\,.$$

(b) On the other hand, suppose that $u \in W^{1,p}(\Omega)$ with $1 \leqslant p < \infty$ and that $\dfrac{u(x)}{x} \in L^p(\Omega)$. Show that $u(0) = 0$.

(**Hint**: *Argue by contradiction.*)

(c) Let $u(x) = \dfrac{1}{1 + |\log x|}$. Show that $u \in W^{1,1}(\Omega)$, $u(0) = 0$, but $\dfrac{u(x)}{x} \notin L^1(\Omega)$.

(d) Suppose that $u \in W^{1,p}(\Omega)$ for $1 \leqslant p < \infty$ and $u(0) = 0$. Let $\xi \in \mathscr{C}^\infty(\mathbb{R})$ denote any function satisfying $\xi(x) = 0$ for all $-\infty < x \leqslant 1$ and $\xi(x) = 1$ for all $x \in [2, \infty)$. Set $\xi_n(x) = \xi(nx)$ and let $u_n(x) = \xi_n(x)u(x)$ for $n \in \mathbb{N}$. Verify that $u_n \in W^{1,p}(\Omega)$ and that $u_n \to u$ in $W^{1,p}(\Omega)$ as $n \to \infty$.

(**Hint**: *Consider the cases* $p = 1$ *and* $p > 1$ *separately.*)

Problem 2.17. Let $\Omega = (0, 1) \subseteq \mathbb{R}$.

(a) Let $u \in W^{2,p}(\Omega)$ with $1 < p < \infty$. Assume that $u(0) = u'(0) = 0$. Show that $\dfrac{u(x)}{x^2} \in L^p(\Omega)$ and $\dfrac{u'(x)}{x} \in L^p(\Omega)$ with

$$\left\| \frac{u}{x^2} \right\|_{L^p(\Omega)} + \left\| \frac{u}{x} \right\|_{L^p(\Omega)} \leqslant \frac{p}{p-1} \|u''\|_{L^p(\Omega)}\,.$$

(b) Show then that $v := \dfrac{u}{x} \in W^{1,p}(\Omega)$ with $v(0) = 0$.

(c) With u as in (a), set $u_n = \xi_n u$ as in Problem 2.16(d). Verify that $u_n \in W^{2,p}(\Omega)$ and that $u_n \to u$ in $W^{2,p}(\Omega)$ as $n \to \infty$.

(d) For integers $k \geqslant 1$ and $1 < p < \infty$, suppose that $u \in X^k$, where

$$X^k = \{u \in W^{k,p}(\Omega) \ : \ D^\alpha u(0) = 0\,, |\alpha| \leqslant k-1\}\,.$$

Show that $\frac{u}{x^k} \in L^p(\Omega)$ and that $\frac{u}{x^{k-1}} \in X^1$.

(**Hint**: *Use an induction argument on* k.)

(e) Assume that $u \in X^k$ and show that

$$w = \frac{D^j u}{x^{k-j-i}} \in X^i \quad \forall \text{ integers } i, j\,,\ j \geqslant 0\,, i \geqslant 1\,, i+j \leqslant k-1\,.$$

(f) With u as in (d) and ξ_n as in (c), show that $\xi_n u \in W^{k,p}(\Omega)$ and that $u_n \to u$ in $W^{k,p}(\Omega)$ as $n \to \infty$.

(g) Let $W_0^{k,p}(\Omega)$ denote the closure of $\mathscr{C}_c^\infty(\Omega)$. Show that

$$W_0^{k,p}(\Omega) = \{u \in W^{k,p}(\Omega) \ : \ u = Du = \cdots = D^{k-1}u = 0 \text{ on } \partial\Omega\}\,.$$

(Note well the difference between $W^{k,p}(\Omega) \cap W_0^{1,p}(\Omega)$ and $W_0^{1,p}(\Omega)$ when $k \geqslant 2$.)

(h) Assume now that $u \in W^{2,1}(\Omega)$ with $u(0) = u'(0) = 0$. Set

$$v(x) := \begin{cases} \dfrac{u}{x} & \text{if} \quad x \in (0,1] \\ 0 & \text{if} \qquad x = 0\,. \end{cases}$$

Verify that $v \in C([0,1])$ and prove that $v \in W^{1,1}(\Omega)$.

(**Hint**: *Use the fact that* $v'(x) = \dfrac{1}{x^2}\displaystyle\int_0^x u''(y)dy$.)

(i) Construct an example of a function $u \in W^{2,1}(\Omega)$ satisfying $u(0) = u'(0) = 0$, but with $\dfrac{u}{x^2} \notin L^1(\Omega)$ and $\dfrac{u'}{x} \notin L^1(\Omega)$.

(**Hint**: *Use Problem* 2.16(c).)

Chapter 3

The Fourier Transform

The Fourier transform is one of the most powerful and fundamental tools in linear analysis, converting constant-coefficient linear differential operators into multiplication by polynomials. In this section, we define the Fourier transform, first on $L^1(\mathbb{R}^n)$ functions, next (and miraculously) on $L^2(\mathbb{R}^n)$ functions, and finally on the space of tempered distributions.

3.1 Fourier Transform on $L^1(\mathbb{R}^n)$ and the Space $\mathscr{S}(\mathbb{R}^n)$

Definition 3.1. For all $f \in L^1(\mathbb{R}^n)$, the Fourier transform of f, denoted by $\mathscr{F}f$ or $\widehat{f}$, is defined by

$$(\mathscr{F}f)(\xi) = \widehat{f}(\xi) = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} f(x) e^{-ix\cdot\xi}\, dx\,.$$

It is clear that $\mathscr{F} : L^1(\mathbb{R}^n) \to L^\infty(\mathbb{R}^n)$. In fact,

$$\|\mathscr{F}f\|_{L^\infty(\mathbb{R}^n)} \leqslant (2\pi)^{-\frac{n}{2}} \|f\|_{L^1(\mathbb{R}^n)}\,.$$

Definition 3.2. The Schwartz space is the collection of smooth functions of rapid decay denoted by

$$\mathscr{S}(\mathbb{R}^n) = \big\{ u \in \mathscr{C}^\infty(\mathbb{R}^n) \,\big|\, x^\beta D^\alpha u \in L^\infty(\mathbb{R}^n) \ \ \forall\, \alpha, \beta \in \mathbb{N}^n \big\}.$$

Elements in $\mathscr{S}(\mathbb{R}^n)$ are called Schwartz functions.

It follows from the definition of the Fourier transform that

$$\mathscr{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)\,,$$

and that, in particular,

$$\xi^\alpha D_\xi^\beta \widehat{f} = (-i)^{|\alpha|}(-1)^{|\beta|}\mathscr{F}(D_x^\alpha x^\beta f)\,.$$

The Schwartz space $\mathscr{S}(\mathbb{R}^{\mathrm{n}})$ is also known as the space of rapidly decreasing functions; thus, after multiplying by any polynomial functions $\mathcal{P}(x)$,

$$\mathcal{P}(x)D^\alpha u(x) \to 0 \text{ as } x \to \infty \text{ for all } \alpha \in \mathbb{N}^{\mathrm{n}}.$$

The classical space of test functions $\mathscr{D}(\mathbb{R}^{\mathrm{n}}) := \mathscr{C}_c^\infty(\mathbb{R}^{\mathrm{n}}) \subseteq \mathscr{S}(\mathbb{R}^{\mathrm{n}})$. The prototype element of $\mathscr{S}(\mathbb{R}^{\mathrm{n}})$ is $e^{-|x|^2}$ which is not compactly supported, but has rapidly decreasing derivatives.

The reader is encouraged to verify the following basic properties of $\mathscr{S}(\mathbb{R}^{\mathrm{n}})$:

1. $\mathscr{S}(\mathbb{R}^{\mathrm{n}})$ is a vector space.
2. $\mathscr{S}(\mathbb{R}^{\mathrm{n}})$ is an algebra under the pointwise product of functions.
3. $\mathcal{P}u \in \mathscr{S}(\mathbb{R}^{\mathrm{n}})$ for all $u \in \mathscr{S}(\mathbb{R}^{\mathrm{n}})$ and all polynomial functions $\mathcal{P}$.
4. $\mathscr{S}(\mathbb{R}^{\mathrm{n}})$ is closed under differentiation.
5. $\mathscr{S}(\mathbb{R}^{\mathrm{n}})$ is closed under translations and multiplication by complex exponentials $e^{ix\cdot\xi}$.
6. $\mathscr{S}(\mathbb{R}^{\mathrm{n}}) \subseteq L^1(\mathbb{R}^{\mathrm{n}})$ (since if $u \in \mathscr{S}(\mathbb{R}^{\mathrm{n}})$, $|u(x)| \leqslant C(1+|x|)^{-(\mathrm{n}+1)}$ for some $C > 0$, and $(1+|x|)^{-(\mathrm{n}+1)}dx$ decays like $|x|^{-2}$ as $|x| \to \infty$).

Definition 3.3. For all $f \in L^1(\mathbb{R}^{\mathrm{n}})$, we define operator $\mathscr{F}^*$ by

$$(\mathscr{F}^* f)(x) = (2\pi)^{-\frac{\mathrm{n}}{2}} \int_{\mathbb{R}^{\mathrm{n}}} f(\xi) e^{ix\cdot\xi} d\xi\,.$$

The function $\mathscr{F}^* f$ sometimes is also denoted by $\check{f}$.

Lemma 3.4. *$(\mathscr{F}u, v)_{L^2(\mathbb{R}^{\mathrm{n}})} = (u, \mathscr{F}^* v)_{L^2(\mathbb{R}^{\mathrm{n}})}$ for all $u, v \in \mathscr{S}(\mathbb{R}^{\mathrm{n}})$.*

Recall that the $L^2(\mathbb{R}^{\mathrm{n}})$ inner-product for complex-valued functions is given by $(u,v)_{L^2(\mathbb{R}^{\mathrm{n}})} = \displaystyle\int_{\mathbb{R}^{\mathrm{n}}} u(x)\overline{v(x)}\,dx$.

Proof. Since $u, v \in \mathscr{S}(\mathbb{R}^n)$, by Fubini's Theorem,

$$\begin{aligned}(\mathscr{F}u, v)_{L^2(\mathbb{R}^n)} &= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} u(x)e^{-ix\cdot\xi}\,dx\,\overline{v(\xi)}\,d\xi \\ &= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} u(x)\overline{e^{ix\cdot\xi}v(\xi)}\,d\xi\,dx \\ &= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} u(x)\int_{\mathbb{R}^n} \overline{e^{ix\cdot\xi}v(\xi)}\,d\xi\,dx = (u, \mathscr{F}^*v)_{L^2(\mathbb{R}^n)}\,. \qquad \square\end{aligned}$$

Theorem 3.5. *$\mathscr{F}^*\mathscr{F} = \mathrm{Id} = \mathscr{F}\mathscr{F}^*$ on $\mathscr{S}(\mathbb{R}^n)$.*

Proof. We first prove that for all $f \in \mathscr{S}(\mathbb{R}^n)$, $(\mathscr{F}^*\mathscr{F}f)(x) = f(x)$.

$$\begin{aligned}(\mathscr{F}^*\mathscr{F}f)(x) &= (2\pi)^{-n}\int_{\mathbb{R}^n} e^{i\xi\cdot x}\Big(\int_{\mathbb{R}^n} e^{-iy\cdot\xi}f(y)dy\Big)d\xi \\ &= (2\pi)^{-n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{i(x-y)\cdot\xi}f(y)\,dy\,d\xi\,.\end{aligned}$$

Since $\mathscr{F}f \in \mathscr{S}(\mathbb{R}^n)$, by the dominated convergence theorem,

$$(\mathscr{F}^*\mathscr{F}f)(x) = \lim_{\epsilon\to 0}(2\pi)^{-n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-\epsilon|\xi|^2}e^{i(x-y)\cdot\xi}f(y)\,dy\,d\xi\,.$$

For all $\epsilon > 0$, the *convergence factor* $e^{-\epsilon|\xi|^2}$ allows us to interchange the order of integration, so that by Fubini's theorem,

$$(\mathscr{F}^*\mathscr{F}f)(x) = \lim_{\epsilon\to 0}(2\pi)^{-n}\int_{\mathbb{R}^n} f(y)\Big(\int_{\mathbb{R}^n} e^{-\epsilon|\xi|^2}e^{i(y-x)\cdot\xi}\,d\xi\Big)dy\,.$$

We define the integral kernel

$$p_\epsilon(x) = (2\pi)^{-n}\int_{\mathbb{R}^n} e^{-\epsilon|\xi|^2+ix\cdot\xi}d\xi$$

Then $\mathscr{F}^*\mathscr{F}f = \lim_{\epsilon\to 0} p_\epsilon * f$. Let $p(x) = p_1(x) = (2\pi)^{-n}\int_{\mathbb{R}^n} e^{-|\xi|^2+ix\cdot\xi}d\xi$. Then

$$p\Big(\frac{x}{\sqrt{\epsilon}}\Big) = (2\pi)^{-n}\int_{\mathbb{R}^n} e^{-|\xi|^2+ix\cdot\xi/\sqrt{\epsilon}}d\xi = (2\pi)^{-n}\int_{\mathbb{R}^n} e^{-|\xi|^2+ix\cdot\xi}\epsilon^{\frac{n}{2}}d\xi = \epsilon^{\frac{n}{2}}p_\epsilon(x)\,.$$

We claim that

$$p_\epsilon(x) = \frac{1}{(4\pi\epsilon)^{\frac{n}{2}}}e^{-\frac{|x|^2}{4\epsilon}} \quad \text{and that} \quad \int_{\mathbb{R}^n} p(x)\,dx = 1\,. \tag{3.1}$$

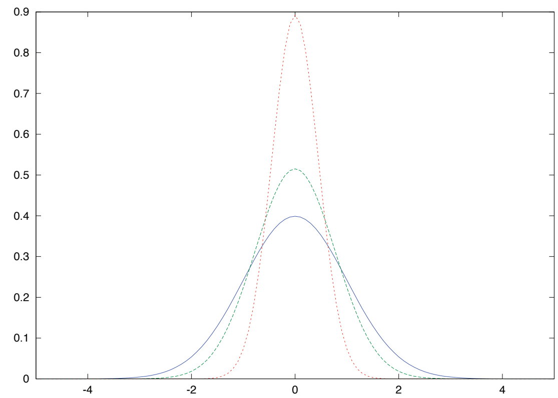


Figure 3.1: As $\epsilon \to 0$, the sequence of functions p_ϵ becomes more localized about the origin.

Given (3.1), then for all $f \in \mathscr{S}(\mathbb{R}^{\rm n})$, $p_\epsilon * f \to f$ uniformly as $\epsilon \to 0$, which shows that $\mathscr{F}^*\mathscr{F} = \mathrm{Id}$, and similar argument shows that $\mathscr{F}\mathscr{F}^* = \mathrm{Id}$. (Note that this follows from the proof of Theorem 1.42, since the standard mollifiers η_ϵ can be replaced by the sequence p_ϵ and all assertions of the theorem continue to hold, for if (3.1) is true, then even though p_ϵ does not have compact support, $\displaystyle\int_{B(0,\delta)^\complement} p_\epsilon(x)\,dx \to 0$ as $\epsilon \to 0$ for all $\delta > 0$.)

Thus, it remains to prove (3.1). It suffices to consider the case $\epsilon = \dfrac{1}{2}$; then by definition

$$p_{\frac{1}{2}}(x) = (2\pi)^{-{\rm n}} \int_{\mathbb{R}^{\rm n}} e^{ix\cdot\xi} e^{-\frac{|\xi|^2}{2}} d\xi = \mathscr{F}\Big((2\pi)^{-\frac{\rm n}{2}} e^{-\frac{|\cdot|^2}{2}}\Big)(x)\,.$$

In order to prove that $p_{\frac{1}{2}}(x) = (2\pi)^{-\frac{\rm n}{2}} e^{-\frac{|x|^2}{2}}$, we must show that with the Gaussian function $G(x) = (2\pi)^{-\frac{\rm n}{2}} e^{-\frac{|x|^2}{2}}$,

$$G(\xi) = \mathscr{F}(G)(\xi)\,.$$

By the multiplicative property of the exponential,

$$e^{-|\xi|^2/2} = e^{-\xi_1^2/2} \cdots e^{-\xi_n^2/2}\,,$$

it suffices to consider the case that ${\rm n} = 1$. Then the Gaussian satisfies the differential equation

$$\frac{d}{dx} G(x) + xG(x) = 0\,.$$

Computing the Fourier transform, we see that

$$-i\frac{d}{d\xi}\widehat{G}(x) - i\xi\widehat{G}(x) = 0\,.$$

Thus,

$$\widehat{G}(\xi) = Ce^{-\frac{\xi^2}{2}}\,.$$

To compute the constant C,

$$C = \widehat{G}(0) = (2\pi)^{-1}\int_{\mathbb{R}} e^{\frac{x^2}{2}}\,dx = (2\pi)^{-\frac{1}{2}}$$

which follows from the fact that

$$\int_{\mathbb{R}} e^{\frac{x^2}{2}}\,dx = (2\pi)^{\frac{1}{2}}\,. \tag{3.2}$$

To prove (3.2), one can again rely on the multiplication property of the exponential to observe that

$$\int_{\mathbb{R}} e^{x_1^2/2}\,dx_1 \int_{\mathbb{R}} e^{x_2^2/2}\,dx_2 = \int_{\mathbb{R}^2} e^{(x_1^2+x_2^2)/2}\,dx = \int_0^{2\pi}\int_0^\infty e^{-\frac{r^2}{2}}r\,dr\,d\theta = 2\pi\,. \qquad \square$$

It follows from Lemma 3.4 that for all $u, v \in \mathscr{S}(\mathbb{R}^{\mathrm{n}})$,

$$(\mathscr{F}u, \mathscr{F}v)_{L^2(\mathbb{R}^{\mathrm{n}})} = (u, \mathscr{F}^*\mathscr{F}v)_{L^2(\mathbb{R}^{\mathrm{n}})} = (u, v)_{L^2(\mathbb{R}^{\mathrm{n}})}\,.$$

Thus, we have established the *Plancheral theorem* on $\mathscr{S}(\mathbb{R}^{\mathrm{n}})$.

Theorem 3.6 (Plancheral's theorem)**.** *$\mathscr{F} : \mathscr{S}(\mathbb{R}^{\mathrm{n}}) \to \mathscr{S}(\mathbb{R}^{\mathrm{n}})$ is an isomorphism with inverse $\mathscr{F}^*$ preserving the $L^2(\mathbb{R}^{\mathrm{n}})$ inner-product.*

3.2 The Topology on $\mathscr{S}(\mathbb{R}^{\mathrm{n}})$ and Tempered Distributions

An alternative to Definition 3.2 can be stated as follows:

Definition 3.7 (The space $\mathscr{S}(\mathbb{R}^{\mathrm{n}})$)**.** Setting $\langle x\rangle = \sqrt{1+|x|^2}$,

$$\mathscr{S}(\mathbb{R}^{\mathrm{n}}) = \left\{u \in \mathscr{C}^\infty(\mathbb{R}^{\mathrm{n}}) \,\middle|\, \langle x\rangle^k |D^\alpha u| \leqslant C_{k,\alpha} \quad \forall\, k \in \mathbb{N}\right\}.$$

The space $\mathscr{S}(\mathbb{R}^{\mathrm{n}})$ has a Fréchet topology determined by semi-norms.

Definition 3.8 (Topology on $\mathscr{S}(\mathbb{R}^n)$)**.** For $k \in \mathbb{N}$, define the semi-norm

$$p_k(u) = \sup_{x\in\mathbb{R}^n, |\alpha|\leqslant k} \langle x\rangle^k |D^\alpha u(x)|\,,$$

and the metric on $\mathscr{S}(\mathbb{R}^n)$

$$d(u,v) = \sum_{k=0}^{\infty} 2^{-k} \frac{p_k(u-v)}{1+p_k(u-v)}\,.$$

The space $(\mathscr{S}(\mathbb{R}^n), d)$ is a Fréchet space.

Definition 3.9 (Convergence in $\mathscr{S}(\mathbb{R}^n)$)**.** A sequence $u_j \to u$ in $\mathscr{S}(\mathbb{R}^n)$ if $p_k(u_j - u) \to 0$ as $j \to \infty$ for all $k \in \mathbb{N}$.

Definition 3.10 (Tempered Distributions)**.** A linear map $T : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ is continuous if for *some* $k \in \mathbb{N}$, there exists some constant C_k such that

$$\big|\langle T, u\rangle\big| \leqslant C_k p_k(u) \quad \forall\, u \in \mathscr{S}(\mathbb{R}^n)\,.$$

The space of continuous linear functionals on $\mathscr{S}(\mathbb{R}^n)$ is denoted by $\mathscr{S}'(\mathbb{R}^n)$. Elements of $\mathscr{S}'(\mathbb{R}^n)$ are called tempered distributions.

Definition 3.11 (Convergence in $\mathscr{S}'(\mathbb{R}^n)$)**.** A sequence $T_j \rightharpoonup T$ in $\mathscr{S}'(\mathbb{R}^n)$ if $\langle T_j, u\rangle \to \langle T, u\rangle$ for all $u \in \mathscr{S}(\mathbb{R}^n)$.

For $1 \leqslant p \leqslant \infty$, there is a natural injection of $L^p(\mathbb{R}^n)$ into $\mathscr{S}'(\mathbb{R}^n)$ given by

$$\langle f, u\rangle = \int_{\mathbb{R}^n} f(x)u(x)\,dx \qquad \forall\, u \in \mathscr{S}(\mathbb{R}^n)\,.$$

Any finite measure on $\mathbb{R}^n$ provides an element of $\mathscr{S}'(\mathbb{R}^n)$. The basic example of such a finite measure is the Dirac delta 'function' defined as follows:

$$\langle \delta_0, u\rangle = u(0) \;\text{ or, more generally, }\; \langle \delta_x, u\rangle = u(x) \quad \forall\, u \in \mathscr{S}(\mathbb{R}^n)\,.$$

We shall often use δ to denote the Dirac delta distribution δ_0.

Definition 3.12. The distributional derivative $D : \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ is defined by the relation

$$\langle DT, u\rangle = -\langle T, Du\rangle \qquad \forall\, u \in \mathscr{S}(\mathbb{R}^n)\,.$$

More generally, the αth distributional derivative exists in $\mathscr{S}'(\mathbb{R}^n)$ and is defined by

$$\langle D^\alpha T, u\rangle = (-1)^{|\alpha|}\langle T, D^\alpha u\rangle \qquad \forall\, u \in \mathscr{S}(\mathbb{R}^n)\,.$$

Multiplication by $f \in \mathscr{S}(\mathbb{R}^n)$ preserves $\mathscr{S}'(\mathbb{R}^n)$; in particular, if $T \in \mathscr{S}'(\mathbb{R}^n)$, then $fT \in \mathscr{S}'(\mathbb{R}^n)$ and is defined by

$$\langle fT, u\rangle = \langle T, fu\rangle \qquad \forall\, u \in \mathscr{S}(\mathbb{R}^n)\,.$$

Example 3.13. Let $H := \mathbf{1}_{[0,\infty)}$ denote the Heavyside function. Then

$$\frac{dH}{dx} = \delta \quad \text{in} \quad \mathscr{S}'(\mathbb{R}^n)\,.$$

This follows since for all $u \in \mathscr{S}(\mathbb{R}^n)$,

$$\langle \frac{dH}{dx}, u\rangle = -\langle H, \frac{du}{dx}\rangle = -\int_0^\infty \frac{du}{dx}\,dx = u(0) = \langle \delta, u\rangle\,.$$

Example 3.14 (Distributional derivative of Dirac measure)**.**

$$\left\langle \frac{d\delta}{dx}, u\right\rangle = -\frac{du}{dx}(0) \;\; \forall\, u \in \mathscr{S}(\mathbb{R}^n)\,.$$

3.3 Fourier Transform on $\mathscr{S}'(\mathbb{R}^n)$

Definition 3.15. Define $\mathscr{F} : \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ by

$$\langle \mathscr{F}T, u\rangle = \langle T, \mathscr{F}u\rangle \qquad \forall\, u \in \mathscr{S}(\mathbb{R}^n)\,,$$

with the analogous definition for $\mathscr{F}^* : \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$.

Theorem 3.16. *$\mathscr{F}^*\mathscr{F} = \mathrm{Id} = \mathscr{F}\mathscr{F}^*$ on $\mathscr{S}'(\mathbb{R}^n)$.*

Proof. By Definition 3.15, for all $u \in \mathscr{S}(\mathbb{R}^n)$

$$\langle \mathscr{F}\mathscr{F}^*T, u\rangle = \langle \mathscr{F}^*w, \mathscr{F}u\rangle = \langle T, \mathscr{F}^*\mathscr{F}u\rangle = \langle T, u\rangle\,,$$

the last equality following from Theorem 3.5. □

Example 3.17 (Fourier transform of δ)**.** We claim that $\mathscr{F}\delta = (2\pi)^{-\frac{n}{2}}$. According to Definition 3.15, for all $u \in \mathscr{S}(\mathbb{R}^n)$,

$$\langle \mathscr{F}\delta, u\rangle = \langle \delta, \mathscr{F}u\rangle = \mathscr{F}u(0) = \int_{\mathbb{R}^n} (2\pi)^{-\frac{n}{2}}\, u(x)\,dx\,,$$

so that $\mathscr{F}\delta = (2\pi)^{-\frac{n}{2}}$.

Example 3.18. The same argument shows that $\mathscr{F}^*(\delta) = (2\pi)^{-\frac{n}{2}}$ so that $\mathscr{F}^*[(2\pi)^{\frac{n}{2}}\delta] = 1$. Using Theorem 3.16, we see that $\mathscr{F}(1) = (2\pi)^{\frac{n}{2}}\delta$. This demonstrates nicely the identity

$$|\xi^\alpha \widehat{u}(\xi)| = |\mathscr{F}(D^\alpha u)(\xi)|.$$

In other the words, the smoother the function $x \mapsto u(x)$ is, the faster $\xi \mapsto \widehat{u}(\xi)$ must decay.

Remark 3.19. A function $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ generates a distribution $f \in \mathscr{S}'(\mathbb{R}^n)$. We now show that the Fourier transform given by Definition 3.15 agrees with the Fourier transform of a function.

For $\varphi \in \mathscr{S}(\mathbb{R}^n)$,

$$\begin{aligned}\langle \widehat{f}, \varphi \rangle \equiv \langle f, \widehat{\varphi} \rangle &= \int_{\mathbb{R}^n} f(\xi)\widehat{\varphi}(\xi)\, d\xi = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(\xi)\varphi(x)e^{-ix\cdot\xi}\, dx\, d\xi \\ &= \lim_{m\to\infty} \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\Omega_m}\int_{\mathbb{R}^n} f(\xi)\varphi(x)e^{-ix\cdot\xi}\, dx\, d\xi\,,\end{aligned}$$

where Ω_m is an increasing sequence of bounded sets such that $\bigcup\limits_{m=1}^{\infty} \Omega_m = \mathbb{R}^n$. Letting $f_m = \mathbf{1}_{\Omega_m} f$ or $\widehat{f_m}(x) = \dfrac{1}{(2\pi)^{\frac{n}{2}}} \displaystyle\int_{\Omega_m} f(\xi)e^{-ix\cdot\xi} d\xi$, we find that

$$\langle \lim_{m\to\infty} \widehat{f_m}, \varphi \rangle = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \lim_{m\to\infty} \int_{\Omega_m} f(\xi)e^{-ix\cdot\xi}\varphi(x)\, d\xi\; dx\,.$$

Therefore, if we define $\widehat{f} = \lim\limits_{m\to\infty} \widehat{f_m}$ whenever the limit makes sense, then we have the following identity

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(\xi)\varphi(x)e^{-ix\cdot\xi} d\xi\; dx = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(\xi)\varphi(x)e^{-ix\cdot\xi}\, dx\, d\xi\,, \tag{3.3}$$

and $\widehat{f}$ agrees with the Fourier transform of a function. Note that (3.3) shows that we can interchange the order of integration even though $f(\xi)\varphi(x)e^{-ix\cdot\xi}$ does not belong to $L^1(\mathbb{R}^{2n})$.

3.4 The Fourier Transform on $L^2(\mathbb{R}^n)$

In Theorem 1.42, we proved that $\mathscr{C}_c^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for $1 \leqslant p < \infty$. Since $\mathscr{C}_c^\infty(\mathbb{R}^n) \subseteq \mathscr{S}(\mathbb{R}^n)$, it follows that $\mathscr{S}(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ as well. Thus, for every

$u \in L^2(\mathbb{R}^n)$, there exists a sequence $\{u_j\}_{j=1}^\infty \subseteq \mathscr{S}(\mathbb{R}^n)$ such that $u_j \to u$ in $L^2(\mathbb{R}^n)$, so that by Plancheral's Theorem 3.6,

$$\|\widehat{u}_j - \widehat{u}_k\|_{L^2(\mathbb{R}^n)} = \|u_j - u_k\|_{L^2(\mathbb{R}^n)} < \epsilon\,.$$

It follows from the completeness of $L^2(\mathbb{R}^n)$ that the sequence $\widehat{u}_j$ *converges* in $L^2(\mathbb{R}^n)$.

Definition 3.20 (Fourier transform on $L^2(\mathbb{R}^n)$)**.** For $u \in L^2(\mathbb{R}^n)$ let $\{u_j\}_{j=1}^\infty$ denote an approximating sequence in $\mathscr{S}(\mathbb{R}^n)$. Define the Fourier transform as follows:

$$\mathscr{F}u = \widehat{u} = \lim_{j\to\infty} \widehat{u}_j\,.$$

Note well that $\mathscr{F}$ on $L^2(\mathbb{R}^n)$ is well-defined, as the limit is independent of the approximating sequence. In particular,

$$\|\widehat{u}\|_{L^2(\mathbb{R}^n)} = \lim_{j\to\infty} \|\widehat{u}_j\|_{L^2(\mathbb{R}^n)} = \lim_{j\to\infty} \|u_j\|_{L^2(\mathbb{R}^n)} = \|u\|_{L^2(\mathbb{R}^n)}\,.$$

By the polarization identity

$$\begin{aligned}(u,v)_{L^2(\mathbb{R}^n)} = \frac{1}{2}\Big(&\|u+v\|^2_{L^2(\mathbb{R}^n)} - i\|u+iv\|^2_{L^2(\mathbb{R}^n)} \\ &- (1-i)\|u\|^2_{L^2(\mathbb{R}^n)} - (1-i)\|v\|^2_{L^2(\mathbb{R}^n)}\Big)\end{aligned}$$

we have proved the Plancheral theorem[1] on $L^2(\mathbb{R}^n)$:

Theorem 3.21. $(u,v)_{L^2(\mathbb{R}^n)} = (\mathscr{F}u, \mathscr{F}v)_{L^2(\mathbb{R}^n)}$ *for all* $u,v \in L^2(\mathbb{R}^n)$.

3.5 Bounds for the Fourier Transform on $L^p(\mathbb{R}^n)$

We have shown that for $u \in L^1(\mathbb{R}^n)$, $\|\widehat{u}\|_{L^\infty(\mathbb{R}^n)} \leqslant (2\pi)^{-\frac{n}{2}}\|u\|_{L^1(\mathbb{R}^n)}$, and that for $u \in L^2(\mathbb{R}^n)$, $\|\widehat{u}\|_{L^2(\mathbb{R}^n)} = \|u\|_{L^2(\mathbb{R}^n)}$. Applying the Marcinkiewicz Interpolation Theorem (Theorem D.4) (by interpolating p between 1 and 2) yields the following result.

Theorem 3.22 (Hausdorff-Young inequality)**.** *If* $u \in L^p(\mathbb{R}^n)$ *for* $1 \leqslant p \leqslant 2$, *then for* $q = \dfrac{p-1}{p}$, *there exists a constant* C *such that*

$$\|\widehat{u}\|_{L^q(\mathbb{R}^n)} \leqslant C\|u\|_{L^p(\mathbb{R}^n)}\,.$$

[1] The unitarity of the Fourier transform is often called Parseval's theorem in science and engineering fields, based on an earlier (but less general) result that was used to prove the unitarity of the Fourier series.

Returning to the case that $u \in L^1(\mathbb{R}^n)$, not only is $\mathscr{F}u \in L^\infty(\mathbb{R}^n)$, but the transformed function decays at infinity.

Theorem 3.23 (Riemann-Lebesgue "lemma"). *For $u \in L^1(\mathbb{R}^n)$, $\mathscr{F}u$ is continuous and $(\mathscr{F}u)(\xi) \to 0$ as $|\xi| \to \infty$.*

Proof. Let $B_M = B(0, M) \subseteq \mathbb{R}^n$. Since $f \in L^1(\mathbb{R}^n)$, for each $\epsilon > 0$, we can choose M sufficiently large such that

$$\Big|\widehat{f}(\xi) - \int_{B_M} e^{-ix\cdot\xi} f(x)\,dx\Big| < \epsilon\,.$$

Using Lemma 1.36, choose a sequence of simple functions $\{\varphi_j\}_{j=1}^\infty$ which converges to f in $L^1(B_M)$. Then for $j \in \mathbb{N}$ chosen sufficiently large,

$$\Big|\widehat{f}(\xi) - \int_{B_M} \varphi_j(x) e^{-ix\cdot\xi}\,dx\Big| < 2\epsilon\,.$$

Writing $\varphi_j(x) = \sum\limits_{\ell=1}^{N} C_\ell \mathbf{1}_{E_\ell}(x)$, we have that

$$\Big|\widehat{f}(\xi) - \sum_{\ell=1}^{N} C_\ell \int_{E_\ell} \varphi_j(x) e^{-ix\cdot\xi}\,dx\Big| < 2\epsilon\,.$$

By the regularity of the Lebesgue measure μ, for each $\ell \in \{1, ..., N\}$, there exists a compact set K_ℓ and an open set O_ℓ such that $K_\ell \subseteq E_\ell \subseteq O_\ell$ and

$$\mu(O_\ell) - \frac{\epsilon}{2} < \mu(E_\ell) < \mu(K_\ell) + \frac{\epsilon}{2}\,.$$

Since O_ℓ is open, $O_\ell = \bigcup\limits_{\alpha \in A_\ell} \mathcal{V}_\alpha^\ell$ for some open rectangle $\mathcal{V}_\alpha^\ell$ and index set A_ℓ. By the compactness of K_ℓ, $K_\ell \subseteq \bigcup\limits_{j=1}^{N_\ell} \mathcal{V}_{\alpha_j}^\ell$ for some $\{\alpha_1, ..., \alpha_{N_\ell}\} \subseteq A_\ell$; thus

$$\mu\big(E_\ell \backslash \bigcup_{j=1}^{N_\ell} \mathcal{V}_{\alpha_j}^\ell\big) + \mu\big(\bigcup_{j=1}^{N_\ell} \mathcal{V}_{\alpha_j}^\ell \backslash E_\ell\big) < \epsilon\,.$$

It then follows that

$$\Big|\int_{E_\ell} e^{-ix\cdot\xi}\,dx - \int_{\bigcup_{j=1}^{N_\ell} \mathcal{V}_{\alpha_j}^\ell} e^{-ix\cdot\xi}\,dx\Big| < \epsilon\,.$$

On the other hand, for each rectangle $\mathcal{V}^{\ell}_{\alpha_j}$, $\left|\int_{\mathcal{V}^{\ell}_{\alpha_j}} e^{-ix\cdot\xi}\,dx\right| \leqslant \dfrac{C}{\xi_1\cdots\xi_n}$, so

$$\widehat{f}(\xi) \leqslant C\Big(\epsilon + \frac{1}{\xi_1\cdots\xi_n}\Big)\,.$$

Since $\epsilon > 0$ is arbitrary, we see that $\widehat{f}(\xi) \to 0$ as $|\xi| \to \infty$. Continuity of $\mathscr{F}u$ follows easily from the dominated convergence theorem. ▫

3.6 Convolution and the Fourier Transform

Theorem 3.24. *If $u, v \in L^1(\mathbb{R}^{\mathrm{n}})$, then $u * v \in L^1(\mathbb{R}^{\mathrm{n}})$ and*

$$\mathscr{F}(u*v) = (2\pi)^{\frac{\mathrm{n}}{2}}(\mathscr{F}u)(\mathscr{F}v)\,,\quad \mathscr{F}^*(u*v) = (2\pi)^{\frac{\mathrm{n}}{2}}(\mathscr{F}^*u)(\mathscr{F}^*v)\,.$$

Proof. Young's inequality (Theorem 1.47) shows that $u*v \in L^1(\mathbb{R}^{\mathrm{n}})$ so that the Fourier transform is well-defined. The assertion then follows from a direct computation:

$$\begin{aligned}
\mathscr{F}(u*v) &= (2\pi)^{-\frac{\mathrm{n}}{2}} \int_{\mathbb{R}^{\mathrm{n}}} e^{-ix\cdot\xi}(u*v)(x)\,dx \\
&= (2\pi)^{-\frac{\mathrm{n}}{2}} \int_{\mathbb{R}^{\mathrm{n}}}\int_{\mathbb{R}^{\mathrm{n}}} u(x-y)v(y)dy\, e^{-ix\cdot\xi}\,dx \\
&= (2\pi)^{-\frac{\mathrm{n}}{2}} \int_{\mathbb{R}^{\mathrm{n}}}\int_{\mathbb{R}^{\mathrm{n}}} u(x-y)e^{-i(x-y)\cdot\xi}\,dx\, v(x)\, e^{-iy\cdot\xi}\,dy \\
&= (2\pi)^{\frac{\mathrm{n}}{2}}\widehat{u}\widehat{v} \qquad \text{(by Fubini's theorem)}\,.
\end{aligned}$$

That $\mathscr{F}^*(u*v) = (2\pi)^{\frac{\mathrm{n}}{2}}(\mathscr{F}^*u)(\mathscr{F}^*v)$ can be proved in a similar way. ▫

By using Young's inequality (Theorem 1.48) together with the Hausdorff-Young inequality, we can generalize the convolution result to the following

Theorem 3.25. *Suppose that $u \in L^p(\mathbb{R}^{\mathrm{n}})$ and $v \in L^q(\mathbb{R}^{\mathrm{n}})$, and let r satisfy $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1$ for $1 \leqslant p, q, r \leqslant 2$. Then $\mathscr{F}(u*v), \mathscr{F}^*(u*v) \in L^{\frac{r}{r-1}}(\mathbb{R}^{\mathrm{n}})$ and*

$$\mathscr{F}(u*v) = (2\pi)^{\frac{\mathrm{n}}{2}}(\mathscr{F}u)(\mathscr{F}v)\,,\quad \mathscr{F}^*(u*v) = (2\pi)^{\frac{\mathrm{n}}{2}}(\mathscr{F}^*u)(\mathscr{F}^*v)\,.$$

Let $\star$ denote the convolution operator defined by $f \star g = (2\pi)^{-\frac{\mathrm{n}}{2}}(f*g)$. Then the theorem above implies that if $u \in L^p(\mathbb{R}^{\mathrm{n}})$ and $v \in L^q(\mathbb{R}^{\mathrm{n}})$,

$$\mathscr{F}(u\star v) = (\mathscr{F}u)(\mathscr{F}v)\,,\quad \mathscr{F}^*(u\star v) = (\mathscr{F}^*u)(\mathscr{F}^*v)\,.$$

Using this notation, we find that for $f, g \in \mathscr{S}(\mathbb{R}^{\mathrm{n}})$,

$$\langle f \star g, \varphi \rangle = \Big(\frac{1}{\sqrt{2\pi}}\Big)^{\mathrm{n}} \int_{\mathbb{R}^{\mathrm{n}}} \int_{\mathbb{R}^{\mathrm{n}}} f(x-y) g(y) \varphi(x)\, dy\, dx = \langle g, \widetilde{f} \star \varphi \rangle ,$$

where $\widetilde{\ } : \mathscr{S}(\mathbb{R}^{\mathrm{n}}) \to \mathscr{S}(\mathbb{R}^{\mathrm{n}})$ is the reflection operator defined by $\widetilde{f}(x) = f(-x)$. This observation motivates the following

Definition 3.26. Let $f \in \mathscr{S}(\mathbb{R}^{\mathrm{n}})$ and $T \in \mathscr{S}'(\mathbb{R}^{\mathrm{n}})$. The convolution $f \star T$ is the tempered distribution defined by

$$\langle f \star T, \varphi \rangle = \langle T, \widetilde{f} \star \varphi \rangle \qquad \forall\, \varphi \in \mathscr{S}(\mathbb{R}^{\mathrm{n}}) ,$$

Theorem 3.27. *Let $f \in \mathscr{S}(\mathbb{R}^{\mathrm{n}})$ and $T \in \mathscr{S}'(\mathbb{R}^{\mathrm{n}})$. Then $\mathscr{F}(f \star T) = \widehat{f} \cdot \widehat{T}$.*

Proof. Since $\mathscr{F}^{-1}(\widehat{f} \star \widehat{\varphi}) = \check{\widetilde{f}}\varphi$, we have $\widetilde{f} \star \widehat{\varphi} = \mathscr{F}\big(\check{\widetilde{f}} \cdot \varphi\big)$. Moreover, $\check{\widetilde{f}} = \widehat{f}$. As a consquence,

$$\langle \mathscr{F}(f \star T), \varphi \rangle = \langle f \star T, \widehat{\varphi} \rangle = \langle T, \widetilde{f} \star \widehat{\varphi} \rangle = \langle \widehat{T}, \check{\widetilde{f}} \cdot \varphi \rangle = \langle \widehat{T}, \widehat{f} \cdot \varphi \rangle = \langle \widehat{f}\widehat{T}, \varphi \rangle . \qquad \square$$

3.7 An Explicit Computation with the Fourier Transform

As we have shown that $\mathcal{F}^*\mathcal{F} = \mathcal{F}\mathcal{F}^* = \mathrm{Id}$, in this section, we shall denote $\mathcal{F}^*$ by $\mathcal{F}^{-1}$.

The computation of the Green's function for the Laplace operator is an important application of the Fourier transform. For this purpose, we will compute $\widehat{f}$ for the following two cases: (1) $f(x) = e^{-t|x|}$, $t > 0$ and (2) $f(x) = |x|^{\alpha}$, $-\mathrm{n} < \alpha < 0$.

Case (1) In this case, f_1 is rapidly decreasing but not in the Schwartz class $\mathcal{S}(\mathbb{R}^{\mathrm{n}})$. We begin with $\mathrm{n} = 1$. It follows that

$$\begin{aligned} \mathscr{F}(e^{-t|\cdot|})(\xi) &= \int_{-\infty}^{\infty} e^{-t|x|} e^{-ix\cdot\xi} d\mu_1(x) = \int_{-\infty}^{0} e^{x(t-i\xi)} d\mu_1(x) + \int_0^{\infty} e^{x(-t-i\xi)} d\mu_1(x) \\ &= \frac{1}{\sqrt{2\pi}} \Big[\frac{e^{x(t-i\xi)}}{t-i\xi}\Big|_{-\infty}^{0} + \frac{e^{x(-t-i\xi)}}{-t-i\xi}\Big|_0^{\infty} \Big] = \sqrt{\frac{2}{\pi}} \frac{t}{t^2+\xi^2} . \end{aligned}$$

By the inversion formula, we then see that $e^{-t|x|} = \dfrac{1}{\pi}\displaystyle\int_{-\infty}^{\infty} \frac{t}{t^2+\xi^2}\, e^{ix\xi} d\xi$.

In order to study the higher-dimensional cases when $\mathrm{n} > 1$, we begin with the observation that

$$\int_0^{\infty} e^{-st^2} e^{-s\xi^2} ds = \frac{e^{-s(t^2+\xi^2)}}{-(t^2+\xi^2)}\Big|_0^{\infty} = \frac{1}{t^2+\xi^2} . \tag{3.4}$$

With $\lambda = |x| > 0$, we use (3.4) to find that

$$\begin{aligned} e^{-t|x|} &= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{t}{t^2+\xi^2}\ e^{i|x|\xi} d\xi = \frac{1}{\pi}\int_{-\infty}^{\infty} t\Big(\int_0^\infty e^{-st^2}e^{-s\xi^2}ds\Big)e^{i\lambda\xi}d\xi \\ &= \frac{1}{\pi}\int_0^\infty t\Big(\int_{-\infty}^{\infty} e^{-s\xi^2}e^{i|x|\xi}d\xi\Big)e^{-st^2}ds = \int_0^\infty \frac{t}{\sqrt{s\pi}}e^{-st^2}e^{-\frac{|x|^2}{4s}}ds\,. \end{aligned}$$

Now, we compute the Fourier transform with respect to the variable $x = (x_1, ..., x_\mathrm{n})$, and find that

$$\begin{aligned} \mathscr{F}(e^{-t|\cdot|})(\xi) &= \int_0^\infty \frac{t}{\sqrt{s\pi}}e^{-st^2}\mathcal{F}(e^{-\frac{|x|^2}{4s}})ds = \int_0^\infty \frac{t}{\sqrt{s\pi}}(2s)^{\frac{\mathrm{n}}{2}}e^{-s(t^2+|\xi|^2)}ds \\ &= \frac{t}{(t^2+|\xi|^2)^{\frac{\mathrm{n}+1}{2}}}\int_0^\infty \frac{1}{\sqrt{\pi s}}(2s)^{\frac{\mathrm{n}}{2}}e^{-s}ds = \frac{C(\mathrm{n})t}{(t^2+|\xi|^2)^{\frac{\mathrm{n}+1}{2}}}\,, \end{aligned}$$

where the constant $C(\mathrm{n}) = \int_0^\infty \frac{1}{\sqrt{\pi s}}(2s)^{\frac{\mathrm{n}}{2}}ds = \sqrt{\frac{2^\mathrm{n}}{\pi}}\Gamma\big(\frac{\mathrm{n}+1}{2}\big)$, where $\Gamma\big(\frac{\mathrm{n}+1}{2}\big)$ denotes the *Gamma* function. It follows that

$$\mathscr{F}^{-1}(e^{-t|\cdot|})(x) = \mathcal{F}(e^{-t|\cdot|})(-x) = \sqrt{\frac{2^\mathrm{n}}{\pi}}\Gamma(\frac{\mathrm{n}+1}{2})\frac{t}{(t^2+|x|^2)^{\frac{\mathrm{n}+1}{2}}}\,. \tag{3.5}$$

Case (2) For this case, we compute $\mathcal{F}(|\cdot|^\alpha)$, when $-\mathrm{n} < \alpha < 0$. Using the definition of the Gamma-function, we see that

$$\int_0^\infty s^{-\frac{\alpha}{2}-1}e^{-s|x|^2}ds = |x|^\alpha\int_0^\infty s^{-\frac{\alpha}{2}-1}e^{-s}ds = |x|^\alpha\Gamma(-\frac{\alpha}{2})\,,$$

Therefore,

$$\begin{aligned} \mathscr{F}(|\cdot|^\alpha)(\xi) &= \frac{1}{\Gamma(-\frac{\alpha}{2})}\int_0^\infty s^{-\frac{\alpha}{2}-1}\mathscr{F}(e^{-s|\cdot|^2})ds = \frac{1}{2^{\frac{\mathrm{n}}{2}}\Gamma(-\frac{\alpha}{2})}\int_0^\infty s^{-\frac{\alpha}{2}-\frac{\mathrm{n}}{2}-1}e^{-\frac{|\xi|^2}{4s}}ds \\ &= \frac{1}{2^{\frac{\mathrm{n}}{2}}\Gamma(-\frac{\alpha}{2})}\int_0^\infty \Big(\frac{|\xi|^2}{4s}\Big)^{-\frac{\alpha}{2}-\frac{\mathrm{n}}{2}-1}e^{-s}\frac{|\xi|^2}{4s^2}ds = \frac{2^{\alpha+\frac{\mathrm{n}}{2}}\Gamma(\frac{\alpha+\mathrm{n}}{2})}{\Gamma(-\frac{\alpha}{2})}|\xi|^{-\alpha-\mathrm{n}}\,, \end{aligned}$$

where we impose the condition $-\mathrm{n} < \alpha < 0$ to ensure the boundedness of the Γ-function. In particular, for $\mathrm{n} = 3$ and $\alpha = -1$,

$$\mathscr{F}(|\cdot|^{-1})(\xi) = \frac{\sqrt{2}\Gamma(1)}{\Gamma(\frac{1}{2})}|\xi|^{-2} = \sqrt{\frac{2}{\pi}}\,|\xi|^{-2}\,,$$

from which it follows that

$$\mathscr{F}^{-1}(|\cdot|^{-2})(x) = \sqrt{\frac{\pi}{2}}\frac{1}{|x|}\,. \tag{3.6}$$

3.8 Applications to the Poisson, Heat, and Wave Equations

The Poisson equation on $\mathbb{R}^3$

In Theorem 2.63, we proved the existence of unique weak solutions to the Dirichlet problem on a bounded domain Ω. We will now provide an explicit representation for solutions to the Poisson problem on $\mathbb{R}^3$. The issue of uniqueness in this setting will be of interest.

Given the Poisson problem

$$\Delta u = f \quad \text{in} \quad \mathscr{S}'(\mathbb{R}^{\mathrm{n}}),$$

we compute the Fourier transform of both sides to obtain that

$$-|\xi|^2\widehat{u}(\xi) = \widehat{f}(\xi). \tag{3.7}$$

Distributional solutions to (3.7) are not unique; for example,

$$\widehat{u}(\xi) = -\frac{\widehat{f}(\xi)}{|\xi|^2} \text{ and } \widehat{u}(\xi) = -\frac{\widehat{f}(\xi)}{|\xi|^2} + \delta$$

are both solutions. By requiring solutions to have enough decay, such as $u \in L^2(\mathbb{R}^{\mathrm{n}})$ so that $\widehat{u} \in L^2(\mathbb{R}^{\mathrm{n}})$, then we do obtain uniqueness.

We will find an explicit representation for the solution to the Poisson problem when $\mathrm{n} = 3$. If $u \in L^2(\mathbb{R}^3)$, then using (3.6) we see that $\widehat{u}(\xi) = \dfrac{\widehat{f}(\xi)}{|\xi|^2}$; thus

$$u(x) = \mathscr{F}^*\Big(\frac{\widehat{f}(\cdot)}{|\cdot|^2}\Big)(x) = \Big[\mathscr{F}^*(|\cdot|^{-2}) \star \mathscr{F}^*(\widehat{f})\Big](x) = (\Phi * f)(x),$$

where $\Phi(x) = \dfrac{1}{4\pi|x|}$. The function Φ is the so-called *fundamental solution*; more precisely, it is the distributional solution of the equation

$$\Delta\Phi = \delta \quad \text{in} \quad \mathscr{S}'(\mathbb{R}^{\mathrm{n}}).$$

Conceptually

$$\begin{aligned} -\Delta(\Phi * f) &= -\Delta\Phi * f \\ &= \delta_0 * f = f \quad \forall\, f \in \mathscr{C}(\mathbb{R}^{\mathrm{n}}) \text{ whenever } \Phi * f \text{ makes sense,} \end{aligned}$$

where the first equality follows from the fact that

$$\begin{aligned}\langle \Delta(\Phi * f), \widehat{\varphi}\rangle &= (2\pi)^{\frac{n}{2}}\langle -|\xi|^2 \widehat{\Phi}\widehat{f}, \varphi\rangle = (2\pi)^{\frac{n}{2}}\langle \mathscr{F}(\Delta\Phi), \widehat{f}\varphi\rangle \\ &= (2\pi)^{\frac{n}{2}}\langle \Delta\Phi, \mathscr{F}(\widehat{f}\varphi)\rangle = \langle \Delta\Phi, \widetilde{f} * \widehat{\varphi}\rangle = \langle \Delta\Phi * f, \widehat{\varphi}\rangle .\end{aligned}$$

Example 3.28. On $\mathbb{R}^2$, $\Delta(e^{x_1}\cos x_2) = 0$. The function $e^{x_1}\cos x_2$ is not a tempered distribution because it grows too fast as $x_1 \to \infty$. As such, the Fourier transfor of $e^{x_1}\cos x_2$ is not defined.

Using Fourier transform to convert PDE to linear algebraic equations only provides those solutions which do not grow too rapidly at ∞.

3.8.1 The Poisson integral formula on the half-space

Let $\Omega = \mathbb{R}^n \times \mathbb{R}_+$, and consider the Dirichlet problem

$$\begin{aligned}\Big(\frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}\Big)u = \Big(\frac{\partial^2}{\partial t^2} + \Delta\Big)u = 0 &\quad \text{in } \ \Omega \times (0,\infty)\,, \\ u(\cdot,0) = f(\cdot) &\quad \text{on } \ \Omega \times \{t=0\}\end{aligned}$$

for some $f \in \mathscr{S}(\mathbb{R}^n)$. Note that for any constant c, ct is always a solution as it is harmonic and vanishes at the boundary $t = 0$, so for uniqueness, we insist that u be bounded. This in turn means u is in $\mathscr{S}'(\mathbb{R}^n)$ and hence we may use the Fourier transform. Applying the Fourier transform (in the x variable) $\mathscr{F}_x$, we see that

$$\begin{aligned}\frac{\partial^2}{\partial t^2}(\mathscr{F}_x u)(\xi,t) = |\xi|^2(\mathscr{F}_x u)(\xi,t) &\qquad \forall\,(\xi,t) \in \mathbb{R}^n \times \mathbb{R}_+\,, \\ (\mathscr{F}_x u)(\xi,0) = \widehat{f}(\xi) &\qquad \forall\,\xi \in \mathbb{R}^n\,.\end{aligned}$$

Therefore, $(\mathscr{F}_x u)(\xi,t) = C_1(\xi)e^{t|\xi|} + C_2(\xi)e^{-t|\xi|}$, and $C_1(\xi) = 0$ by the growth condition imposed on u. Then $(\mathscr{F}_x u)(\xi,t) = \widehat{f}(\xi)e^{-t|\xi|}$ and hence using (3.5),

$$\begin{aligned}u(x,t) &= \mathscr{F}_x^*(\widehat{f}(\cdot)e^{-t|\cdot|})(x) = \big[\mathscr{F}_x^*(e^{-t|\cdot|}) * f\big](x) \\ &= \frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}}\int_{\mathbb{R}^n} \frac{t f(y)}{(t^2 + |x-y|^2)^{\frac{n+1}{2}}}\,dy\,.\end{aligned}$$

This is the *Poisson integral formula on the half-space.*

If f is bounded; that is, $f \in L^\infty(\mathbb{R}^n)$, then the integral converges and $u \in L^\infty(\mathbb{R}^n \times \mathbb{R}_+)$. Therefore, $u \in \mathscr{C}^\infty(\mathbb{R}^n \times \mathbb{R}_+) \cap L^\infty(\mathbb{R}^n \times \mathbb{R}_+)$.

3.8.2 The Heat equation

Let $t \geqslant 0$ denote time, and x denote a point in space $\mathbb{R}^{\mathrm{n}}$. The function $u(x,t)$ denotes the temperature at time t and position x, and $g \in \mathscr{S}(\mathbb{R}^{\mathrm{n}})$ denotes the initial temperature distribution. We wish to solve the *heat equation*

$$u_t(x,t) = \Delta u(x,t) \qquad \forall\, (x,t) \in \mathbb{R}^{\mathrm{n}} \times (0,\infty)\,, \tag{3.8a}$$

$$u(x,0) = g(x) \qquad \forall\, x \in \mathbb{R}^{\mathrm{n}}. \tag{3.8b}$$

Taking the Fourier transform of (3.8), we find that

$$\begin{aligned} \partial_t \widehat{u}(\xi,t) &= -|\xi|^2 \widehat{u}(\xi,t)\,, \\ \widehat{u}(\xi,0) &= \widehat{g}(\xi)\,. \end{aligned}$$

Therefore, $\widehat{u}(\xi,t) = \widehat{g}(\xi)e^{-|\xi|^2 t}$ and hence

$$\begin{aligned} u(x,t) &= \mathscr{F}^*\big(\widehat{g}(\cdot)e^{-|\cdot|^2 t}\big)(x) = \big[\mathscr{F}^*(e^{-|\cdot|^2 t}) \star g\big](x) \\ &= \frac{1}{(4\pi t)^{\mathrm{n}/2}} \int_{\mathbb{R}^{\mathrm{n}}} e^{-\frac{|x-y|^2}{4t}} g(y)\,dy \quad \big(\equiv (\mathscr{H}(\cdot,t) * g)(x)\big)\,, \end{aligned} \tag{3.9}$$

where $\mathscr{H}(x,t) = \dfrac{1}{(4\pi t)^{\mathrm{n}/2}} \exp\Big(-\dfrac{|x|^2}{4t}\Big)$ is called the *heat kernel*.

Theorem 3.29. *If $g \in L^\infty(\mathbb{R}^{\mathrm{n}})$, then the solution u to (3.8) is in $\mathscr{C}^\infty(\mathbb{R}^{\mathrm{n}} \times (0,\infty))$.*

Proof. The function $\dfrac{e^{-|x|^2/4t}}{(4\pi t)^{n/2}}$ is $\mathscr{C}^\infty(\mathbb{R}^{\mathrm{n}} \times [\alpha,\infty))$ for all $\alpha > 0$. □

Remark 3.30. The representation formula (3.9) shows that whenever g is bounded, continuous, and positive, the solution $u(x,t)$ to (3.8) is positive everywhere for $t > 0$.

The representation formula (3.9) can also be used to prove the following

Theorem 3.31. *Assume that $g \in \mathscr{C}(\mathbb{R}^{\mathrm{n}}) \cap L^\infty(\mathbb{R}^{\mathrm{n}})$. Then u defined by (3.9) is continuous at $t = 0$; that is,*

$$\lim_{(x,t)\to(x_0,0^+)} u(x,t) = g(x_0) \qquad \forall\, x_0 \in \mathbb{R}^{\mathrm{n}}\,.$$

In order to study the *Inhomogeneous heat equation*

$$u_t(x,t) - \Delta u(x,t) = f(x,t) \qquad \forall\,(x,t) \in \mathbb{R}^{\mathrm{n}} \times (0,\infty)\,, \tag{3.10a}$$
$$u(x,0) = 0 \qquad \forall\, x \in \mathbb{R}^{\mathrm{n}}\,. \tag{3.10b}$$

we introduce the parameter $s > 0$, and consider the following problem for U:

$$\begin{aligned} U_t(x,t,s) &= \Delta U(x,t,s)\,, \\ U(x,s,s) &= f(x,s)\,. \end{aligned}$$

Then by (3.9),

$$U(x,t,s) = \int_{\mathbb{R}^{\mathrm{n}}} \mathscr{H}(x-y,t-s) f(y,s)dy\,.$$

We next invoke *Duhamel's principle* to find a solution $u(x,t)$ to (3.10):

$$u(x,t) = \int_0^t U(x,t,s)ds = \int_0^t \int_{\mathbb{R}^{\mathrm{n}}} \mathscr{H}(x-y,t-s)f(y,s)dyds\,. \tag{3.11}$$

The principle of linear superposition then shows that the solution of the equations

$$\begin{aligned} u_t(x,t) - \Delta u(x,t) &= f(x,t) \qquad \forall\,(x,t) \in \mathbb{R}^{\mathrm{n}} \times (0,\infty)\,, \\ u(x,0) &= g(x) \qquad \forall\, x \in \mathbb{R}^{\mathrm{n}}\,, \end{aligned}$$

is the sum of (3.9) and (3.11):

$$\begin{aligned} u(x,t) &= \int_0^t \int_{\mathbb{R}^{\mathrm{n}}} \mathscr{H}(x-y,t-s)f(y,s)\,dyds + \int_{\mathbb{R}^{\mathrm{n}}} \mathscr{H}(x-y,t)g(y)\,dy \\ &= \big[\mathscr{H}(\cdot,t) * g\big](x) + \int_0^t \big[\mathscr{H}(\cdot,t-s) * f(\cdot,s)\big](x)\,ds\,. \end{aligned} \tag{3.12}$$

3.8.3 The Wave equation

For wave speed $c > 0$, and for $x \in \mathbb{R}^{\mathrm{n}}$, $t \in \mathbb{R}$, consider the following second-order linear wave equation:

$$\begin{aligned} u_{tt}(x,t) &= c^2 \Delta u(x,t) \qquad \forall\,(x,t) \in \mathbb{R}^{\mathrm{n}} \times (0,\infty)\,, \\ u(x,0) &= f(x) \qquad \forall\, x \in \mathbb{R}^{\mathrm{n}}, \\ u_t(x,0) &= g(x) \qquad \forall\, x \in \mathbb{R}^{\mathrm{n}}. \end{aligned}$$

Taking the Fourier transform of (3.13), we find that

$$\begin{aligned} \widehat{u}_{tt}(\xi,t) &= -c^2|\xi|^2\widehat{u}(\xi,t) && \forall\,(\xi,t)\in\mathbb{R}^{\mathrm{n}}\times(0,\infty)\,,\\ \widehat{u}(\xi,0) &= \widehat{f}(\xi) && \forall\,\xi\in\mathbb{R}^{\mathrm{n}},\\ \widehat{u}_t(\xi,0) &= \widehat{g}(\xi) && \forall\,\xi\in\mathbb{R}^{\mathrm{n}}. \end{aligned}$$

The general solution of this second-order ordinary differential equations is given by

$$\widehat{u}(\xi,t) = C_1(\xi)\cos c|\xi|t + C_2(\xi)\sin c|\xi|t\,.$$

Solving for C_1 and C_2 by using the initial conditions, we find that

$$\widehat{u}(\xi,t) = \widehat{f}(\xi)\cos c|\xi|t + \widehat{g}(\xi)\frac{\sin c|\xi|t}{c|\xi|}\,.$$

Therefore,

$$\begin{aligned} u(x,t) &= \Big[\mathscr{F}^*(\cos c|\cdot|t) \star f + \mathscr{F}^*\Big(\frac{\sin c|\cdot|t}{c|\cdot|}\Big)\star g\Big](x)\\ &= \frac{1}{c}\Big[\frac{d}{dt}\mathscr{F}^*\Big(\frac{\sin c|\cdot|t}{|\cdot|}\Big)\star f + \mathscr{F}^*\Big(\frac{\sin c|\cdot|t}{|\cdot|}\Big)\star g\Big](x)\,. \end{aligned}$$

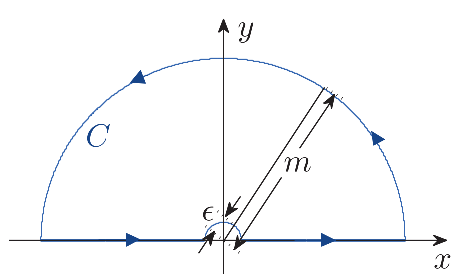
The 1-dimensional case. For the case that $\mathrm{n}=1$,

$$\int_{-m}^{m}\frac{\sin ct\lambda}{\lambda}e^{ix\lambda}d\lambda = \int_{-m}^{m}\frac{e^{i(x+ct)\lambda}-e^{i(x-ct)\lambda}}{2i\lambda}\,d\lambda\,.$$

By the Cauchy integral formula and the residue theorem,

$$\begin{aligned} \int_{-m}^{m}\frac{e^{iz}}{z}dz &= \lim_{\epsilon\searrow 0^+}\Big(\int_{-m}^{-\epsilon}+\int_{\epsilon}^{m}\Big)\frac{e^{iz}}{z}dz\\ &= \oint_C\frac{e^{iz}}{z}dz - i\int_0^{\pi}e^{ime^{i\theta}}d\theta - i\lim_{\epsilon\searrow 0^+}\int_{\pi}^{0}e^{i\epsilon e^{i\theta}}d\theta\\ &= -i\int_0^{\pi}e^{-m\sin\theta+im\cos\theta}d\theta + i\pi\,, \end{aligned}$$

where C is the contour shown below.



By the fact that $\lim\limits_{m\to\infty}\int_0^\pi e^{-m\sin\theta}d\theta = 0$ (which follows from the dominated convergence theorem), we find that

$$\lim_{m\to\infty}\int_{-m}^m \frac{e^{iz}}{z}\,dz = i\pi\,.$$

Therefore, for all $t>0$,

$$\lim_{m\to\infty}\frac{1}{\pi}\int_{-m}^m \frac{\sin ct\lambda}{\lambda}\,e^{ix\lambda}d\lambda = \chi_{\{|x|<ct\}}(x) = \begin{cases} 1 & |x|<ct, \\ 0 & |x|\geqslant ct. \end{cases}$$

Corollary 3.32. $\mathscr{F}^*\Big(\dfrac{\sin c|\cdot|t}{|\cdot|}\Big)(x) = \sqrt{\dfrac{\pi}{2}}\,\chi_{\{|x|<ct\}}(x)$ *in* $\mathscr{S}'(\mathbb{R})$.

Proof. We first note that $\Big|\displaystyle\int_{-m}^m \frac{e^{iz}}{z}dz\Big| \leqslant 2\pi$ for all $m>0$; thus

$$\Big|\int_{-m}^m \frac{\sin c|\xi|t}{|\xi|}e^{ix\xi}d\xi\Big| = \Big|\int_{-m}^m \frac{e^{i(x+ct)\xi}-e^{i(x-ct)\xi}}{2i\xi}d\xi\Big| \leqslant 2\pi \quad \forall\, m>0\,. \tag{3.14}$$

Now for all $\varphi\in\mathscr{S}(\mathbb{R})$,

$$\begin{aligned}\int_{\mathbb{R}}\mathscr{F}^*\Big(\frac{\sin c|\cdot|t}{|\cdot|}\Big)(x)\varphi(x)dx &= \int_{\mathbb{R}}\frac{\sin c|\xi|t}{|\xi|}\mathscr{F}^{-1}(\varphi)(\xi)d\xi \\ &= \lim_{m\to\infty}\frac{1}{\sqrt{2\pi}}\int_{-m}^m\int_{\mathbb{R}}\frac{\sin c|\xi|t}{|\xi|}e^{ix\xi}\varphi(x)dxd\xi\,.\end{aligned}$$

By Fubini's theorem,

$$\int_{-m}^m\int_{\mathbb{R}}\frac{\sin c|\xi|t}{|\xi|}e^{ix\xi}\varphi(x)dxd\xi = \int_{\mathbb{R}}\int_{-m}^m\frac{\sin c|\xi|t}{|\xi|}e^{ix\xi}\varphi(x)d\xi dx\,;$$

thus estimate (3.14) together with the dominated convergence theorem implies that

$$\begin{aligned}\lim_{m\to\infty}\int_{-m}^m\int_{\mathbb{R}}\frac{\sin c|\xi|t}{|\xi|}\,e^{ix\xi}\varphi(x)\,dxd\xi &= \lim_{m\to\infty}\int_{\mathbb{R}}\int_{-m}^m\frac{\sin c|\xi|t}{|\xi|}\,e^{ix\xi}\varphi(x)\,d\xi dx \\ &= \int_{\mathbb{R}}\lim_{m\to\infty}\int_{-m}^m\frac{\sin c|\xi|t}{|\xi|}\,e^{ix\xi}\varphi(x)\,d\xi dx = \pi\int_{\mathbb{R}}\chi_{\{|x|<ct\}}(x)\varphi(x)\,dx\,.\end{aligned}$$ □

We have thus established *d'Alembert's formula* for the solution of the the 1-D wave equation:

$$\begin{aligned}u(x,t) &= \frac{1}{c}\sqrt{\frac{\pi}{2}}\frac{1}{\sqrt{2\pi}}\Big[\frac{d}{dt}\int_{\mathbb{R}}f(x-y)\chi_{\{|y|<ct\}}(y)dy + \int_{\mathbb{R}}g(x-y)\chi_{\{|y|<ct\}}(y)dy\Big] \\ &= \frac{1}{2c}\frac{d}{dt}\int_{-ct}^{ct}f(x-y)dy + \frac{1}{2c}\int_{-ct}^{ct}g(x-y)dy \\ &= \frac{f(x-ct)+f(x+ct)}{2} + \frac{1}{2c}\int_{x-ct}^{x+ct}g(y)dy\,.\end{aligned}$$

The 3-dimensional case. Formally, we want to compute

$$\mathscr{F}^*\Big(\frac{\sin c|\cdot|t}{|\cdot|}\Big)(x) = \lim_{m\to\infty}\int_{|\xi|\leqslant m}\frac{\sin c|\xi|t}{|\xi|}\,e^{ix\cdot\xi}d\mu_3(\xi)\,.$$

Note that if O is a 3×3 orthonormal matrix, $|\mathrm{O}^{\mathrm{T}}\xi| = |\xi|$ for all $\xi\in\mathbb{R}^3$; thus the change of variables formula implies that

$$\begin{aligned}\mathscr{F}^*\Big(\frac{\sin c|\cdot|t}{|\cdot|}\Big)(x) &= \lim_{m\to\infty}\int_{|\xi|\leqslant m}\frac{\sin c|\mathrm{O}^{\mathrm{T}}\xi|t}{|\mathrm{O}^{\mathrm{T}}\xi|}\,e^{ix\cdot\xi}d\mu_3(\xi)\\ {\scriptstyle(\mathrm{O}^{\mathrm{T}}\xi=\eta)} &= \lim_{m\to\infty}\int_{|\xi|\leqslant m}\frac{\sin c|\eta|t}{|\eta|}\,e^{ix\cdot(\mathrm{O}\eta)}d\mu_3(\eta)\\ &= \lim_{m\to\infty}\int_{|\xi|\leqslant m}\frac{\sin c|\eta|t}{|\eta|}\,e^{i(\mathrm{O}^{\mathrm{T}}x)\cdot\eta}d\mu_3(\eta)\,.\end{aligned}$$

Now, for each $x\in\mathbb{R}^{\mathrm{n}}$, choose a 3×3 orthonormal matrix O such that $\mathrm{O}^{\mathrm{T}}x = (0,0,|x|)$. Using spherical coordinates, we obtain that

$$\begin{aligned}\mathscr{F}^{-1}\Big(\frac{\sin c|\cdot|t}{|\cdot|}\Big)(x) &= \lim_{m\to\infty}\int_{|\xi|\leqslant m}\frac{\sin c|\eta|t}{|\eta|}\,e^{i|x|\eta_3}d\mu_3(\eta)\\ &= \lim_{m\to\infty}\Big(\frac{1}{\sqrt{2\pi}}\Big)^3\int_0^m\int_0^\pi\int_0^{2\pi}\frac{\sin c\rho t}{\rho}\,e^{i|x|\rho\cos\phi}\rho^2\sin\phi\,d\theta d\phi\,d\rho\\ &= \frac{1}{\sqrt{2\pi}}\lim_{m\to\infty}\int_0^m\frac{\sin c\rho t}{-i|x|}\,e^{i|x|\rho\cos\phi}\Big|_{\phi=0}^{\phi=\pi}d\rho\\ &= \frac{1}{\sqrt{2\pi}}\lim_{m\to\infty}\int_0^m\frac{e^{-ic\rho t}-e^{ic\rho t}}{2|x|}\big(e^{i|x|\rho}-e^{-i|x|\rho}\big)d\rho\\ &= \frac{1}{2|x|}\int_{-\infty}^\infty\big(e^{i\rho(|x|-ct)}-e^{i\rho(|x|+ct)}\big)d\mu_1(\rho)\,.\end{aligned}$$

By the fact that

$$f(ct) = \int_{-\infty}^\infty\int_{-\infty}^\infty f(r)e^{i(r-ct)\rho}d\mu_1(\rho)d\mu_1(r)\,,$$

we find that for $t>0$,

$$\begin{aligned}&\Big[\mathscr{F}^{-1}\Big(\frac{\sin c|\cdot|t}{|\cdot|}\Big)\star\varphi\Big](x)\\ &\quad= \frac{1}{4\pi}\int_{\partial B(0,1)}\int_0^\infty\int_{-\infty}^\infty\frac{1}{r}\big(e^{i\rho(r-ct)}-e^{i\rho(|x|+ct)}\big)\varphi(x-r\omega)r^2d\mu_1(\rho)\,d\mu_1(r)\,dS_\omega\\ &\quad= \frac{1}{4\pi}\int_{\partial B(0,1)}\int_{-\infty}^\infty\int_{-\infty}^\infty\big(e^{i\rho(r-ct)}-e^{i\rho(|x|+ct)}\big)\varphi(x-r\omega)r\chi_{\{r>0\}}(r)d\mu_1(\rho)\,d\mu_1(r)\,dS_\omega\\ &\quad= \frac{ct}{4\pi}\int_{\partial B(0,1)}\varphi(x-ctw)\,dS_\omega = \frac{1}{4\pi ct}\int_{\partial B(x,ct)}\varphi(y)\,dS_y\,.\end{aligned}$$

Therefore,

$$u(x,t) = \frac{1}{4\pi c^2 t}\int_{\partial B(x,ct)} g(y)\, dS_y + \frac{\partial}{\partial t}\Big[\frac{1}{4\pi c^2 t}\int_{\partial B(x,ct)} f(y)\, dS_y\Big]. \tag{3.15}$$

We have just used the Fourier transform to find explicit solutions to the fundamental linear elliptic, parabolic, and hyperbolic equations. More generally, the Fourier transform is a powerful tool for the analysis of many other constant coefficient linear partial differential equations.

3.9 Exercises

PROBLEM 3.1. (a) For $f \in L^1(\mathbb{R})$, set $S_R f(x) = (2\pi)^{-\frac{1}{2}} \int_{-R}^{R} \widehat{f}(\xi) e^{ix\xi} d\xi$. Show that

$$S_R f(x) = K_R * f(x) = \int_{-\infty}^{\infty} K_R(x-y) f(y) dy$$

where

$$K_R(x) = (2\pi)^{-1}\int_{-R}^{R} e^{ix\xi} d\xi = \frac{\sin Rx}{\pi x}.$$

(b) Show that if $f \in L^2(\mathbb{R})$, then $S_R f \to f$ in $L^2(\mathbb{R})$ as $R \to \infty$.

PROBLEM 3.2. Show that for any $R \in (0,\infty)$, there exists $f \in L^1(\mathbb{R})$ such that $S_R f \notin L^1(\mathbb{R})$.
Hint: Note that $K_R \notin L^1(\mathbb{R})$.

PROBLEM 3.3. Assume $w \in \mathscr{S}'(\mathbb{R}^{\mathrm{n}}) \cap L^1_{\mathrm{loc}}(\mathbb{R}^{\mathrm{n}})$ and $w(x) \geqslant 0$. Show that if $\widehat{w} \in L^\infty(\mathbb{R}^{\mathrm{n}})$, then $w \in L^1(\mathbb{R}^{\mathrm{n}})$ and

$$\|\widehat{w}\|_{L^\infty(\mathbb{R}^{\mathrm{n}})} = (2\pi)^{-\frac{\mathrm{n}}{2}} \|w\|_{L^1(\mathbb{R}^{\mathrm{n}})}$$

Hint: Consider $w_j(x) = \psi\big(\frac{x}{j}\big) w(x)$ with $\psi \in \mathscr{C}_c^\infty(\mathbb{R}^{\mathrm{n}})$ and $\psi(0) = 1$.

PROBLEM 3.4. Consider the Poisson equation on $\mathbb{R}^1$: $u_{xx} = f$.

(a) Show that $\varphi(x) = \dfrac{x+|x|}{2}$ and $\psi(x) = \dfrac{|x|}{2}$ are both distributional solutions to $u_{xx} = \delta_0$.

(b) Let f be continuous with compact support in $\mathbb{R}$. Show that $u(x) = \int_{\mathbb{R}} \varphi(x-y)f(y)dy$ and $v(x) = \int_{\mathbb{R}} \psi(x-y)f(y)dy$ both solve the Poisson equation $w_{xx}(x) = f(x)$ (without relying upon distribution theory).

Problem 3.5. Let $T \in \mathscr{S}'(\mathbb{R}^{\rm n})$ and $f \in S(\mathbb{R}^{\rm n})$. Show that the Leibniz rule for distributional derivatives holds; that is, show that $\dfrac{\partial}{\partial x_i}(fT) = f\dfrac{\partial T}{\partial x_i} + \dfrac{\partial f}{\partial x_i}T$ in the sense of distribution.

Problem 3.6. Let $f(x) = e^{-s|x|^2}$ and $g(x) = e^{-t|x|^2}$. Find the Fourier transform of f (and g) and use the inversion formula to compute $f * g$.

Problem 3.7. Let d_r denote the map given by $d_r f(x) = f(rx)$. Show that

$$\mathscr{F}(d_r f) = r^{-\rm n} d_{1/r}\mathscr{F}(f).$$

Problem 3.8. Show that a function $f \in L^2(\mathbb{R}^{\rm n})$ is real if and only if $\widehat{f}(-\xi) = \overline{\widehat{f}(\xi)}$.

Problem 3.9. Find the Fourier transform of the function $f(x) = xe^{tx^2}$ for $t < 0$.

Problem 3.10. Find the Fourier transform of $\mathbf{1}_{(-a,a)}$, the characteristic (or indicator) function of the set $(-a, a)$.

Problem 3.11. Let $f(x) = \mathbf{1}_{(0,\infty)}(x)e^{-tx}$; that is,

$$f(x) = \begin{cases} e^{-tx} & \text{if } x > 0, \\ 0 & \text{if } x \leqslant 0. \end{cases}$$

Find the Fourier transform of f for $t > 0$.

Problem 3.12. Find the Fourier transform of the function $f(x) = x_1|x|^\alpha$, where x_1 is the first component of x and $-n-2 < \alpha < -2$.
Hint: Use the fact that for $-{\rm n} < \alpha < 0$,

$$\mathscr{F}(|x|^\alpha)(\xi) = \frac{\Gamma(\frac{{\rm n}+\alpha}{2})}{\Gamma(-\frac{\alpha}{2})} 2^{\alpha+\frac{\rm n}{2}} |\xi|^{-(\alpha+{\rm n})}$$

and $f(x) = \dfrac{1}{\alpha+2}\dfrac{\partial}{\partial x_1}|x|^{\alpha+2}$.

Problem 3.13. Let $\alpha > 0$ be given. Show that the Fourier transform of the function

$$f(x) = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-t} e^{-t|x|^2} dt$$

is positive.

Problem 3.14. Let $f \in L^1(\mathbb{R})$. Show that the anti-derivative of f can be written as the convolution of f and a function $\varphi \in L^1_{\mathrm{loc}}(\mathbb{R})$.

Problem 3.15. Let f be a continuous function with period 2π , and $\widehat{f}$ be the Fourier transform of f . Show that

$$\widehat{f}(\xi) = \sum_{n=-\infty}^{\infty} (\sqrt{2\pi} f_n) \tau_{-n}\delta$$

in the sense of distribution, where f_n is the Fourier coefficient defined by

$$f_n = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx \,.$$

Problem 3.16. Using Definition 3.15, compute the Fourier transform of the function

$$\mathrm{R}(x) = \begin{cases} x & \text{if } x \geqslant 0\,, \\ 0 & \text{otherwise}\,, \end{cases}$$

by completing the following:

(1) Let H be the Heaviside function. Show that $\widehat{H}(\xi) = \text{p.v.}\frac{1}{\sqrt{2\pi} i\xi} + C\delta(\xi)$ for some constant C , where $\text{p.v.}\frac{1}{\xi}$ is defined as

$$\Big\langle \text{p.v.}\frac{1}{\xi}, \varphi \Big\rangle = \lim_{\epsilon \to 0^+} \int_{\mathbb{R}\backslash[-\epsilon,\epsilon]} \frac{\varphi(\xi)}{\xi} d\xi = \lim_{\epsilon \to 0^+} \Big(\int_{-\infty}^{-\epsilon} + \int_{\epsilon}^{\infty} \Big) \frac{\varphi(\xi)}{\xi} d\xi \,.$$

Note that the integral above always exists as long as $\varphi \in \mathscr{S}(\mathbb{R})$.

(2) Let $S(x) = H(x) - \frac{1}{2}$. Then S is an odd function, and show that $\widehat{S}(\xi) = -\widehat{S}(-\xi)$.

(3) Use (2) to determine the constant C in (1).

(4) By the definition of Fourier transform, show that $\langle \widehat{R}, \varphi \rangle = -i \langle \widehat{H}, \varphi' \rangle$, and as a consequence

$$\widehat{R}(\xi) = i \frac{d}{d\xi} \widehat{H}(\xi) \,.$$

Problem 3.17. The Hilbert transform of a function $f : \mathbb{R} \to \mathbb{R}$ is defined (formally) by

$$(\mathscr{H} f)(x) = \frac{1}{\pi}\,\text{p.v.}\frac{1}{x} * f = \frac{1}{\pi}\lim_{\epsilon\to 0^+}\int_{|y-x|>\epsilon}\frac{f(y)}{x-y}\,dy\,.$$

1. Show that $\mathscr{F}(\mathscr{H} f)(\xi) = i\,\text{sgn}(\xi)\widehat{f}(\xi)$ for all $f \in L^2(\mathbb{R})$.

2. Show that $\mathscr{H} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is a surjective isometry, and $\mathscr{H}^2 = \text{Id}$.

Problem 3.18. By (3.15), the solution of the 3-dimensional wave equation

$$u_{tt}(x,t) = c^2\Delta u(x,t) \qquad \text{in} \quad \mathbb{R}^3 \times (0,\infty)\,, \tag{3.16a}$$

$$u(x,0) = f(x) \qquad \text{on} \quad \mathbb{R}^3 \times \{t=0\}\,, \tag{3.16b}$$

$$u_t(x,0) = g(x) \qquad \text{on} \quad \mathbb{R}^3 \times \{t=0\}\,, \tag{3.16c}$$

can be expressed as

$$u(x,t) = \frac{1}{4\pi c^2 t}\int_{\partial B(x,ct)} g(y)\,dS_y + \frac{\partial}{\partial t}\Big[\frac{1}{4\pi c^2 t}\int_{\partial B(x,ct)} f(y)\,dS_y\Big]\,.$$

Suppose that $f \in \mathscr{C}_c^2(\mathbb{R}^3)$ and $g \in \mathscr{C}_c^1(\mathbb{R}^3)$ so that they provide a solution $u \in \mathscr{C}^2(\mathbb{R}^3 \times (0,\infty))$. Show that there exists a constant $K > 0$ so that

$$|u(x,t)| \leqslant \frac{K}{t} \qquad \forall\, t > 0\,.$$

Draw the same conclusion if $f \in W^{2,1}(\mathbb{R}^3)$ and $g \in W^{1,1}(\mathbb{R}^3)$, and show that in this case

$$|u(x,t)| \leqslant \frac{C}{t}\Big[\|g\|_{W^{1,1}(\mathbb{R}^3)} + \|f\|_{W^{2,1}(\mathbb{R}^3)}\Big] \qquad \forall\, t > 0\,.$$

Hint: First rewrite (3.15) as

$$u(x,t) = \frac{1}{4\pi^2 c^2 t^2}\int_{|y-x|=ct}\Big[tg(y) + f(y) + \sum_{i=1}^{3} f_{y_i}(y)(y_i - x_i)\Big]dS_y$$

and convert the integral into an integral over $B(x,ct)$.

Problem 3.19. Let us consider the BBM equation

$$u_t + u_x + uu_x - u_{xxt} = 0 \qquad \forall\, x \in \mathbb{R}\,, t \in (0,T]\,, \tag{3.17a}$$

$$u(x,0) = g(x) \qquad \forall\, x \in \mathbb{R}\,. \tag{3.17b}$$

1. Use the Fourier transform to show that a bounded solution to (3.17) satisfies

$$u(x,t) = g(x) + \int_0^t \int_{-\infty}^{\infty} K(x-y)\Big[u(y,s) + \frac{1}{2}u^2(y,s)\Big]dyds\,, \tag{3.18}$$

where K is defined by

$$K(x) = \frac{1}{2}\,\mathrm{sgn}(x)e^{-|x|}\,.$$

2. Write (3.18) as $u = F(u)$; that is, treat the right-hand side of (3.18) as a function of u. Show that for $T > 0$ small enough, F has a fixed-point in the space of bounded continuous functions. (Hint: similar to the proof of the fundamental theorem of ODE, you can try to show that the map F is a contraction mapping if T is small enough, and then apply the contraction mapping theorem.)

Proof. (1) Take the Fourier transform of (3.17), we find that

$$(1+\xi^2)\widehat{u}_t(\xi,t) + i\xi\widehat{u}(\xi,t) + \frac{1}{2}i\xi\widehat{u^2}(\xi,t) = 0 \qquad \forall\, \xi \in \mathbb{R}\,, t \in (0,T]\,, \tag{3.19a}$$

$$\widehat{u}(\xi,0) = \widehat{g}(\xi) \qquad \forall\, \xi \in \mathbb{R}\,. \tag{3.19b}$$

If we can show that the inverse Fourier transform of $\dfrac{i\xi}{1+\xi^2}$ is $\sqrt{2\pi}K$, then (3.18) follows from taking the inverse Fourier transform of $\dfrac{(??a)}{1+\xi^2}$ and then time integrating the resulting equality over the time interval $(0,t)$. By making a change of variable $\xi x = z$, we find that

$$\int_{-\infty}^{\infty} \frac{i\xi}{1+\xi^2}e^{i\xi x}d\xi = \mathrm{sgn}(x)\int_{-\infty}^{\infty} \frac{iz}{z^2+x^2}e^{iz}dz\,.$$

Let C_R be the oriented boundary of the region $\big\{z \in \mathbb{C}\,\big|\,|z| < R, \mathrm{Im}(z) > 0\big\}$. Then by the residue theorem, for $R > |x|$,

$$\oint_{C_R} \frac{iz}{z^2+x^2}e^{iz}dz = \oint_{C_R} \frac{iz}{(z+ix)(z-ix)}e^{iz}dz$$

$$= \begin{cases} 2\pi i\dfrac{i(ix)}{ix+ix}e^{i(ix)} = -\pi e^{-x} & \text{if } x > 0\,, \\[2ex] 2\pi i\dfrac{i(-ix)}{-ix-ix}e^{i(-ix)} = -\pi e^{x} & \text{if } x < 0\,. \end{cases}$$

On the other hand,

$$\begin{aligned}
&\oint_{C_R} \frac{iz}{z^2+x^2} e^{iz} dz \\
&\quad = \int_{-R}^{R} \frac{iz}{z^2+x^2} e^{iz} dz + \int_0^\pi \frac{iRe^{i\theta}}{R^2e^{2i\theta}+x^2} e^{-R\sin\theta + iR\cos\theta} iRe^{i\theta} d\theta \\
&\quad = \int_{-R}^{R} \frac{iz}{z^2+x^2} e^{iz} dz - \int_0^\pi \frac{R^2e^{2i\theta}}{R^2e^{2i\theta}+x^2} e^{-R\sin\theta + iR\cos\theta} d\theta \,.
\end{aligned}$$

However, for any $\delta \in (0, \pi/2)$,

$$\begin{aligned}
&\int_0^\pi \frac{R^2e^{2i\theta}}{R^2e^{2i\theta}+x^2} e^{-R\sin\theta + iR\cos\theta} d\theta \\
&\quad = \Big(\int_{[0,\pi]\backslash(\delta,\pi-\delta)} + \int_\delta^{\pi-\delta} \Big) \frac{R^2e^{2i\theta}}{R^2e^{2i\theta}+x^2} e^{-R\sin\theta + iR\cos\theta} d\theta \,;
\end{aligned}$$

hence by the fact that the integral over $(\delta, \pi - \delta)$ converges to zero as $R \to \infty$,

$$\limsup_{R\to\infty} \Big| \int_0^\pi \frac{R^2e^{2i\theta}}{R^2e^{2i\theta}+x^2} e^{-R\sin\theta + iR\cos\theta} d\theta \Big| \leqslant 2\delta$$

and by the arbitrariness of δ, we eventually find that

$$\Big(\frac{i\xi}{1+\xi^2}\Big)^\vee = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{i\xi}{1+\xi^2} e^{i\xi x} d\xi = -\operatorname{sgn}(x) \sqrt{\frac{\pi}{2}}\, e^{-|x|} = \sqrt{2\pi} K(x) \,.$$

(2) Let $M = \Big\{ w \in C_b(\mathbb{R} \times [0,T]) \Big| \max\limits_{\mathbb{R}\times[0,T]} |w| \leqslant 2 \max\limits_{\mathbb{R}} |g| \Big\}$ for some T to be determined. Then M is a closed subset of $C_b(\mathbb{R} \times [0,T])$; hence M is complete. We claim that for $T > 0$ small enough, $F : M \to M$ is a contraction mapping, then as a consequence, F has a fixed point in M which provides a unique solution to the BBM equation.

Now suppose that $u \in M$. Then

$$\begin{aligned}
|F(u)| &\leqslant \max_{\mathbb{R}} |g| + \frac{1}{2} \int_0^T e^{-|x-y|} \Big[|u(y,s)| + \frac{1}{2} u^2(y,s) \Big] dyds \\
&\leqslant \max_{\mathbb{R}} |g| + T \max_{\mathbb{R}} (|g| + |g|^2) \leqslant 2 \max_{\mathbb{R}} |g|
\end{aligned}$$

if $T \leqslant \dfrac{1}{1+\max_{\mathbb{R}}|g|}$. Moreover, if $u, v \in M$,

$$\begin{aligned}
&|F(u)-F(v)| \\
&\quad \leqslant \frac{1}{2}\int_0^T e^{-|x-y|}\Big[|u-v|(y,s)+\frac{1}{2}(|u+v|)(|u-v|)(y,s)\Big]dyds \\
&\quad \leqslant \frac{T}{2}\Big[1+2\max_{\mathbb{R}}|g|\Big]\max_{\mathbb{R}\times[0,T]}|u-v| \leqslant \kappa \max_{\mathbb{R}\times[0,T]}|u-v|
\end{aligned}$$

for some constant $\kappa \in (0,1)$ if $T \leqslant \dfrac{2\kappa}{1+2\max_{\mathbb{R}}|g|}$. Therefore, $F: M \to M$ is a contraction mapping as long as $T \leqslant \dfrac{2\kappa}{1+2\max_{\mathbb{R}}|g|}$. □

Problem 3.20. In some occasions (especially in engineering applications), the Fourier transform and inverse Fourier transform of a (Schwartz) function f are defined by

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x)e^{-i2\pi x\cdot\xi}\,dx \quad \text{and} \quad \check{f}(x) = \int_{\mathbb{R}^n} f(\xi)e^{i2\pi x\cdot\xi}d\xi\,.$$

In this problem, we adopt this definition. Complete the following.

1. Show that $\check{\widehat{f}} = \widehat{\check{f}} = f$ for all $f \in \mathscr{S}(\mathbb{R}^n)$.

2. Show the Poisson summation formula

$$\sum_{k=-\infty}^{\infty}\widehat{f}(k) = \sum_{n=-\infty}^{\infty} f(n) \qquad \forall\, f \in \mathscr{S}(\mathbb{R})\,.$$

3. Suppose that $f \in \mathscr{S}(\mathbb{R})$. Show that

$$\sum_{k=-\infty}^{\infty}\widehat{f}(\xi-\frac{k}{T}) = T\sum_{n=-\infty}^{\infty} f(nT)e^{-i2\pi nT\xi}\,.$$

In particular, if $f \in \mathscr{S}(\mathbb{R})$ and $\mathrm{spt}(\widehat{f}) \subseteq \big[0,\frac{1}{T}\big]$,

$$\widehat{f}(\xi) = \sum_{n=-\infty}^{\infty} f(nT)e^{-i2\pi nT\xi} \qquad \forall\, \xi \in \big[0,\frac{1}{T}\big]\,.$$

This suggests that if $\widehat{f}$ has compact support in $\big[0,\frac{1}{T}\big]$, f can be reconstructed based on partial knowledge of f, namely $f(nT)$.

Chapter 4

The Sobolev Spaces $H^s(\mathbb{R}^n)$, $s \in \mathbb{R}$

4.1 $H^s(\mathbb{R}^n)$ via the Fourier Transform

The Fourier transform allows us to generalize the Hilbert spaces $H^k(\mathbb{R}^n)$ for $k \in \mathbb{N}$ to $H^s(\mathbb{R}^n)$ for all $s \in \mathbb{R}$, and hence study functions which possess fractional derivatives (and anti-derivatives) which are square integrable.

Definition 4.1. For any $s \in \mathbb{R}^n$, let $\langle \xi \rangle = \sqrt{1 + |\xi|^2}$, and set

$$H^s(\mathbb{R}^n) = \left\{ u \in \mathscr{S}'(\mathbb{R}^n) \,\middle|\, \langle \xi \rangle^s \hat{u} \in L^2(\mathbb{R}^n) \right\} = \left\{ u \in \mathscr{S}'(\mathbb{R}^n) \,\middle|\, \Lambda^s u \in L^2(\mathbb{R}^n) \right\},$$

where $\Lambda^s u = \mathscr{F}^* \big(\langle \cdot \rangle^s \hat{u} \big)$.

The operator Λ^s can be thought of as a "differential operator" of order s, yielding the isomorphism

$$H^s(\mathbb{R}^n) \cong \Lambda^{-s} L^2(\mathbb{R}^n).$$

Definition 4.2. The inner-product on $H^s(\mathbb{R}^n)$ is given by

$$(u, v)_{H^s(\mathbb{R}^n)} = (\Lambda^s u, \Lambda^s v)_{L^2(\mathbb{R}^n)} \quad \forall\, u, v \in H^s(\mathbb{R}^n).$$

and the norm on $H^s(\mathbb{R}^n)$ is

$$\|u\|^2_{H^s(\mathbb{R}^n)} = (u, u)_{H^s(\mathbb{R}^n)} \quad \forall\, u \in H^s(\mathbb{R}^n).$$

The completeness of $H^s(\mathbb{R}^n)$ with respect to the $\|\cdot\|_{H^s(\mathbb{R}^n)}$ is induced by the completeness of $L^2(\mathbb{R}^n)$.

Theorem 4.3. *For $s \in \mathbb{R}$, $(H^s(\mathbb{R}^n), \|\cdot\|_{H^s(\mathbb{R}^n)})$ is a Hilbert space.*

Example 4.4 ($H^1(\mathbb{R}^n)$)**.** The $H^1(\mathbb{R}^n)$-norm in Fourier representation is exactly the same as the that given by Definition 2.15:

$$
\begin{aligned}
\|u\|^2_{H^1(\mathbb{R}^n)} &= \int_{\mathbb{R}^n} \langle\xi\rangle^2 |\widehat{u}(\xi)|^2 d\xi = \int_{\mathbb{R}^n} (1+|\xi|^2)|\widehat{u}(\xi)|^2 d\xi \\
&= \int_{\mathbb{R}^n} (|u(x)|^2 + |Du(x)|^2)\, dx\,,
\end{aligned}
$$

the last equality following from the Plancheral theorem.

Example 4.5 ($H^{\frac{1}{2}}(\mathbb{R}^n)$)**.** The space $H^{\frac{1}{2}}(\mathbb{R}^n)$ can be viewed as interpolating between the decay required for $\widehat{u} \in L^2(\mathbb{R}^n)$ and $\widehat{u} \in H^1(\mathbb{R}^n)$:

$$
H^{\frac{1}{2}}(\mathbb{R}^n) = \left\{ u \in L^2(\mathbb{R}^n) \;\Big|\; \int_{\mathbb{R}^n} \sqrt{1+|\xi|^2}\,|\widehat{u}(\xi)|^2\, d\xi < \infty \right\}.
$$

Example 4.6 ($H^{-1}(\mathbb{R}^n)$)**.** The space $H^{-1}(\mathbb{R}^n)$ can be heuristically described as those distributions whose anti-derivative is in $L^2(\mathbb{R}^n)$; in terms of the Fourier representation, elements of $H^{-1}(\mathbb{R}^n)$ have Fourier "modes" that can grow linearly at infinity:

$$
H^{-1}(\mathbb{R}^n) = \left\{ u \in \mathscr{S}'(\mathbb{R}^n) \;\Big|\; \int_{\mathbb{R}^n} \frac{|\widehat{u}(\xi)|^2}{1+|\xi|^2}\, d\xi < \infty \right\}.
$$

For $T \in H^{-s}(\mathbb{R}^n)$ and $u \in H^s(\mathbb{R}^n)$, the duality pairing is given by

$$
\langle T, u\rangle = (\Lambda^{-s}T, \Lambda^s u)_{L^2(\mathbb{R}^n)}\,,
$$

from which the following result follows.

Proposition 4.7. *For all $s \in \mathbb{R}$, $H^s(\mathbb{R}^n)' = H^{-s}(\mathbb{R}^n)$.*

The ability to define fractional-order Sobolev spaces $H^s(\mathbb{R}^n)$ allows us to refine the estimates of the trace of a function which we previously stated in Theorem 2.44. That result, based on the Gauss-Green theorem, stated that the trace operator was continuous from $H^1(\mathbb{R}^n_+)$ into $L^2(\mathbb{R}^{n-1})$. In fact, the trace operator is continuous from $H^1(\mathbb{R}^n_+)$ into $H^{\frac{1}{2}}(\mathbb{R}^{n-1})$.

To demonstrate the idea, we take $n = 2$. Given a continuous function $u : \mathbb{R}^2 \to \mathbb{R}$, we define the operator

$$
\tau u = u(0, x_2)\,.
$$

The trace theorem asserts that we can extend τ to a continuous linear map from $H^1(\mathbb{R}^2)$ into $H^{\frac{1}{2}}(\mathbb{R})$ so that we only lose one-half of a derivative.

Theorem 4.8. $\tau : H^1(\mathbb{R}^2) \to H^{\frac{1}{2}}(\mathbb{R})$, *and there is a constant C such that*

$$\|\tau u\|_{H^{\frac{1}{2}}(\mathbb{R})} \leqslant C\|u\|_{H^1(\mathbb{R}^2)}\,.$$

Before we proceed with the proof, we state a very useful result.

Lemma 4.9. *Suppose that $u \in \mathscr{S}(\mathbb{R}^2)$ and define $f(x_2) = u(0, x_2)$. Then*

$$\widehat{f}(\xi_2) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}_{\xi_1}} \widehat{u}(\xi_1, \xi_2) d\xi_1\,.$$

Proof. $\widehat{f}(\xi_2) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{\mathbb{R}} \widehat{u}(\xi_1, \xi_2) d\xi_1$ if and only if

$$f(x_2) = \frac{1}{\sqrt{2\pi}} \mathscr{F}^*\Big(\int_{\mathbb{R}} \widehat{u}(\xi_1, \cdot) d\xi_1 \Big)(x_2) = \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \widehat{u}(\xi_1, \xi_2) d\xi_1 e^{ix_2\xi_2} d\xi_2\,.$$

On the other hand,

$$u(x_1, x_2) = \mathscr{F}^*(\widehat{u})(x_1, x_2) = \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \widehat{u}(\xi_1, \xi_2) e^{ix_1\xi_1 + ix_2\xi_2} d\xi_1 d\xi_2\,,$$

so that

$$u(0, x_2) = \mathscr{F}^*(\widehat{u})(0, x_2) = \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \widehat{u}(\xi_1, \xi_2) e^{ix_2\xi_2} d\xi_1 d\xi_2\,. \qquad \square$$

Proof of Theorem 4.8. Suppose that $u \in \mathscr{S}(\mathbb{R}^2)$ and set $f(x_2) = u(0, x_1)$. According to Lemma 4.9,

$$\begin{aligned}\widehat{f}(\xi_2) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{u}(\xi_1, \xi_2) d\xi_1 = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{u}(\xi_1, \xi_2) \langle \xi \rangle \langle \xi \rangle^{-1} d\xi_1 \\ &\leqslant \frac{1}{\sqrt{2\pi}} \Big(\int_{\mathbb{R}} |\widehat{u}(\xi_1, \xi_2)|^2 \langle \xi \rangle^2 d\xi_1 \Big)^{\frac{1}{2}} \Big(\int_{\mathbb{R}} \langle \xi \rangle^{-2} d\xi_1 \Big)^{\frac{1}{2}}\,,\end{aligned}$$

and hence

$$|\hat{f}(\xi_2)|^2 \leqslant C \int_{\mathbb{R}} |\widehat{u}(\xi_1, \xi_2)|^2 \langle \xi \rangle^2 d\xi_1 \int_{\mathbb{R}} \langle \xi \rangle^{-2} d\xi_1\,.$$

The key to this trace estimate is the explicit evaluation of the integral $\displaystyle\int_{\mathbb{R}} \langle \xi \rangle^{-2} d\xi_1$:

$$\int_{\mathbb{R}} \frac{1}{1 + \xi_1^2 + \xi_2^2} d\xi_1 = \frac{\tan^{-1}\Big(\frac{\xi_1}{\sqrt{1+\xi_2^2}}\Big)}{\sqrt{1 + \xi_2^2}} \Bigg|_{\xi_1 = -\infty}^{\xi_1 = +\infty} \leqslant \pi(1 + \xi_2^2)^{-\frac{1}{2}}\,. \tag{4.1}$$

It follows that $\displaystyle\int_{\mathbb{R}} (1 + \xi_2^2)^{\frac{1}{2}} |\widehat{f}(\xi_2)|^2 d\xi_2 \leqslant C \int_{\mathbb{R}} |\widehat{u}(\xi_1, \xi_2)|^2 \langle \xi \rangle^2 d\xi_1$, so that integration of this inequality over the set $\{\xi_2 \in \mathbb{R}\}$ yields the result. Using the density of $\mathscr{S}(\mathbb{R}^2)$ in $H^1(\mathbb{R}^2)$ completes the proof. $\square$

The proof of the trace theorem for general Sobolev spaces $H^s(\mathbb{R}^n)$ spaces replacing $H^1(\mathbb{R}^n)$ proceeds in a similar fashion; the only difference is that the integral $\int_{\mathbb{R}} \langle \xi \rangle^{-2} d\xi_1$ is replaced by $\int_{\mathbb{R}^{n-1}} \langle \xi \rangle^{-2s} d\xi_1 \cdots d\xi_{n-1}$, and its anti-derivative can also be computed. The result is the following general trace theorem.

Theorem 4.10 (The trace theorem for $H^s(\mathbb{R}^n)$)**.** *For $s > \frac{1}{2}$, the trace operator $\tau : H^s(\mathbb{R}^n) \to H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})$ is continuous.*

We can extend this result to open, bounded, $\mathscr{C}^\infty$-domains $\Omega \subseteq \mathbb{R}^n$.

Definition 4.11. Let $\partial\Omega$ denote a closed $\mathscr{C}^\infty$-manifold, and let $\{\omega_\ell\}_{\ell=1}^K$ denote an open covering of $\partial\Omega$, such that for each $\ell \in \{1, 2, ..., K\}$, there exist $\mathscr{C}^\infty$-class *charts* ϑ_ℓ which satisfy

$$\vartheta_\ell : B(0, r_\ell) \subseteq \mathbb{R}^{n-1} \to \omega_\ell \quad \text{is a } \mathscr{C}^\infty\text{-diffeomorphism}\,.$$

Next, for each $1 \leqslant \ell \leqslant K$, let $0 \leqslant \varphi_\ell \in \mathscr{C}_c^\infty(\mathcal{U}_\ell)$ denote a partition of unity so that $\sum_{\ell=1}^K \varphi_\ell(x) = 1$ for all $x \in \partial\Omega$. For all real $s \geqslant 0$, we define

$$H^s(\partial\Omega) = \left\{ u \in L^2(\partial\Omega) \,\middle|\, \|u\|_{H^s(\partial\Omega)} < \infty \right\},$$

where for all $u \in H^s(\partial\Omega)$,

$$\|u\|_{H^s(\partial\Omega)}^2 = \sum_{\ell=1}^K \|(\varphi_\ell u) \circ \vartheta_\ell\|_{H^s(\mathbb{R}^{n-1})}^2 \,.$$

The space $(H^s(\partial\Omega), \|\cdot\|_{H^s(\partial\Omega)})$ is a Hilbert space by virtue of the completeness of $H^s(\mathbb{R}^{n-1})$; furthermore, any system of charts for $\partial\Omega$ with subordinate partition of unity will produce an equivalent norm.

Theorem 4.12 (The trace map on $H^s(\Omega)$)**.** *For $s > \frac{1}{2}$, the trace operator $\tau : H^s(\Omega) \to H^{s-\frac{1}{2}}(\partial\Omega)$ is continuous.*

Proof. Let $\{\mathcal{U}_\ell\}_{\ell=1}^K$ denote an n-dimensional open cover of $\partial\Omega$ such that $\mathcal{U}_\ell \cap \partial\Omega = \omega_\ell$. Define charts $\vartheta_\ell : \mathcal{V}_\ell \to \mathcal{U}_\ell$, as in (2.25) but with each chart being a $\mathscr{C}^\infty$-map, such that ϑ_ℓ is equal to the restriction of ϑ_ℓ to the $(n-1)$-dimensional ball $B(0, r_\ell) \subseteq \mathbb{R}^{n-1}$. Also, choose a partition of unity $0 \leqslant \zeta_\ell \in \mathscr{C}_c^\infty(\mathcal{U}_\ell)$ subordinate to the covering $\mathcal{U}_\ell$ such that $\varphi_\ell = \zeta_\ell|_{\omega_\ell}$.

Then by Theorem 4.10, for $s > \dfrac{1}{2}$,

$$\|u\|^2_{H^{s-\frac{1}{2}}(\partial\Omega)} = \sum_{\ell=1}^{K} \|(\varphi_\ell u)\circ\vartheta_\ell\|^2_{H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})} \leqslant C\sum_{\ell=1}^{K} \|(\varphi_\ell u)\circ\vartheta_\ell\|^2_{H^s(\mathbb{R}^n)} \leqslant C\|u\|^2_{H^s(\Omega)}\,. \quad \square$$

One may then ask if the trace operator τ is onto; namely, given $f \in H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})$ for $s > \dfrac{1}{2}$, does there exist a $u \in H^s(\mathbb{R}^n)$ such that $f = \tau u$? By essentially reversing the order of the proof of Theorem 4.8, it is possible to answer this question in the affirmative. We first consider the case that $n = 2$ and $s = 1$.

Theorem 4.13. *The trace operator $\tau : H^1(\mathbb{R}^2) \to H^{\frac{1}{2}}(\mathbb{R})$ is a surjection.*

Proof. With $\xi = (\xi_1, \xi_2)$, we define (as one of many possible choices) the function u on $\mathbb{R}^2$ via its Fourier representation:

$$\widehat{u}(\xi_1,\xi_2) = K\widehat{f}(\xi_1)\frac{\langle\xi_1\rangle}{\langle\xi\rangle^2}\,,$$

for a constant $K \neq 0$ to be determined shortly. To verify that $\|u\|_{H^1(\mathbb{R}^2)} \leqslant \|f\|_{H^{\frac{1}{2}}(\mathbb{R})}$, note that

$$\begin{aligned}\int_{\mathbb{R}^2} |\widehat{u}(\xi)|^2\langle\xi\rangle^2 d\xi &= K^2\int_{-\infty}^{\infty} |\widehat{f}(\xi_1)|^2(1+\xi_1^2)\int_{-\infty}^{\infty}\frac{1}{1+\xi_1^2+\xi_2^2}d\xi_2\,d\xi_1\\ &= \pi K^2\int_{-\infty}^{\infty}|\widehat{f}(\xi_1)|^2\langle\xi_1\rangle d\xi_1 \leqslant C\|f\|^2_{H^{\frac{1}{2}}(\mathbb{R})}\,,\end{aligned}$$

where we have used the estimate (4.1) for the inequality above.

It remains to prove that $u(x_1, 0) = f(x_1)$, but by Lemma 4.9, it suffices that

$$\int_{-\infty}^{\infty}\widehat{u}(\xi_1,\xi_2)d\xi_2 = \sqrt{2\pi}\widehat{f}(\xi_1)\,.$$

Integrating $\widehat{u}$, we find that

$$\int_{-\infty}^{\infty}\widehat{u}(\xi_1,\xi_2)d\xi_2 = K\widehat{f}(\xi_1)\sqrt{1+\xi_1^2}\int_{-\infty}^{\infty}\frac{1}{1+\xi_1^2+\xi_2^2}d\xi_2 = K\pi\widehat{f}(\xi_1)$$

so setting $K = \sqrt{2\pi}/\pi$ completes the proof. $\square$

A similar construction yields the general result.

Theorem 4.14. *For $s > \dfrac{1}{2}$, the trace operator $\tau : H^s(\mathbb{R}^n) \to H^{s-\frac{1}{2}}(\mathbb{R}^{n-1})$ is a surjection.*

By using the system of charts employed for the proof of Theorem 4.12, we also have the surjectivity of the trace map on bounded domains.

Theorem 4.15. *For $s > \frac{1}{2}$, the trace operator $\tau : H^s(\Omega) \to H^{s-\frac{1}{2}}(\partial\Omega)$ is a surjection.*

The Fourier representation provides a very easy proof of a simple version of the Sobolev embedding theorem.

Theorem 4.16. *For $s > \frac{n}{2}$, if $u \in H^s(\mathbb{R}^n)$, then u is continuous and*

$$\max |u(x)| \leqslant C \|u\|_{H^s(\mathbb{R}^n)}.$$

Proof. By Theorem 3.6, $u = \mathscr{F}^* \widehat{u}$; thus according to Hölder's inequality and the Riemann-Lebesgue lemma (Theorem 3.23), it suffices to show that

$$\|\widehat{u}\|_{L^1(\mathbb{R}^n)} \leqslant C \|u\|_{H^s(\mathbb{R}^n)} \,. \tag{4.2}$$

But this follows from the Cauchy-Schwarz inequality since

$$\begin{aligned} \int_{\mathbb{R}^n} |\widehat{u}(\xi)| d\xi &= \int_{\mathbb{R}^n} |\widehat{u}(\xi)| \langle \xi \rangle^s \langle \xi \rangle^{-s} d\xi \\ &\leqslant \Big(\int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 \langle \xi \rangle^{2s} d\xi \Big)^{\frac{1}{2}} \Big(\int_{\mathbb{R}^n} \langle \xi \rangle^{-2s} d\xi \Big)^{\frac{1}{2}} \leqslant C \|u\|_{H^s(\mathbb{R}^n)} \,, \end{aligned}$$

the latter inequality holding whenever $s > \frac{n}{2}$. □

Hölder's inequality can be used to prove the following

Theorem 4.17 (Interpolation inequality)**.** *Let $0 < r < t < \infty$, and $s = \alpha r + (1-\alpha)t$ for some $\alpha \in (0,1)$. Then*

$$\|u\|_{H^s(\mathbb{R}^n)} \leqslant C \|u\|^{\alpha}_{H^r(\mathbb{R}^n)} \|u\|^{1-\alpha}_{H^t(\mathbb{R}^n)} \,. \tag{4.3}$$

Example 4.18 (Euler equation on $\mathbb{T}^2$)**.** On some time interval $[0,T]$ suppose that $u(x,t)$, $x \in \mathbb{T}^2, t \in [0,T]$, is a smooth solution of the Euler equations:

$$\begin{aligned} \partial_t u + (u \cdot D)u + Dp = 0 \qquad &\text{in} \quad \mathbb{T}^2 \times (0,T] \,, \\ \operatorname{div} u = 0 \qquad &\text{in} \quad \mathbb{T}^2 \times (0,T] \,, \end{aligned}$$

with smooth initial condition $u|_{t=0} = u_0$. Written in components, $u = (u^1, u^2)$ satisfies $u^i_t + u^i{,_j}\, j^j + p{,_i} = 0$ for $i = 1, 2$, where we are using the Einstein summation convention for summing repeated indices from 1 to 2 and where $u^i{,_j} = \dfrac{\partial u^i}{\partial x_j}$ and $p{,_i} = \dfrac{\partial p}{\partial x_i}$.

Computing the $L^2(\mathbb{T}^2)$ inner-product of the Euler equations with u yields the equality

$$\frac{1}{2}\frac{d}{dt}\int_{\mathbb{T}^2} |u(x,t)|^2\, dx + \underbrace{\int_{\mathbb{T}^2} u^i{,_j}\, u^j u^i\, dx}_{\mathcal{I}_1} + \underbrace{\int_{\mathbb{T}^2} p{,_i}\, u^i dx}_{\mathcal{I}_2} = 0\,.$$

Notice that

$$\mathcal{I}_1 = \frac{1}{2}\int_{\mathbb{T}^2} (|u|^2){,_j}\, u^j\, dx = \frac{1}{2}\int_{\mathbb{T}^2} |u|^2 \operatorname{div} u\, dx = 0\,,$$

the second equality arising from integration by parts with respect to $\partial/\partial x_j$. Integration by parts in the integral $\mathcal{I}_2$ shows that $\mathcal{I}_2 = 0$ as well, from which the conservation law $\dfrac{d}{dt}\|u(\cdot,t)\|^2_{L^2(\mathbb{T}^2)}$ follows.

To estimate the rate of change of higher-order Sobolev norms of u relies on the use of the Sobolev embedding theorem. In particular, we claim that on a short enough time interval $[0,T]$, we have the inequality

$$\frac{d}{dt}\|u(\cdot,t)\|^2_{H^3(\mathbb{T}^2)} \leqslant C\|u(\cdot,t)\|^3_{H^3(\mathbb{T}^2)} \tag{4.4}$$

from which it follows that $\|u(\cdot,t)\|^2_{H^3(\mathbb{T}^2)} \leqslant M$ for some constant $M < \infty$.

To prove (4.4), we compute the $H^3(\mathbb{T}^2)$ inner-product of the Euler equations with u:

$$\frac{1}{2}\frac{d}{dt}\|u(\cdot,t)\|^2_{H^3(\mathbb{T}^2)} + \sum_{|\alpha|\leqslant 3}\int_{\mathbb{T}^2} D^\alpha(u^i{,_j}\, u^j) D^\alpha u^i\, dx + \sum_{|\alpha|\leqslant 3}\int_{\mathbb{T}^2} D^\alpha p{,_i}\, D^\alpha u^i\, dx = 0\,.$$

The third integral vanishes by integration by parts and the fact that $D^\alpha \operatorname{div} u = 0$; thus, we focus on the nonlinearity, and in particular, on the highest-order derivatives $|\alpha| = 3$, and use D^3 to denote all third-order partial derivatives, as well as the notation l.o.t. for lower-order terms. We see that

$$\begin{aligned}\int_{\mathbb{T}^2} D^3(u^i{,_j}\, u^j) D^3 u^i\, dx &= \underbrace{\int_{\mathbb{T}^2} D^3 u^i{,_j}\; u^j\, D^3 u^i\, dx}_{\mathcal{K}_1} \\ &\quad + \underbrace{\int_{\mathbb{T}^2} u^i{,_j}\; D^3 u^j\, D^3 u^i\, dx}_{\mathcal{K}_2} + \int_{\mathbb{T}^2} \text{l.o.t.}\; dx\,.\end{aligned}$$

By definition of being lower-order terms, $\int_{\mathbb{T}^2} \text{l. o. t. } dx \leqslant C\|u\|^3_{H^3(\mathbb{T}^2)}$, so it remains to estimate the integrals $\mathcal{K}_1$ and $\mathcal{K}_2$. But the integral $\mathcal{K}_1$ vanishes by the same argument that proved $\mathcal{I}_1 = 0$. On the other hand, the integral $\mathcal{K}_2$ is estimated by Hölder's inequality:

$$|\mathcal{K}_2| \leqslant \|u^i,_j\|_{L^\infty(\mathbb{T}^2)} \, \|D^3 u^j\|_{H^3(\mathbb{T}^2)} \, \|D^3 u^i\|_{H^3(\mathbb{T}^2)} \, .$$

Thanks to the Sobolev embedding theorem, for $s = 2$ (s needs only to be greater than 1),

$$\|u^i,_j\|_{L^\infty(\mathbb{T}^2)} \leqslant C\|u^i,_j\|_{H^2(\mathbb{T}^2)} \leqslant \|u\|_{H^3(\mathbb{T}^2)} \, ,$$

from which it follows that $\mathcal{K}_2 \leqslant C\|u\|^3_{H^3(\mathbb{T}^2)}$, and this proves the claim.

Note well, that it is the Sobolev embedding theorem that requires the use of the space $H^3(\mathbb{T}^2)$ for this analysis; for example, it would not have been possible to establish the inequality (4.4) with the $H^2(\mathbb{T}^2)$ norm replacing the $H^3(\mathbb{T}^2)$ norm.

4.2 Fractional-Order Sobolev Spaces via Difference Quotient Norms

The case that $s > 0$

Lemma 4.19. *For $0 < s < 1$, $u \in H^s(\mathbb{R}^{\mathrm{n}})$ is equivalent to*

$$u \in L^2(\mathbb{R}^{\mathrm{n}}) \, , \qquad \iint_{\mathbb{R}^{\mathrm{n}} \times \mathbb{R}^{\mathrm{n}}} \frac{|u(x) - u(y)|^2}{|x-y|^{\mathrm{n}+2s}} \, dxdy < \infty \, .$$

Proof. The Fourier transform shows that for $h \in \mathbb{R}^{\mathrm{n}}$,

$$\int_{\mathbb{R}^{\mathrm{n}}} |u(x+h) - u(x)|^2 \, dx = \int_{\mathbb{R}^{\mathrm{n}}} |e^{ih\cdot\xi} - 1|^2 |\hat{u}(\xi)|^2 d\xi = \int_{\mathbb{R}^{\mathrm{n}}} \sin^2 \frac{h\cdot\xi}{2} |\hat{u}(\xi)|^2 d\xi \, .$$

It follows that

$$\begin{aligned}
\iint_{\mathbb{R}^{\mathrm{n}} \times \mathbb{R}^{\mathrm{n}}} \frac{|u(x) - u(y)|^2}{|x-y|^{\mathrm{n}+2s}} \, dxdy &= \iint_{\mathbb{R}^{\mathrm{n}} \times \mathbb{R}^{\mathrm{n}}} \frac{\sin^2 \frac{h\cdot\xi}{2}}{|h|^{\mathrm{n}+2s}} |\hat{u}(\xi)|^2 d\xi dh \\
&= \int_{\mathbb{R}^{\mathrm{n}}} |\hat{u}(\xi)|^2 \Big[\int_{\mathbb{R}^{\mathrm{n}}} \frac{\sin^2 \frac{h\cdot\xi}{2}}{|h|^{\mathrm{n}+2s}} dh \Big] d\xi \\
{\scriptstyle (\text{letting } h \,=\, 2|\xi|^{-1} z)} &= 2^{-2s} \int_{\mathbb{R}^{\mathrm{n}}} |\xi|^{2s} |\hat{u}(\xi)|^2 \Big[\int_{\mathbb{R}^{\mathrm{n}}} \frac{\sin^2(z \cdot \frac{\xi}{|\xi|})}{|z|^{\mathrm{n}+2s}} dz \Big] d\xi \, .
\end{aligned}$$

As the integral inside of the square brackets is rotationally invariant, it is independent of the direction of $\xi/|\xi|$; as such we set $\xi/|\xi| = e_1$ and let $z_1 = z \cdot e_1$ denote the first component of the vector z. It follows that

$$\iint_{\mathbb{R}^n \times \mathbb{R}^n} \frac{|u(x) - u(y)|^2}{|x-y|^{n+2s}} \, dxdy = C \int_{\mathbb{R}^n} |\xi|^{2s} |\widehat{u}(\xi)|^2 d\xi \,,$$

where $C = \displaystyle\int_{\mathbb{R}^n} \frac{\sin^2 z_1}{|z|^{n+2s}} dz < \infty$ since $0 < s < 1$. □

Corollary 4.20. *For* $0 < s < 1$,

$$\|u\|_{H^s(\mathbb{R}^n)} = \Big[\|u\|^2_{L^2(\mathbb{R}^n)} + \iint_{\mathbb{R}^n \times \mathbb{R}^n} \frac{|u(x) - u(y)|^2}{|x-y|^{n+2s}} \, dxdy \Big]^{\frac{1}{2}}$$

is an equivalent norm on $H^s(\mathbb{R}^n)$.

For real $s \geqslant 0$, $u \in H^s(\mathbb{R}^n)$ if and only if $D^\alpha u \in L^2(\mathbb{R}^n)$ for all $|\alpha| \leqslant [s]$ (where $[s]$ denotes the greatest integer that is not bigger than s), and

$$\iint_{\mathbb{R}^n \times \mathbb{R}^n} \frac{|D^\alpha u(x) - D^\alpha u(y)|^2}{|x-y|^{n+2(s-[s])}} \, dxdy < \infty$$

for all $|\alpha| = [s]$. Moreover, an equivalent norm on $H^s(\mathbb{R}^n)$ is given by

$$\begin{aligned} \|u\|_{H^s(\mathbb{R}^n)} &= \Big[\sum_{|\alpha| \leqslant [s]} \|D^\alpha u\|^2_{L^2(\Omega(\mathbb{R}^n)} + \iint_{\mathbb{R}^n \times \mathbb{R}^n} \frac{|D^\alpha u(x) - D^\alpha u(y)|^2}{|x-y|^{n+2(s-[s])}} \, dxdy \Big]^{\frac{1}{2}} \\ &= \Big[\|u\|^2_{H^{[s]}(\mathbb{R}^n)} + \iint_{\mathbb{R}^n \times \mathbb{R}^n} \frac{|D^\alpha u(x) - D^\alpha u(y)|^2}{|x-y|^{n+2(s-[s])}} \, dxdy \Big]^{\frac{1}{2}} . \end{aligned} \tag{4.5}$$

If $u \in H^k(\mathbb{R}^n)$, $k \in \mathbb{N}$, and $\varphi \in \mathscr{S}(\mathbb{R}^n)$, application of the product rule shows that $\varphi u \in H^k(\mathbb{R}^n)$. When $s \notin \mathbb{N}$, however, the product rule is not directly applicable and we must rely on other means to show that $\varphi u \in H^s(\mathbb{R}^n)$.

Lemma 4.21. *Suppose that* $u \in H^s(\mathbb{R}^n)$ *for some* $s \geqslant 0$ *and* $\varphi \in \mathscr{S}(\mathbb{R}^n)$. *Then* $\varphi u \in H^s(\mathbb{R}^n)$.

Proof. We first conisder the case that $0 \leqslant s < 1$.

By Corollary 4.19, since φu is clearly an $L^2(\mathbb{R}^n)$-function, it suffices to show that

$$\iint\limits_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|(\varphi u)(x) - (\varphi u)(y)|^2}{|x-y|^{n+2s}}\,dxdy < \infty\,.$$

Since $|(\varphi u)(x) - (\varphi u)(y)| \leqslant |\varphi(x) - \varphi(y)||u(x)| + |u(x) - u(y)||\varphi(y)|$,

$$\begin{aligned}
&\iint\limits_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|(\varphi u)(x) - (\varphi u)(y)|^2}{|x-y|^{n+2s}}\,dxdy \\
&\qquad \leqslant 2\iint\limits_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|\varphi(x)-\varphi(y)|^2|u(x)|^2 + |u(x)-u(y)|^2|\varphi(y)|^2}{|x-y|^{n+2s}}\,dxdy \\
&\qquad \leqslant \underbrace{2\iint\limits_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|\varphi(x)-\varphi(y)|^2|u(x)|^2}{|x-y|^{n+2s}}\,dxdy}_{\mathcal{I}_1} + \underbrace{2\|\varphi\|_{L^\infty(\mathbb{R}^n)}\iint\limits_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|u(x)-u(y)|^2}{|x-y|^{n+2s}}\,dxdy}_{\mathcal{I}_2}\,.
\end{aligned}$$

Since $u \in H^s(\mathbb{R}^n)$, $\mathcal{I}_2 < \infty$. On the other hand,

$$\mathcal{I}_1 = \Big[\int_{\mathbb{R}^n}\int_{|x-y|\leqslant 1} + \int_{\mathbb{R}^n}\int_{|x-y|\geqslant 1}\Big]\frac{|\varphi(x)-\varphi(y)|^2|u(x)|^2}{|x-y|^{n+2s}}\,dxdy\,.$$

For the integral over $|x-y| \leqslant 1$, since $\varphi \in \mathscr{S}(\mathbb{R}^n)$, $|\varphi(x) - \varphi(y)| \leqslant C|x-y|$ for some constant C. Therefore,

$$\begin{aligned}
&\int_{\mathbb{R}^n}\int_{|x-y|\leqslant 1} \frac{|\varphi(x)-\varphi(y)|^2|u(x)|^2}{|x-y|^{n+2s}}\,dxdy \\
&\qquad \leqslant C\int_{\mathbb{R}^n}\int_{|x-y|\leqslant 1} |x-y|^{2-n-2s}|u(x)|^2\,dxdy \\
&\qquad \leqslant C\int_{|z|\leqslant 1}|z|^{2-n-2s}dz\int_{\mathbb{R}^n}|u(x)|^2\,dx < \infty \qquad \text{if } s < 1\,.
\end{aligned}$$

For the remaining integral,

$$\begin{aligned}
&\int_{\mathbb{R}^n}\int_{|x-y|\geqslant 1} \frac{|\varphi(x)-\varphi(y)|^2|u(x)|^2}{|x-y|^{n+2s}}\,dxdy \\
&\qquad \leqslant 4\|\varphi\|^2_{L^\infty(\mathbb{R}^n)}\int_{\mathbb{R}^n}\int_{|x-y|\leqslant 1} |x-y|^{-n-2s}|u(x)|^2\,dxdy \\
&\qquad \leqslant 4\|\varphi\|^2_{L^\infty(\mathbb{R}^n)}\int_{|z|\geqslant 1}|z|^{-n-2s}dz\int_{\mathbb{R}}^{n}|u(x)|^2 < \infty \qquad \text{if } s > 0\,.
\end{aligned}$$

The general case of $s \geqslant 0$ can be proved in a similar fashion, and we leave the details to the reader. $\square$

The following theorem shows that $H^s(\mathbb{R}^{\rm n})$ is a *multiplicative-algebra*; that is, $fg \in H^s(\mathbb{R}^{\rm n})$ if $f, g \in H^s(\mathbb{R}^{\rm n})$, provided that $s > \dfrac{\rm n}{2}$.

Theorem 4.22. *Let $s > \dfrac{\rm n}{2}$ be a real number. Then there exists a generic constant $\mathrm{C}_s > 0$ such that*

$$\|uv\|_{H^s(\mathbb{R}^{\rm n})} \leqslant \mathrm{C}_s \|u\|_{H^s(\mathbb{R}^{\rm n})} \|v\|_{H^s(\mathbb{R}^{\rm n})} \qquad \forall\, u, v \in H^s(\mathbb{R}^{\rm n})\,. \tag{4.6}$$

Proof. Assume that $u, v \in H^s(\mathbb{R}^{\rm n})$. Since

$$\begin{aligned}\langle\xi\rangle^s &= (1+|\xi|^2)^{\frac{s}{2}} \leqslant (1+2|\xi-\eta|^2+2|\eta|^2)^{\frac{s}{2}} \\ &\leqslant 2^{\frac{s}{2}}\big(\langle\xi-\eta\rangle^2+\langle\eta\rangle^2\big)^{\frac{s}{2}} \leqslant C_s\big[\langle\xi-\eta\rangle^s+\langle\eta\rangle^s\big]\,,\end{aligned}$$

where C_s can be chosen as $2^{\frac{s}{2}}$ if $\mathrm{n} \leqslant 4$, or 2^{s-1} if $\mathrm{n} > 4$, by the definition of the convolution we find that

$$\begin{aligned}\langle\xi\rangle^s(\widehat{u}\star\widehat{v})(\xi) &= \int_{\mathbb{R}^{\rm n}} \langle\xi\rangle^s\big|\widehat{u}(\xi-\eta)\widehat{v}(\eta)\big|d\eta \\ &\leqslant C_s \int_{\mathbb{R}^{\rm n}} \big[\langle\xi-\eta\rangle^s+\langle\eta\rangle^s\big]\big|\widehat{u}(\xi-\eta)\widehat{v}(\eta)\big|d\eta \\ &= C_s\Big[\int_{\mathbb{R}^{\rm n}} \big|\langle\xi-\eta\rangle^s\widehat{u}(\xi-\eta)\big|\big|\widehat{v}(\eta)\big|d\eta + \int_{\mathbb{R}^{\rm n}} \big|\widehat{u}(\xi-\eta)\big|\big|\langle\eta\rangle^s\widehat{v}(\eta)\big|d\eta\Big] \\ &= C_s\Big[\big(|\widehat{u_s}|\star|\widehat{v}|\big)(\xi)+\big(|\widehat{u}|\star|\widehat{v_s}|\big)(\xi)\Big]\,,\end{aligned}$$

where $w_s(x) \equiv \displaystyle\int_{\mathbb{R}^{\rm n}} \langle\xi\rangle^s\widehat{w}(\xi)e^{ix\cdot\xi}d\mu_{\rm n}(\xi)$ is the inverse Fourier transform of $\langle\cdot\rangle^s\widehat{w}(\cdot)$. As a consequence,

$$\begin{aligned}\|uv\|_{H^s(\mathbb{R}^{\rm n})} &= \Big[\int_{\mathbb{R}^{\rm n}} \langle\xi\rangle^{2s}|\widehat{uv}(\xi)|^2 d\xi\Big]^{\frac{1}{2}} = \Big[\int_{\mathbb{R}^{\rm n}} \langle\xi\rangle^{2s}|(\widehat{u}\star\widehat{v})(\xi)|^2 d\xi\Big]^{\frac{1}{2}} \\ &= \frac{1}{\sqrt{2\pi}^{\rm n}}\Big[\int_{\mathbb{R}^{\rm n}} \langle\xi\rangle^{2s}|(\widehat{u}*\widehat{v})(\xi)|^2 d\xi\Big]^{\frac{1}{2}} \\ &\leqslant \frac{C_s}{\sqrt{2\pi}^{\rm n}}\big\||\widehat{u_s}|*|\widehat{v}|+|\widehat{u}|*\|\widehat{v_s}|\big\|_{L^2(\mathbb{R}^{\rm n})}\,,\end{aligned}$$

while Plancherel's formula and Young's inequality further imply that

$$\begin{aligned}\big\||\widehat{u_s}|*|\widehat{v}|+|\widehat{u}|*\|\widehat{v_s}|\big\|_{L^2(\mathbb{R}^{\rm n})} &\leqslant \big\||\widehat{u_s}|*|\widehat{v}|\big\|_{L^2(\mathbb{R}^{\rm n})}+\big\||\widehat{u}|*\|\widehat{v_s}|\big\|_{L^2(\mathbb{R}^{\rm n})} \\ &\leqslant \|\widehat{u_s}\|_{L^2(\mathbb{R}^{\rm n})}\|\widehat{v}\|_{L^1(\mathbb{R}^{\rm n})}+\|\widehat{u}\|_{L^1(\mathbb{R}^{\rm n})}\|\widehat{v_s}\|_{L^2(\mathbb{R}^{\rm n})} \\ &= \|u_s\|_{L^2(\mathbb{R}^{\rm n})}\|\widehat{v}\|_{L^1(\mathbb{R}^{\rm n})}+\|\widehat{u}\|_{L^1(\mathbb{R}^{\rm n})}\|v_s\|_{L^2(\mathbb{R}^{\rm n})} \\ &\leqslant 2\|\langle\cdot\rangle^{-s}\|_{L^2(\mathbb{R}^{\rm n})}\|u\|_{H^s(\mathbb{R}^{\rm n})}\|v\|_{H^s(\mathbb{R}^{\rm n})}\,,\end{aligned}$$

where (4.2) is used to conclude the last inequality. Estimate (4.6) is then established by choosing $\mathrm{C}_s = \dfrac{2C_s\|\langle\cdot\rangle^{-s}\|_{L^2(\mathbb{R}^n)}}{\sqrt{2\pi}^{\,n}}$. $\square$

4.2.1 The dual space of $H^s(\mathbb{R}^n)$ for $s > 0$

For $s > 0$, let $H^s(\mathbb{R}^n)'$ denote the dual space of $H^s(\mathbb{R}^n)$ with corresponding dual space norm (or operator norm) defined by

$$\|u\|_{H^s(\mathbb{R}^n)'} = \sup_{v \in H^s(\mathbb{R}^n)} \frac{\langle u, v\rangle}{\|v\|_{H^s(\mathbb{R}^n)}} = \sup_{\|v\|_{H^s(\mathbb{R}^n)}=1} \langle u, v\rangle .$$

Let $u \in H^s(\mathbb{R}^n)'$ be given. Since $H^s(\mathbb{R}^n)$ is a Hilbert space, the Riesz representation theorem implies that there exists a unique $w \in H^s(\mathbb{R}^n)$ satisfying

$$\langle u, v\rangle = (w, v)_{H^s(\mathbb{R}^n)} \qquad \forall\, v \in H^s(\mathbb{R}^n)\,, \tag{4.7}$$

and the operator norm of u is the same as the $H^s(\mathbb{R}^n)$-norm of w; that is,

$$\|u\|_{H^s(\mathbb{R}^n)'} = \|w\|_{H^s(\mathbb{R}^n)}\,. \tag{4.8}$$

Moreover, since $\mathscr{S}(\mathbb{R}^n) \subseteq H^s(\mathbb{R}^n)$, $u \in H^s(\mathbb{R}^n)'$ is a tempered distribution; thus the definition of the Fourier transform of a tempered distribution implies that

$$\big\langle \langle\cdot\rangle^{-s}\widehat{u}, \varphi\big\rangle = \big\langle \widehat{u}, \langle\cdot\rangle^{-s}\varphi\big\rangle = \big\langle u, \widehat{\langle\cdot\rangle^{-s}\varphi}\big\rangle = \big(w, \widehat{\langle\cdot\rangle^{-s}\varphi}\big)_{H^s(\mathbb{R}^n)} \qquad \forall\, \varphi \in \mathscr{S}(\mathbb{R}^n)\,.$$

Using $\widehat{f} = \check{\tilde{f}}$, where $^{\sim}$ denotes the reflection operator, we find that

$$\begin{aligned}\big\langle \langle\cdot\rangle^{-s}\widehat{u}, \varphi\big\rangle &= \int_{\mathbb{R}^n} \langle\xi\rangle^{2s}\widehat{w}(\xi)\overline{\langle-\xi\rangle^{-s}\varphi(-\xi)}\,d\xi = \int_{\mathbb{R}^n} \langle\xi\rangle^{s}\widehat{w}(-\xi)\bar{\varphi}(\xi)\,d\xi \\ &= \big(\widetilde{\langle\cdot\rangle^{s}\widehat{w}}, \varphi\big)_{L^2(\mathbb{R}^n)} \qquad \forall\, \varphi \in \mathscr{S}(\mathbb{R}^n)\,.\end{aligned}$$

Since $w \in H^s(\mathbb{R}^n)$, $\langle\cdot\rangle^{s}\widehat{w} \in L^2(\mathbb{R}^n)$; thus by the fact that $\mathscr{S}(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n)$, the equality above implies that $\langle\cdot\rangle^{-s}\widehat{u} \in L^2(\mathbb{R}^n)'$ and

$$\begin{aligned}\|\langle\cdot\rangle^{-s}\widehat{u}\|_{L^2(\mathbb{R}^n)} &= \sup_{\|\varphi\|_{L^2(\mathbb{R}^n)}=1} \big\langle \langle\cdot\rangle^{-s}\widehat{u}, \varphi\big\rangle = \sup_{\|\varphi\|_{L^2(\mathbb{R}^n)}=1} \big(\widetilde{\langle\cdot\rangle^{s}\widehat{w}}, \varphi\big)_{L^2(\mathbb{R}^n)} \\ &= \big\|\widetilde{\langle\cdot\rangle^{s}\widehat{w}}\big\|_{L^2(\mathbb{R}^n)} = \big\|\langle\cdot\rangle^{s}\widehat{w}\big\|_{L^2(\mathbb{R}^n)} = \|w\|_{H^s(\mathbb{R}^n)}\,.\end{aligned}$$

As a consequence, we conclude from (4.8) that

$$\|u\|_{H^s(\mathbb{R}^n)'} = \Big[\int_{\mathbb{R}^n} \langle\xi\rangle^{-2s}|\widehat{u}(\xi)|^2 d\xi\Big]^{\frac{1}{2}} = \|u\|_{H^{-s}(\mathbb{R}^n)}\,. \tag{4.9}$$

In other words, for $s > 0$ the space $H^{-s}(\mathbb{R}^n) = H^s(\mathbb{R}^n)'$.

4.3 The Interpolation Spaces[1]

Given $t > 0$, let

$$K(t, u; r, s) \equiv \inf_{\substack{u = u_r + u_s \\ u_r \in H^r(\mathbb{R}^{\mathrm{n}}), u_s \in H^s(\mathbb{R}^{\mathrm{n}})}} \Big[\|u_r\|^2_{H^r(\mathbb{R}^{\mathrm{n}})} + t^2 \|u_s\|^2_{H^s(\mathbb{R}^{\mathrm{n}})} \Big]^{1/2}. \tag{4.10}$$

For $\alpha \in (0,1)$, define the interpolation space $\big(H^r(\mathbb{R}^{\mathrm{n}}), H^s(\mathbb{R}^{\mathrm{n}})\big)_\alpha$ by

$$\big(H^r(\mathbb{R}^{\mathrm{n}}), H^s(\mathbb{R}^{\mathrm{n}})\big)_\alpha \equiv \Big\{ u \,\Big|\, \int_0^\infty t^{-1-2\alpha} K(t, u; r, s)^2 dt < \infty \Big\}$$

equipped with norm

$$\|u\|_{(H^r(\mathbb{R}^{\mathrm{n}}), H^s(\mathbb{R}^{\mathrm{n}}))_\alpha} \equiv \Big[\int_0^\infty t^{-1-2\alpha} K(t, u; r, s)^2 dt \Big]^{1/2}.$$

Our first goal in this section is to show that the space $\big(H^r(\mathbb{R}^{\mathrm{n}}), H^s(\mathbb{R}^{\mathrm{n}})\big)_\alpha$ is the same as $H^{\alpha s + (1-\alpha) r}(\mathbb{R}^{\mathrm{n}})$. To be more precise, we shall prove the following

Proposition 4.23. *Let $0 < r < s < \infty$, and $\alpha \in (0,1)$. Then*

$$\int_0^\infty t^{-1-2\alpha} K(t, u; r, s)^2 dt = C_\alpha \|u\|^2_{H^{\alpha r + (1-\alpha)s}(\mathbb{R}^{\mathrm{n}})}, \tag{4.11}$$

where $C_\alpha = \displaystyle\int_0^\infty \frac{t^{1-2\alpha}}{t^2+1} dt = \frac{\pi}{2 \sin \alpha\pi} < \infty.$

Proof. By the definition of the Sobolev space $H^r(\mathbb{R}^{\mathrm{n}})$,

$$K(t, u; r, s) = \inf_{\substack{u = u_r + u_s \\ u_r \in H^r(\mathbb{R}^{\mathrm{n}}), u_s \in H^s(\mathbb{R}^{\mathrm{n}})}} \Big[\int_{\mathbb{R}^{\mathrm{n}}} \big(|\widehat{u_r}(\xi)|^2 \langle \xi \rangle^{2r} + t^2 |\widehat{u_s}(\xi)|^2 \langle \xi \rangle^{2s} \big) d\xi \Big]^{1/2},$$

where we recall that $\langle \xi \rangle = (1 + |\xi|^2)^{1/2}$. For each $\xi \in \mathbb{R}^{\mathrm{n}}$, choose $\lambda(\xi) = r(\xi) e^{i\theta(\xi)}$ minimizing

$$\begin{aligned} &|\lambda(\xi)|^2 \langle \xi \rangle^{2r} + t^2 |\widehat{u}(\xi) - \lambda(\xi)|^2 \langle \xi \rangle^{2s} \\ &\quad = r(\xi)^2 \big(\langle \xi \rangle^{2r} + t^2 \langle \xi \rangle^{2s} \big) - 2\, t^2 \mathrm{Re}\big(\widehat{u}(\xi) e^{-i\theta(\xi)} \big) \langle \xi \rangle^{2s} r(\xi) + t^2 |\widehat{u}(\xi)|^2 \langle \xi \rangle^{2s}. \end{aligned}$$

Such an $\big(r(\xi), \theta(\xi)\big)$ must satisfy

$$r(\xi) = \frac{t^2 \langle \xi \rangle^{2s}}{\langle \xi \rangle^{2r} + t^2 \langle \xi \rangle^{2s}} \mathrm{Re}\big(\widehat{u}(\xi)\, e^{i\theta(\xi)} \big) \quad \text{and} \quad \theta(\xi) = \arg\big(\widehat{u}(\xi) \big).$$

[1]Readers should skip this section on first reading.

In other words, if u_r and u_s are given by

$$\widehat{u_r}(\xi) = \lambda(\xi) = \frac{t^2\langle\xi\rangle^{2s}e^{i\theta(\xi)}}{\langle\xi\rangle^{2r} + t^2\langle\xi\rangle^{2s}}\mathrm{Re}\big(\widehat{u}(\xi)\, e^{-i\theta(\xi)}\big) = \frac{t^2\langle\xi\rangle^{2s}}{\langle\xi\rangle^{2r} + t^2\langle\xi\rangle^{2s}}\widehat{u}(\xi)$$

and

$$\widehat{u_s}(\xi) = \widehat{u}(\xi) - \widehat{u_r}(\xi) = \frac{\langle\xi\rangle^{2r}}{\langle\xi\rangle^{2r} + t^2\langle\xi\rangle^{2s}}\widehat{u}(\xi)\,,$$

then

$$\begin{aligned} K(t,u;r,s) &= \Big[\int_{\mathbb{R}^n} \big(|\widehat{u_r}(\xi)|^2\langle\xi\rangle^{2r} + t^2|\widehat{u_s}(\xi)|^2\langle\xi\rangle^{2s}\big)d\xi\Big]^{1/2} \\ &= \Big[\int_{\mathbb{R}^n} \frac{t^2\langle\xi\rangle^{2r}}{t^2 + \langle\xi\rangle^{2(r-s)}}|\widehat{u}(\xi)|^2 d\xi\Big]^{1/2}. \end{aligned}$$

As a consequence, by Tonelli's theorem

$$\begin{aligned} \int_0^\infty t^{-1-2\alpha}K(t,u;r,s)^2 dt &= \int_{\mathbb{R}^n}\int_0^\infty \frac{t^{1-2\alpha}\langle\xi\rangle^{2r}}{t^2 + \langle\xi\rangle^{2(r-s)}}|\widehat{u}(\xi)|^2 dt d\xi \\ {\scriptstyle (t\,=\,\langle\xi\rangle^{r-s}t')} &= C_\alpha\int_{\mathbb{R}^n}\langle\xi\rangle^{2\alpha s+2(1-\alpha)r}|\widehat{u}(\xi)|^2 d\xi \\ &= C_\alpha\|u\|^2_{H^{\alpha s+(1-\alpha)r}(\mathbb{R}^n)}\,, \end{aligned}$$

where the constant C_α is given by $\displaystyle\int_0^\infty \frac{t^{1-2\alpha}}{t^2+1}\,dt$. □

Theorem 4.24. *Suppose that $0 < r_1 < s_1 < \infty$ and $0 < r_2 < s_2 < \infty$. Let $A \in \mathscr{B}(H^{s_1}(\mathbb{R}^n), H^{s_2}(\mathbb{R}^n)) \cap \mathscr{B}(H^{r_1}(\mathbb{R}^n), H^{r_2}(\mathbb{R}^n))$; that is, A is linear and satisfies*

$$\|Au\|_{H^{r_2}(\mathbb{R}^n)} \leqslant \mathrm{M}_0\|u\|_{H^{r_1}(\mathbb{R}^n)}\,, \quad \|Au\|_{H^{s_2}(\mathbb{R}^n)} \leqslant \mathrm{M}_1\|u\|_{H^{s_1}(\mathbb{R}^n)}\,.$$

Then $A \in \mathscr{B}(H^{\alpha s_1+(1-\alpha)r_1}(\mathbb{R}^n), H^{\alpha s_2+(1-\alpha)r_2}(\mathbb{R}^n))$, and

$$\|Au\|_{H^{\alpha s_2+(1-\alpha)r_2}(\mathbb{R}^n)} \leqslant \sqrt{2}\,\mathrm{M}_0^{1-\alpha}\mathrm{M}_1^{\alpha}\|u\|_{H^{\alpha s_1+(1-\alpha)r_1}(\mathbb{R}^n)}\,. \tag{4.12}$$

Proof. Let $u \in H^{\alpha s_1+(1-\alpha)r_1}(\mathbb{R}^n)$. By Proposition 4.23

$$\begin{aligned} G_1(t,u) \equiv &\inf_{\substack{u=u_s+u_r \\ u_r\in H^{r_1}(\mathbb{R}^n),u_s\in H^{s_1}(\mathbb{R}^n)}} \Big[\|u_r\|_{H^{r_1}(\mathbb{R}^n)} + t\|u_s\|_{H^{s_1}(\mathbb{R}^n)}\Big] \\ &\leqslant \sqrt{2}\,K(t,u;r_1,s_1) < \infty \end{aligned}$$

for almost all $t \in (0,\infty)$. For each decomposition $u = u_s + u_r$ with $u_s \in H^{s_1}(\mathbb{R}^n)$ and $u_r \in H^{r_1}(\mathbb{R}^n)$, we have $Au = Au_r + Au_s$; thus

$$\begin{aligned}
G_2(t, Au) &\equiv \inf_{\substack{Au=v_r+v_s\\ v_r\in H^{r_2}(\mathbb{R}^n), v_s\in H^{s_2}(\mathbb{R}^n)}} \Big[\|v_r\|_{H^{r_2}(\mathbb{R}^n)} + t\|v_s\|_{H^{s_2}(\mathbb{R}^n)}\Big]\\
&\leqslant \|Au_r\|_{H^{r_2}(\mathbb{R}^n)} + t\|Au_s\|_{H^{s_2}(\mathbb{R}^n)} \leqslant \mathrm{M}_0\|u_r\|_{H^{r_1}(\mathbb{R}^n)} + t\mathrm{M}_1\|u_s\|_{H^{s_1}(\mathbb{R}^n)}\\
&= \mathrm{M}_0\Big(\|u_s\|_{H^{s_1}(\mathbb{R}^n)} + \frac{t\mathrm{M}_1}{\mathrm{M}_0}\|u_r\|_{H^{r_1}(\mathbb{R}^n)}\Big)\,.
\end{aligned}$$

Taking the infimum over all decompositions of u, we find that

$$K(t, Au; r_2, s_2) \leqslant G_2(t, Au) \leqslant \mathrm{M}_0 G_1\Big(\frac{t\mathrm{M}_1}{\mathrm{M}_0}, u\Big) \leqslant \sqrt{2}\,\mathrm{M}_0 K\Big(\frac{t\mathrm{M}_1}{\mathrm{M}_0}, u; r_1, s_1\Big)\,.$$

Therefore, (4.11) suggests that

$$\begin{aligned}
C_\alpha\|Au\|^2_{H^{\alpha s_2+(1-\alpha)r_2}(\mathbb{R}^n)} &= \int_0^\infty t^{-1-2\alpha} K(t, Au; r_2, s_2)^2 dt\\
&\leqslant 2\int_0^\infty t^{-1-2\alpha}\mathrm{M}_0^2 K\Big(\frac{t\mathrm{M}_1}{\mathrm{M}_0}, u; r_1, s_1\Big)^2 dt\\
&= 2\int_0^\infty \mathrm{M}_1^{2\alpha}\mathrm{M}_0^{2-2\alpha}\tilde{t}^{-1-2\alpha} K(\tilde{t}, u; r_1, s_1)^2 d\tilde{t}\\
&= 2C_\alpha \mathrm{M}_0^{2-2\alpha}\mathrm{M}_1^{2\alpha}\|u\|^2_{H^{\alpha s_1+(1-\alpha)r_1}(\mathbb{R}^n)}\,.
\end{aligned}$$

□

Corollary 4.25. *Let $s > \frac{n}{2}$ be a real number. Then there exists a generic constant $\mathrm{C}_s > 0$ such that for all $0 \leqslant r \leqslant s$,*

$$\|uv\|_{H^r(\mathbb{R}^n)} \leqslant C_s\|u\|_{H^s(\mathbb{R}^n)}\|v\|_{H^r(\mathbb{R}^n)} \qquad \forall\, u \in H^s(\mathbb{R}^n) \text{ and } v \in H^r(\mathbb{R}^n)\,. \tag{4.13}$$

Proof. Let $u \in H^s(\mathbb{R}^n)$ be given, and define a linear map A by $Av = uv$. Then Theorem 4.22 implies that $A \in \mathscr{B}(H^s(\mathbb{R}^n), H^s(\mathbb{R}^n))$ with estimate

$$\|Av\|_{H^s(\mathbb{R}^n)} \leqslant \mathrm{C}_s\|u\|_{H^s(\mathbb{R}^n)}\|v\|_{H^s(\mathbb{R}^n)} \qquad \forall\, v \in H^s(\mathbb{R}^n)\,.$$

Moreover, by the Sobolev embedding (Theorem 4.16),

$$\|Av\|_{L^2(\mathbb{R}^n)} \leqslant \|u\|_{L^\infty(\mathbb{R}^n)}\|v\|_{L^2(\mathbb{R}^n)} \leqslant C\|u\|_{H^s(\mathbb{R}^n)}\|v\|_{L^2(\mathbb{R}^n)} \qquad \forall\, v \in L^2(\mathbb{R}^n)$$

which implies that $A \in \mathscr{B}(L^2(\mathbb{R}^{\mathrm{n}}), L^2(\mathbb{R}^{\mathrm{n}}))$. Therefore, Theorem 4.24 implies that $A \in \mathscr{B}(H^r(\mathbb{R}^{\mathrm{n}}), H^r(\mathbb{R}^{\mathrm{n}}))$, and for $v \in H^r(\mathbb{R}^{\mathrm{n}})$,

$$\begin{aligned}
\|Av\|_{H^r(\mathbb{R}^{\mathrm{n}})} &\leqslant \sqrt{2}\big(\mathrm{C}_s\|u\|_{H^s(\mathbb{R}^{\mathrm{n}})}\big)^{\frac{s-r}{s}}\big(C\|u\|_{H^s(\mathbb{R}^{\mathrm{n}})}\big)^{\frac{r}{s}}\|v\|_{H^r(\mathbb{R}^{\mathrm{n}})} \\
&= C_s\|u\|_{H^s(\mathbb{R}^{\mathrm{n}})}\|v\|_{H^r(\mathbb{R}^{\mathrm{n}})}\,,
\end{aligned}$$

where $C_s = \max\limits_{r\in[0,s]} \sqrt{2}C^{r/s}\mathrm{C}_s^{(s-r)/s} < \infty$. ◻

Chapter 5

Fractional-Order Sobolev Spaces on Domains with Boundary

5.1 The Space $H^s(\mathbb{R}^n_+)$

Let $\mathbb{R}^n_+ = \mathbb{R}^{n-1} \times \mathbb{R}_+$ denote the upper half space of $\mathbb{R}^n$.

The case $s = k \in \mathbb{N}$

The space $H^k(\mathbb{R}^n_+)$ is the collection of all $L^2(\mathbb{R}^n_+)$-functions so that the α-th weak derivatives belong to $L^2(\mathbb{R}^n_+)$ for all $|\alpha| \leqslant k$; that is,

$$H^k(\mathbb{R}^n_+) = \Big\{ u \in L^2(\mathbb{R}^n_+) \;\Big|\; D^\alpha u \in L^2(\mathbb{R}^n_+) \; \forall\, |\alpha| \leqslant k \Big\}$$

with norm

$$\|u\|_{H^k(\mathbb{R}^n_+)} = \Big[\sum_{|\alpha| \leqslant k} \|D^\alpha u\|^2_{L^2(\mathbb{R}^n_+)} \Big]^{\frac{1}{2}} . \tag{5.1}$$

Note that we are not able to directly use the Fourier transform to define the $H^k(\mathbb{R}^n_+)$.

Definition 5.1 (Extension operator E)**.** Fix $N \in \mathbb{N}$. Let $(a_1, \cdots, a_N)$ solve

$$\sum_{j=1}^{N} (-)^\ell 2^{(1-j)\ell} a_j = 1 , \qquad \ell = 0, \cdots, N-1 .$$

We denote by $E : \mathscr{C}(\overline{\mathbb{R}^n_+}) \to \mathscr{C}(\mathbb{R}^n)$ the function

$$(Eu)(x) = \begin{cases} u(x) & \text{if} \quad x_n \geqslant 0 , \\ \sum\limits_{j=1}^{N} a_j u(x', -2^{1-j} x_n) & \text{if} \quad x_n < 0 . \end{cases} \tag{5.2}$$

Note that the the coefficients a_j solve a linear system of N equations for N unknowns which is always solvable since the determinant never vanishes.

Theorem 5.2 (Sobolev extension theorem)**.** *For any $N \in \mathbb{N}$, the operator E defined in* (5.2) *has a continuous extension to an operator $E : H^k(\mathbb{R}^n_+) \to H^k(\mathbb{R}^n)$ for all $k \leqslant N-1$.*

Proof. We must show that all derivatives of u of order not bigger than $N-1$ are continuous at $x_n = 0$. We compute $D^\ell_{x_n} Eu$:

$$D^\ell_{x_n}(Eu)(x) = \begin{cases} D^\ell_{x_n} u(x) & \text{if} \quad x_n > 0\,, \\ \sum\limits_{j=1}^{N} (-1)^\ell 2^{(1-j)\ell} a_j (D^\ell_{x_n} u)(x', -2^{1-j} x_n) & \text{if} \quad x_n < 0\,. \end{cases}$$

By the definition of a_j, $\lim\limits_{x_1 \to 0^+} D^\ell_{x_1}(Eu)(x) = \lim\limits_{x_1 \to 0^-} D^\ell_{x_1}(Eu)(x)$. So $Eu \in H^k(\mathbb{R}^n)$. Finally, the continuity of E is concluded by the following inequality:

$$\|Eu\|_{H^k(\mathbb{R}^n)} \leqslant C \|u\|_{H^k(\mathbb{R}^n_+)}\,. \qquad \square$$

Remark 5.3. The extension operator E given by (5.2) also has the property that

$$\|Eu\|_{H^k(B)} \leqslant C \|u\|_{H^k(B^+)} \qquad \forall\, u \in \mathscr{C}^k(\overline{B^+}) \cap H^k(B^+)\,,$$

where $B \subseteq \mathbb{R}^n$ denotes a ball in $\mathbb{R}^n$, B^+ is the upper half part of B; that is, $B^+ = \big\{ y = (y_1, \cdots, y_n) \in B \,\big|\, y_n > 0 \big\}$.

Lemma 5.4. *For $k \in \mathbb{N}$, each $u \in H^k(\mathbb{R}^n_+)$ is the restriction of some $w \in H^k(\mathbb{R}^n)$ to $\mathbb{R}^n_+$, that is, $u = w|_{\mathbb{R}^n_+}$.*

Proof. We define the restriction map $\varrho : H^k(\mathbb{R}^n) \to H^k(\mathbb{R}^n_+)$. By Theorem 5.2, the restriction map is onto, since $\varrho E = \mathrm{Id}$ on $H^k(\mathbb{R}^n_+)$. $\square$

The case s is not an integer

Next, suppose that $N - 2 < s < N - 1$ for some $N \in \mathbb{N}$ given in (5.2), and let E continue to denote the Sobolev extension operator.

We define the space $H^s(\mathbb{R}^n_+)$ as the restriction of $H^s(\mathbb{R}^n)$ to $\mathbb{R}^n_+$ with norm

$$\|u\|_{H^s(\mathbb{R}^n_+)} := \|Eu\|_{H^s(\mathbb{R}^n)}\,. \tag{5.3}$$

When $s = k \in \mathbb{N}$, it may not be immediately clear that the $H^s(\mathbb{R}^{\mathrm{n}}_+)$-norm defined by (5.3) is equivalent to the $H^k(\mathbb{R}^{\mathrm{n}}_+)$-norm defined by (5.1). Let $\|\cdot\|_1$ be the norm defined by (5.1) and $\|\cdot\|_2$ be the norm defined by (5.3). It is clear that $\|u\|_1 \leqslant \|u\|_2$, and by the continuity of E, $\|u\|_2 \leqslant C\|u\|_1$; therefore, $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent if $s \in \mathbb{N}$.

For $s \notin \mathbb{N}$, motivated by Lemma 4.19 (or (4.5)), we in fact have the following

Theorem 5.5. *For $s > 0$ and $s \notin \mathbb{N}$, then $\|\cdot\|_{H^s(\mathbb{R}^{\mathrm{n}}_+)}$ is equivalent to the norm*

$$|\!|\!| u |\!|\!|_{H^s(\mathbb{R}^{\mathrm{n}}_+)} := \Big[\|u\|^2_{H^{[s]}(\mathbb{R}^{\mathrm{n}}_+)} + \sum_{|\alpha|=[s]} \iint\limits_{\mathbb{R}^{\mathrm{n}}_+\times\mathbb{R}^{\mathrm{n}}_+} \frac{|D^\alpha u(x) - D^\alpha u(y)|^2}{|x-y|^{\mathrm{n}+2(s-[s])}}\, dxdy\Big]^{\frac{1}{2}}.$$

Proof. Recall that the norm $|\!|\!| w |\!|\!|_{H^s(\mathbb{R}^{\mathrm{n}})}$ defined by

$$|\!|\!| w |\!|\!|_{H^s(\mathbb{R}^{\mathrm{n}})} := \Big[\|w\|^2_{H^{[s]}(\mathbb{R}^{\mathrm{n}})} + \sum_{|\alpha|=[s]} \iint\limits_{\mathbb{R}^{\mathrm{n}}\times\mathbb{R}^{\mathrm{n}}} \frac{|D^\alpha w(x) - D^\alpha w(y)|^2}{|x-y|^{\mathrm{n}+2(s-[s])}}\, dxdy\Big]^{\frac{1}{2}}$$

is equivalent to the norm $\|w\|_{H^s(\mathbb{R}^{\mathrm{n}})}$, so it is clear that $|\!|\!| u |\!|\!|_{H^s(\mathbb{R}^{\mathrm{n}}_+)} \leqslant C_1\|Eu\|_{H^s(\mathbb{R}^{\mathrm{n}})}$ for some constant $C_1 > 0$ since $Eu = u$ on $\mathbb{R}^{\mathrm{n}}_+$.

For the reversed inequality, since

$$\begin{aligned}
&\iint\limits_{\mathbb{R}^{\mathrm{n}}\times\mathbb{R}^{\mathrm{n}}} \frac{|D^\alpha(Eu)(x) - D^\alpha(Eu)(y)|^2}{|x-y|^{\mathrm{n}+2(s-[s])}}\, dxdy \\
&\quad = \Big(\iint\limits_{\mathbb{R}^{\mathrm{n}}_+\times\mathbb{R}^{\mathrm{n}}_+} + \iint\limits_{\mathbb{R}^{\mathrm{n}}_+\times\mathbb{R}^{\mathrm{n}}_-} + \iint\limits_{\mathbb{R}^{\mathrm{n}}_-\times\mathbb{R}^{\mathrm{n}}_+} + \iint\limits_{\mathbb{R}^{\mathrm{n}}_-\times\mathbb{R}^{\mathrm{n}}_-} \Big) \frac{|D^\alpha(Eu)(x) - D^\alpha(Eu)(y)|^2}{|x-y|^{\mathrm{n}+2(s-[s])}}\, dxdy\,,
\end{aligned}$$

by the boundedness of the extension operator we find that

$$\begin{aligned}
\|Eu\|^2_{H^s(\mathbb{R}^{\mathrm{n}})} &= \|Eu\|^2_{H^{[s]}(\mathbb{R}^{\mathrm{n}})} + \sum_{|\alpha|=[s]} \iint\limits_{\mathbb{R}^{\mathrm{n}}\times\mathbb{R}^{\mathrm{n}}} \frac{|D^\alpha(Eu)(x) - D^\alpha(Eu)(y)|^2}{|x-y|^{\mathrm{n}+2(s-[s])}}\, dxdy \\
&\leqslant C\Big[\|u\|^2_{H^s(\mathbb{R}^{\mathrm{n}}_+)} + \sum_{1\leqslant j\leqslant|\alpha|=[s]} \iint\limits_{\mathbb{R}^{\mathrm{n}}_-\times\mathbb{R}^{\mathrm{n}}_-} \frac{|(D^\alpha u)(x', -2^{1-j}x_n) - (D^\alpha u)(y', -2^{1-j}y_n)|^2}{|x-y|^{\mathrm{n}+2(s-[s])}}\, dxdy \\
&\qquad + \sum_{1\leqslant j\leqslant|\alpha|=[s]} \iint\limits_{\mathbb{R}^{\mathrm{n}}_-\times\mathbb{R}^{\mathrm{n}}_+} \frac{|(D^\alpha u)(x', -2^{1-j}x_n) - (D^\alpha u)(y', y_n)|^2}{|(x', x_{\mathrm{n}}) - (y', y_{\mathrm{n}})|^{\mathrm{n}+2(s-[s])}}\, dxdy\Big] \\
&\leqslant C\Big[\|u\|^2_{H^s(\mathbb{R}^{\mathrm{n}}_+)} + \sum_{|\alpha|=[s]} \iint\limits_{\mathbb{R}^{\mathrm{n}}_+\times\mathbb{R}^{\mathrm{n}}_+} \frac{|D^\alpha u(x) - D^\alpha u(y)|^2}{|x-y|^{\mathrm{n}+2(s-[s])}}\, dxdy\Big]
\end{aligned}$$

which implies $\|Eu\|_{H^s(\mathbb{R}^{\mathrm{n}})} \leqslant C|\!|\!| Eu |\!|\!|_{H^s(\mathbb{R}^{\mathrm{n}})} \leqslant C_2|\!|\!| u |\!|\!|_{H^s(\mathbb{R}^{\mathrm{n}}_+)}$ for some constant $C_2 > 0$; thus the equivalence of these two norms is established. □

5.2 The Sobolev Space $H^s(\Omega)$

We can now define the Sobolev spaces $H^s(\Omega)$ for any open and bounded domain $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ with smooth boundary $\partial\Omega$.

Definition 5.6 (Smoothness of the boundary)**.** We say $\partial\Omega$ is $\mathscr{C}^k$ if for each point $x_0 \in \partial\Omega$ there exist $r > 0$ and a $\mathscr{C}^k$-function $\gamma : \mathbb{R}^{\mathrm{n}-1} \to \mathbb{R}$ such that - upon relabeling and reorienting the coordinates axes if necessary - we have

$$\Omega \cap B(x_0, r) = \big\{x \in B(x_0, r) \,\big|\, x_{\mathrm{n}} > \gamma(x_1, \cdots, x_{\mathrm{n}-1})\big\}\,.$$

$\partial\Omega$ is $\mathscr{C}^\infty$ if $\partial\Omega$ is $\mathscr{C}^k$ for all $k \in \mathbb{N}$, and Ω is said to have smooth boundary if $\partial\Omega$ is $\mathscr{C}^\infty$.

Definition 5.7 (Partition of unity)**.** Let X be a topological space. A partition of unity is a collection of continuous functions $\big\{\chi_j : X \to [0,1]\big\}$ such that $\sum_j \chi_j(x) = 1$ for all $x \in X$. A partition of unity is locally finite if each x in X is contained in an open set on which only a finite number of χ_j are non-zero. A partition of unity is subordinate to an open cover $\{\mathcal{U}_j\}$ of X if each χ_j is zero on the complement of $\mathcal{U}_j$.

Proposition 5.8. *Let $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ be a bounded set, and $\{\mathcal{U}_m\}_{m=1}^K$ be an open cover of $\bar{\Omega}$. Then there exists a partition of unity $\{\zeta_m\}_{m=1}^K$ subordinate to $\{\mathcal{U}_m\}_{m=1}^K$ such that $\big\{\sqrt{\zeta_m}\big\}_{m=1}^K \subseteq \mathscr{C}_c^\infty(\mathbb{R}^{\mathrm{n}})$.*

Proof. For an open set $\mathcal{U}$ and $\delta > 0$, define $\mathcal{U}^{(\delta)}$ as the collection of interior points x of $\mathcal{U}$ such that $\operatorname{dist}(x, \partial\Omega) > \delta$. Then $\mathcal{U}^{(\delta)}$ is open. We first show that there exists $\delta > 0$ such that $\big\{\mathcal{U}_m^{(\delta)}\big\}_{m=1}^K$ is still an open cover of $\bar{\Omega}$. If not, then for each $k \in \mathbb{N}$, there exists $x_k \in \Omega$ such that $x_k \notin \bigcup\limits_{m=1}^K \mathcal{U}_m^{(1/k)}$. Since Ω is bounded, $\{x_k\}_{k=1}^\infty$ has a convergent subsequence $\big\{x_{k_j}\big\}_{j=1}^\infty$ converging to $x \in \bar{\Omega}$. This limit x cannot belong to any $\mathcal{U}_m$, a contradiction to that $\{\mathcal{U}_m\}_{m=1}^K$ is an open cover of $\bar{\Omega}$.

Now suppose that $\bar{\Omega} \subseteq \bigcup\limits_{m=1}^K \mathcal{U}_m^{(\delta)}$ for some $\delta > 0$. Let χ_m be the characteristic function of $\mathcal{U}_m^{(\delta)}$; that is,

$$\chi_m(x) = \begin{cases} 1 & \text{if } x \in \mathcal{U}_m^{(\delta)}\,, \\ 0 & \text{otherwise,} \end{cases}$$

and $\{\eta_\epsilon\}_{\epsilon>0}$ be the standard sequence of mollifiers. For $m \in \{1, \cdots, K\}$, we define

$$\xi_m = \frac{\eta_{\frac{\delta}{2}} * \chi_m}{\sum_{j=1}^K \eta_{\frac{\delta}{2}} * \chi_j} \qquad \text{and} \qquad \zeta_m = \frac{\xi_m^2}{\sum_{j=1}^K \xi_j^2}\,.$$

Then $0 \leqslant \zeta_m \leqslant 1$, $\operatorname{spt}(\zeta_m) \subseteq \mathcal{U}_m$, and $\sqrt{\zeta_m} \in \mathscr{C}_c^\infty(\mathbb{R}^n)$ for all $1 \leqslant m \leqslant K$, and $\sum\limits_{m=1}^K \xi_m = \sum\limits_{m=1}^K \zeta_m = 1$. In other words, $\{\zeta_m\}_{m=1}^K$ is a partition of unity subordinate to $\{\mathcal{U}_m\}_{m=1}^K$ satisfying that $\sqrt{\zeta_m} \in \mathscr{C}_c^\infty(\mathbb{R}^n)$ for all $1 \leqslant m \leqslant K$. $\square$

For a bounded $\mathscr{C}^k$-domain Ω, there exist $p_1, \cdots p_K \in \partial\Omega$, $r_1, \cdots r_K > 0$, $\mathscr{C}^k$-maps $\gamma_1, \cdots, \gamma_K$ such that - upon relabeling and reorienting the coordinates axes if necessary - we have

$$\Omega \cap B(p_m, r_m) = \big\{x \in B(p_m, r_m) \,\big|\, x_n > \gamma_m(x_1, \cdots, x_{n-1})\big\} \qquad \forall\, m \in \{1, \cdots, N\}\,.$$

Then for each $m \in \{1, \cdots, N\}$, the function ϑ_m defined by

$$\vartheta_m(x) = (x_1, \cdots, x_{n-1}, \gamma_m(x_1, \cdots, x_{n-1}) + x_n) \tag{5.4}$$

is a $\mathscr{C}^k$-diffeomorphism between a small neighborhood $\mathcal{V}_m$ of $\mathbb{R}^n$ and $B(p_m, r_m)$. Moreover, ϑ_m also maps $\mathcal{V}_m \cap \{x_n > 0\}$ diffeomorphically to the upper half part of $B(p_m, r_m)$. On the other hand, if such $\mathscr{C}^k$-maps $\vartheta_1, \cdots, \vartheta_K$ exist (such that the union of images of ϑ_m covers $\partial\Omega$), then Ω is of class $\mathscr{C}^k$. Therefore, Ω is a bounded $\mathscr{C}^k$-domain if and only if there exist an open cover $\{\mathcal{U}_m\}_{m=1}^K \subseteq \mathbb{R}^n$ of $\partial\Omega$ and a collection of $\mathscr{C}^k$-maps $\{\phi_m\}_{m=1}^K$ (each ϕ_m is the inverse of ϑ_m) such that for each $1 \leqslant m \leqslant K$,

$$\phi_m : \mathcal{U}_m \cap \partial\Omega \to \mathcal{V}_m \subseteq \mathbb{R}^{n-1}$$

is one-to-one, onto, and has a $\mathscr{C}^k$-inverse map for some open subset $\mathcal{V}_m$ of $\mathbb{R}^{n-1}$.

Let $\{\zeta_j\}_{j=0}^N$ be a partition of unity subordinate to the open cover $\{\mathcal{U}_j\}_{j=0}^N$ such that $\sqrt{\zeta_j} \in \mathscr{C}_c^\infty(\mathcal{U}_j)$, and define $v_0 = \zeta_0 u$ and $v_j = (\zeta_j u) \circ \vartheta_j$. Then v_0 can be treated as a function defined on $\mathbb{R}^n$, and v_j can be treated as a function defined on $\mathbb{R}^n_+$. We then have the following

Definition 5.9. The space $H^s(\Omega)$ for $s > 0$ is the collection of all measurable u such that the function $\zeta_0 u \in H^s(\mathbb{R}^n)$ and $(\zeta_j u) \circ \vartheta_j \in H^s(\mathbb{R}^n_+)$. The $H^s(\Omega)$-norm is defined by

$$\|u\|_{H^s(\Omega)} = \Big[\|\zeta_0 u\|^2_{H^s(\mathbb{R}^n)} + \sum_{j=1}^N \big\|(\zeta_j u) \circ \vartheta_j\big\|^2_{H^s(\mathbb{R}^n_+)}\Big]^{1/2}.$$

Theorem 5.10 (Extension). *Let Ω be a bounded, smooth domain. For any open set $\mathcal{U}$ such that $\Omega \subset\subset \mathcal{U}$, there exists a bounded linear operator $E : H^s(\Omega) \to H^s(\mathbb{R}^n)$ such that*

(i) *$Eu = u$ a.e. in Ω,*

(ii) *Eu has support within $\mathcal{U}$,*

(iii) *$\|\mathrm{E}u\|_{H^r(\mathbb{R}^n)} \leqslant C\|u\|_{H^r(\Omega)}$ for all $0 \leqslant r \leqslant s$, where the constant C depends only on s, Ω and $\mathcal{U}$.*

Proof. Let $\{\mathcal{U}_j\}_{j=1}^N$ be an open cover of $\partial\Omega$ such that for each $j \in \{1, \cdots, N\}$, $\mathcal{U}_j \subseteq \mathcal{U}$ and there exists a collection of smooth maps $\{\psi_j\}_{j=1}^N$ such that $\psi_j : \mathcal{U}_j \to \mathbb{R}^n$ is a diffeomorphism between a small neighborhood of $\mathbb{R}^n$. Choose $\mathcal{U}_0 \subset\subset \Omega$ so that $\{\mathcal{U}_j\}_{j=0}^K$ is an open cover of $\bar{\Omega}$, and let $\{\zeta_j\}_{j=0}^K$ be a partition of unity subordinate to $\{\mathcal{U}_j\}_{j=1}^K$ such that $\sqrt{\zeta_j} \in \mathscr{C}_c^\infty(\mathbb{R}^n)$ whose existence is guaranteed by Proposition 5.8. Define

$$\mathrm{E}u = \zeta_0 u + \sum_{j=1}^N \sqrt{\zeta_j}\left[E\big((\sqrt{\zeta_j}\,u) \circ \vartheta_j\big)\right] \circ \vartheta_j^{-1}\Big],$$

where $E : H^k(\mathbb{R}^n_+) \to H^k(\mathbb{R}^n)$ is the continuous extension defined by (5.2) for some $k \geqslant s$. Then E satisfies properties (i)-(iii), and the proofs of these three properties are left to the readers. □

Theorem 5.11 (Rellich's theorem in H^s-spaces). *Suppose that a sequence $\{u_j\}_{j=1}^\infty$ satisfies for $s \in \mathbb{R}$ and $\delta > 0$,*

$$\sup_j \|u_j\|_{H^{s+\delta}(\Omega)} \leqslant M < \infty$$

for some constant M independent of j. Then there exists a subsequence $u_{j_k} \to u$ in $H^s(\Omega)$.

Proof. Let u_j be a bounded sequence in $H^{s+\delta}(\Omega)$. We show that there exists a subsequence u_{j_k} of u_j and $u \in H^s(\Omega)$ such that $\lim\limits_{j\to\infty} \|u_{j_k} - u\|_{H^s(\Omega)} = 0$.

Let E be the extension operator defined in Theorem 5.10, and $v_j = Eu_j$, $v_j^\epsilon = \eta_\epsilon * v_j$, where η_ϵ is the standard mollifier. We first claim that $v_j^\epsilon \to v_j$ in $H^s(\mathbb{R}^n)$ uniformly in

n. In fact,

$$\begin{aligned}\|v_j^\epsilon - v_j\|^2_{H^s(\mathbb{R}^n)} &= \int_{\mathbb{R}^n} \big|\sqrt{2\pi}^{\,n}\widehat{\eta}(\epsilon\xi) - 1\big|^2 \langle\xi\rangle^{2s} \big|\widehat{v}_j(\xi)\big|^2 d\xi \\ &\leqslant \Big[\frac{4\epsilon^\delta}{(1+\epsilon)^\delta} + 4\sin^2\frac{\sqrt{\epsilon}}{2}\Big]\|v_j\|^2_{H^{s+\delta}(\mathbb{R}^n)}\end{aligned}$$

for all $\epsilon \in (0,1)$, where the inequality follows from

$$\begin{aligned}\frac{\big|\sqrt{2\pi}^{\,n}\widehat{\eta}(\epsilon\xi) - 1\big|^2}{\langle\xi\rangle^{2\delta}} &= \Big|\int_{\mathbb{R}^n} \eta(x)\frac{(e^{-ix\cdot\epsilon\xi} - 1)}{\langle\xi\rangle^\delta}\,dx\Big|^2 \\ &\leqslant \begin{cases} \dfrac{4}{(1+R^2)^\delta} & \text{if } |\xi| > R\,, \\ 4\sin^2\dfrac{\epsilon R}{2} & \text{if } |\xi| \leqslant R \leqslant \dfrac{1}{\epsilon}\Big(\leqslant \dfrac{\pi}{\epsilon}\Big)\,. \end{cases}\end{aligned}$$

Therefore, for any given $\epsilon' > 0$, there exists $\epsilon > 0$ such that

$$\|v_j^\epsilon - v_j\|_{H^s(\mathbb{R}^n)} < \frac{\epsilon'}{3} \qquad \forall\, j \in \mathbb{N}\,. \tag{5.5}$$

Now, for this particular $\epsilon > 0$, v_j^ϵ is uniformly bounded and equi-continuous since

$$|v_j^\epsilon| \leqslant \|\eta_\epsilon\|_{L^2(\mathbb{R}^n)}\|v_j^\epsilon\|_{L^2(\mathbb{R}^n)} \leqslant C_\epsilon\,, \quad |Dv_j^\epsilon| \leqslant \|D\eta_\epsilon\|_{L^2(\mathbb{R}^n)}\|v_j^\epsilon\|_{L^2(\mathbb{R}^n)} \leqslant C_\epsilon\,.$$

Therefore, by Arzela-Ascoli theorem, there exists a subsequence $v_{j_k}^\epsilon$ converges uniformly in $\mathscr{C}^0(\mathbb{R}^n)$ (or $\mathscr{C}^0(\mathcal{U})$ to be more precise since the support of v_j^ϵ can be chosen to be inside a bounded open set $\mathcal{U}$), or in particular

$$\limsup_{k,\ell\to\infty} \|v_{j_k}^\epsilon - v_{j_\ell}^\epsilon\|_{L^2(\mathcal{U})} = 0\,.$$

Moreover, by standard properties of convolution and the boundedness of E,

$$\|v_{j_k}^\epsilon - v_{j_\ell}^\epsilon\|_{H^{s+\delta}(\mathbb{R}^n)} \leqslant C\|v_{j_k} - v_{j_\ell}\|_{H^{s+\delta}(\mathbb{R}^n)} \leqslant C\|u_{j_k} - u_{j_\ell}\|_{H^{s+\delta}(\Omega)} \leqslant C\,;$$

hence interpolation inequality (4.3) implies that

$$\limsup_{k,\ell\to\infty} \|v_{j_k}^\epsilon - v_{j_\ell}^\epsilon\|_{H^s(\mathbb{R}^n)} \leqslant C \limsup_{k,\ell\to\infty} \|v_{j_k}^\epsilon - v_{j_\ell}^\epsilon\|^{\frac{\delta}{s+\delta}}_{L^2(\mathbb{R}^n)} = 0\,.$$

As a consequence, there exists $N > 0$ such that

$$\|v_{j_k}^\epsilon - v_{j_\ell}^\epsilon\|_{H^s(\mathbb{R}^n)} < \frac{\epsilon'}{3} \quad \text{whenever} \quad k,\ell \geqslant N\,. \tag{5.6}$$

The triangle inequality together with (5.5) and (5.6) then suggests that v_{j_k} is a Cauchy sequence in $H^s(\mathbb{R}^n)$; hence $v_{j_k} \to v$ in $H^s(\mathbb{R}^n)$ as $k \to \infty$. This implies $u_{j_k} \to u$ in $H^s(\Omega)$, where u is the restriction of v to Ω. $\square$

Using the extension argument, the following theorems are direct consequences of Theorem 4.17, Theorem 4.24 and Corollary 4.25. The proofs for these two theorems are left as an exercise.

Theorem 5.12 (Interpolation inequality). *Let $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ be a bounded smooth domain, $0 < r < t < \infty$, and $s = \alpha r + (1-\alpha)t$ for some $\alpha \in (0,1)$. Then*

$$\|u\|_{H^s(\Omega)} \leqslant C\|u\|^{\alpha}_{H^r(\Omega)}\|u\|^{1-\alpha}_{H^t(\Omega)}\,. \tag{5.7}$$

Theorem 5.13. *Suppose that $0 < r_1 < s_1 < \infty$ and $0 < r_2 < s_2 < \infty$. Let $A \in \mathscr{B}(H^{s_1}(\Omega), H^{s_2}(\Omega)) \cap \mathscr{B}(H^{r_1}(\Omega), H^{r_2}(\Omega))$; that is, A is linear and*

$$\|Au\|_{H^{r_2}(\Omega)} \leqslant \mathrm{M}_0\|u\|_{H^{r_1}(\Omega)}\,, \quad \|Au\|_{H^{s_2}(\Omega)} \leqslant \mathrm{M}_1\|u\|_{H^{s_1}(\Omega)}\,.$$

Then $A \in \mathscr{B}(H^{\alpha s_1+(1-\alpha)r_1}(\Omega), H^{\alpha s_2+(1-\alpha)r_2}(\Omega))$, and

$$\|Au\|_{H^{\alpha s_2+(1-\alpha)r_2}(\Omega)} \leqslant C\mathrm{M}_0^{1-\alpha}\mathrm{M}_1^{\alpha}\|u\|_{H^{\alpha s_1+(1-\alpha)r_1}(\Omega)} \tag{5.8}$$

for some generic constant $C > 0$ (independent of u).

Theorem 5.14. *Let $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ be a bounded smooth domain, and $s > \dfrac{\mathrm{n}}{2}$ be a real number. Then there exists a generic constant $\mathrm{C}_s > 0$ such that for all $0 \leqslant r \leqslant s$,*

$$\|uv\|_{H^r(\Omega)} \leqslant C_s\|u\|_{H^s(\Omega)}\|v\|_{H^r(\Omega)} \qquad \forall\, u \in H^s(\Omega) \text{ and } v \in H^r(\Omega)\,. \tag{5.9}$$

Chapter 6

The Sobolev Spaces $H^s(\mathbb{T}^{\mathrm{n}})$, $s \in \mathbb{R}$

6.1 The Fourier Series: Revisited

Definition 6.1. For $u \in L^1(\mathbb{T}^{\mathrm{n}})$, define

$$(\mathscr{F}u)(k) = \widehat{u}_k = (2\pi)^{-\frac{\mathrm{n}}{2}} \int_{\mathbb{T}^{\mathrm{n}}} e^{-ik\cdot x} u(x)\, dx\,,$$

and for $\widehat{u} \in \ell^1(\mathbb{Z}^{\mathrm{n}})$, define

$$(\mathscr{F}^*\widehat{u})(x) = (2\pi)^{-\frac{\mathrm{n}}{2}} \sum_{k\in\mathbb{Z}^{\mathrm{n}}} \widehat{u}_k e^{ik\cdot x}\,.$$

Note that $\mathscr{F} : L^1(\mathbb{T}^{\mathrm{n}}) \to \ell^\infty(\mathbb{Z}^{\mathrm{n}})$. If u is sufficiently smooth, then integration by parts yields

$$\mathscr{F}(D^\alpha u) = i^{|\alpha|} k^\alpha \widehat{u}_k, \quad k^\alpha = k_1^{\alpha_1} \cdots k_n^{\alpha_n}\,.$$

Example 6.2. Suppose that $u \in \mathscr{C}^1(\mathbb{T}^{\mathrm{n}})$. Then for $j \in \{1, ..., n\}$,

$$\begin{aligned}\mathscr{F}\Big(\frac{\partial u}{\partial x_j}\Big)(k) &= (2\pi)^{-\frac{\mathrm{n}}{2}} \int_{\mathbb{T}^{\mathrm{n}}} \frac{\partial u}{\partial x_j} e^{-ik\cdot x}\, dx \\ &= -(2\pi)^{-\mathrm{n}} \int_{\mathbb{T}^{\mathrm{n}}} u(x)\,(-ik_j)\, e^{-ik\cdot x}\, dx = ik_j \widehat{u}_k\,.\end{aligned}$$

Note that $\mathbb{T}^{\mathrm{n}}$ is a closed manifold without boundary; alternatively, one may identify $\mathbb{T}^{\mathrm{n}}$ with $[0,1]^{\mathrm{n}}$ with periodic boundary conditions; that is, with opposite faces identified.

Definition 6.3. Let $\mathfrak{s} = \mathscr{S}(\mathbb{Z}^{\mathrm{n}})$ denote the space of rapidly decreasing functions $\widehat{u}$ on $\mathbb{Z}^{\mathrm{n}}$ such that for each $N \in \mathbb{N}$,

$$p_N(u) = \sup_{k\in\mathbb{Z}^{\mathrm{n}}} \langle k\rangle^N |\widehat{u}_k| < \infty\,,$$

where $\langle k\rangle = \sqrt{1+|k|^2}$.

Then

$$\mathscr{F} : \mathscr{C}^\infty(\mathbb{T}^n) \to \mathfrak{s}\,, \quad \mathscr{F}^* : \mathfrak{s} \to \mathscr{C}^\infty(\mathbb{T}^n)\,,$$

and $\mathscr{F}^*\mathscr{F} = \mathrm{Id}$ on $\mathscr{C}^\infty(\mathbb{T}^n)$ and $\mathscr{F}\mathscr{F}^* = \mathrm{Id}$ on $\mathfrak{s}$. These properties smoothly extend to the Hilbert space setting:

$$\begin{array}{ll} \mathscr{F} : L^2(\mathbb{T}^n) \to \ell^2(\mathbb{Z}^n) & \quad \mathscr{F}^* : \ell^2(\mathbb{Z}^n) \to L^2(\mathbb{T}^n) \\ \mathscr{F}^*\mathscr{F} = \mathrm{Id} \text{ on } L^2(\mathbb{T}^n) & \quad \mathscr{F}\mathscr{F}^* = \mathrm{Id} \text{ on } \ell^2(\mathbb{Z}^n)\,. \end{array}$$

Definition 6.4. The inner-products on $L^2(\mathbb{T}^n)$ and $\ell^2(\mathbb{Z}^n)$ are

$$(u, v)_{L^2(\mathbb{T}^n)} = \int_{\mathbb{T}^n} u(x)\overline{v(x)}\, dx$$

and

$$(\widehat{u}, \widehat{v})_{\ell^2(\mathbb{Z}^n)} = \sum_{k \in \mathbb{Z}^n} \widehat{u}_k \bar{\widehat{v}}_k\,,$$

respectively.

Parseval's identity shows that $\|u\|_{L^2(\mathbb{T}^n)} = \|\widehat{u}\|_{\ell^2(\mathbb{Z}^n)}$.

Definition 6.5. We set

$$\mathscr{D}'(\mathbb{T}^n) = \mathscr{C}^\infty(\mathbb{T}^n)'\,.$$

The space $\mathscr{D}'(\mathbb{T}^n)$ is termed the space of periodic distributions.

In the same manner that we extended the Fourier transform from $\mathscr{S}(\mathbb{R}^n)$ to $\mathscr{S}'(\mathbb{R}^n)$ by duality, we may produce a similar extension to the periodic distributions:

$$\begin{array}{ll} \mathscr{F} : \mathscr{D}'(\mathbb{T}^n) \to \mathfrak{s}' & \quad \mathscr{F}^* : \mathfrak{s}' \to \mathscr{D}'(\mathbb{T}^n) \\ \mathscr{F}^*\mathscr{F} = \mathrm{Id} \text{ on } \mathscr{D}'(\mathbb{T}^n) & \quad \mathscr{F}\mathscr{F}^* = \mathrm{Id} \text{ on } \mathfrak{s}'\,. \end{array}$$

Definition 6.6 (Sobolev spaces $H^s(\mathbb{T}^n)$)**.** For all $s \in \mathbb{R}$, the Hilbert spaces $H^s(\mathbb{T}^n)$ are defined as follows:

$$H^s(\mathbb{T}^n) = \left\{ u \in \mathscr{D}'(\mathbb{T}^n) \,\middle|\, \|u\|_{H^s(\mathbb{T}^n)} < \infty \right\},$$

where the norm on $H^s(\mathbb{T}^n)$ is defined as

$$\|u\|^2_{H^s(\mathbb{T}^n)} = \sum_{k \in \mathbb{Z}^n} |\widehat{u}_k|^2 \langle k \rangle^{2s}\,.$$

The space $(H^s(\mathbb{T}^n), \|\cdot\|_{H^s(\mathbb{T}^n)})$ is a Hilbert space, and we have that

$$H^{-s}(\mathbb{T}^n) = H^s(\mathbb{T}^n)'.$$

For any $s \in \mathbb{R}$, we define the operator Λ^s as follows: for $u \in \mathscr{D}'(\mathbb{T}^n)$,

$$\Lambda^s u(x) = \sum_{k \in \mathbb{Z}^n} \widehat{u}_k \langle k \rangle^s e^{ik\cdot x}.$$

It follows that

$$H^s(\mathbb{T}^n) = \Lambda^{-s} L^2(\mathbb{T}^n),$$

and for $r, s \in \mathbb{R}$,

$$\Lambda^s : H^r(\mathbb{T}^n) \to H^{r-s}(\mathbb{T}^n) \text{ is an isomorphism}.$$

Notice then that for any $\delta > 0$,

$$\Lambda^{-\delta} : H^s(\mathbb{T}^n) \to H^s(\mathbb{T}^n) \text{ is a compact operator},$$

as it is an operator-norm limit of finite-rank operators. (In particular, the eigenvalues of $\Lambda^{-\delta}$ tend to zero in this limit.) Hence, the inclusion map $H^{s+\delta}(\mathbb{T}^n) \hookrightarrow H^s(\mathbb{T}^n)$ is compact, and we have the following

Theorem 6.7 (Rellich's theorem on $\mathbb{T}^n$). *Suppose that a sequence $\{u_j\}_{j=1}^\infty$ satisfies for $s \in \mathbb{R}$ and $\delta > 0$,*

$$\sup_j \|u_j\|_{H^{s+\delta}(\mathbb{T}^n)} \leqslant M < \infty$$

for some constant M independent of j. Then there exists a subsequence $u_{j_k} \to u$ in $H^s(\mathbb{T}^n)$.

6.2 The Poisson Integral Formula and the Laplace Operator

For $f : \mathbb{S}^1 \to \mathbb{R}$, denote by $\mathrm{PI}(f)$ the harmonic function on the unit disk $\mathrm{D} = \{x \in \mathbb{R}^2 \,|\, |x| < 1\}$ with trace f:

$$\begin{aligned} \Delta\,\mathrm{PI}(f) &= 0 \quad \text{in} \quad \mathrm{D}, \\ \mathrm{PI}(f) &= f \quad \text{on} \quad \partial\mathrm{D} = \mathbb{S}^1. \end{aligned}$$

$\mathrm{PI}(f)$ has an explicit representation via the Fourier series

$$\mathrm{PI}(f)(r,\theta) = \sum_{k\in\mathbb{Z}} \widehat{f}_k r^{|k|} e^{ik\theta} \quad \forall\, r < 1, 0 \leqslant \theta < 2\pi\,, \tag{6.1}$$

as well as the integral representation

$$\mathrm{PI}(f)(r,\theta) = \frac{1-r^2}{2\pi}\int_{\mathbb{S}^1} \frac{f(\varphi)}{r^2 - 2r\cos(\theta-\varphi)+1} d\varphi \ \ \forall\, r < 1, 0 \leqslant \theta < 2\pi\,. \tag{6.2}$$

The dominated convergence theorem shows that if $f \in \mathscr{C}^0(\mathbb{S}^1)$, then $\mathrm{PI}(f) \in \mathscr{C}^\infty(\mathrm{D}) \cap \mathscr{C}^0(\bar{\mathrm{D}})$.

Theorem 6.8. PI *extends to a continuous map from* $H^{k-\frac{1}{2}}(\mathbb{S}^1)$ *to* $H^k(\mathrm{D})$ *for all* $k \in \mathbb{N} \cup \{0\}$.

Proof. Define $u = \mathrm{PI}(f)$.
Step 1. The case that $k = 0$. Assume that $f \in H^{-\frac{1}{2}}(\Gamma)$ so that

$$\sum_{\ell\in\mathbb{Z}} |\widehat{f}_\ell|^2 \langle \ell \rangle^{-1} \leqslant M_0 < \infty\,.$$

Since the functions $\big\{e^{i\ell\theta} \,\big|\, \ell \in \mathbb{Z}\big\}$ are orthogonal with respect to the $L^2(\mathbb{S}^1)$ inner-product,

$$\begin{aligned}\|u\|^2_{L^2(\mathrm{D})} &= \int_0^1 \Big(\int_0^{2\pi} \Big|\sum_{\ell\in\mathbb{Z}} \widehat{f}_\ell r^{|\ell|} e^{i\ell\theta}\Big|^2 d\theta\Big) r dr \\ &= 2\pi \sum_{\ell\in\mathbb{Z}} |\widehat{f}_\ell|^2 \int_0^1 r^{2|\ell|+1} dr = \pi \sum_{\ell\in\mathbb{Z}} |\widehat{f}_\ell|^2 (1+|\ell|)^{-1} \leqslant \pi \|f\|^2_{H^{-\frac{1}{2}}(\mathbb{S}^1)}\,,\end{aligned}$$

where we have used the monotone convergence theorem for the first inequality.
Step 2. The case that $k = 1$. Note that in polar coordinate, the gradient operator ∇ is given by

$$\nabla = \Big(\cos\theta \frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}, \sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial\theta}\Big)\,.$$

To show that $u \in H^1(\mathrm{D})$, it suffices to show that u_r and $\dfrac{1}{r}u_\theta \in L^2(\mathrm{D})$. Since the functions $\big\{e^{i\ell\theta} \,\big|\, \ell \in \mathbb{Z}\big\}$ are orthogonal with respect to the $L^2(\mathbb{S}^1)$ inner-product,

$$\begin{aligned}\|u_r\|^2_{L^2(\mathrm{D})} &= \int_0^1 \Big(\int_0^{2\pi} \Big|\sum_{\ell\neq 0} |\ell| \widehat{f}_\ell r^{|\ell|-1} e^{i\ell\theta}\Big|^2 d\theta\Big) r dr \\ &= 2\pi \sum_{\ell\neq 0} |\ell|^2 |\widehat{f}_\ell|^2 \int_0^1 r^{2|\ell|-1} dr = \pi \sum_{\ell\neq 0} |\ell| |\widehat{f}_\ell|^2 \leqslant \pi \|f\|^2_{H^{\frac{1}{2}}(\mathbb{S}^1)}\end{aligned}$$

and similarly,

$$\begin{aligned}\Big\|\frac{1}{r}u_\theta\Big\|^2_{L^2(\mathrm{D})} &= \int_0^1 \Big(\int_0^{2\pi} \Big|\sum_{\ell\neq 0} i\ell \widehat{f}_\ell r^{|\ell|-1} e^{i\ell\theta}\Big|^2 d\theta\Big) r dr \\ &= 2\pi \sum_{\ell\neq 0} |\ell|^2 |\widehat{f}_\ell|^2 \int_0^1 r^{2|\ell|-1} dr \leqslant \pi \|f\|^2_{H^{\frac{1}{2}}(\mathbb{S}^1)}.\end{aligned}$$

Therefore, combining the estimate from **Step 1**,

$$\|u\|_{H^1(\mathrm{D})} \leqslant C\big[\|u\|_{L^2(\mathrm{D})} + \|u_r\|_{L^2(\mathrm{D})} + \Big\|\frac{1}{r}u_\theta\Big\|_{L^2(\mathrm{D})}\big] \leqslant C\|f\|_{H^{\frac{1}{2}}(\mathbb{S}^1)}.$$

Step 3. The case that $k \geqslant 2$. For general $k \geqslant 2$, we need to show that $\partial_r^k u$ and $\frac{1}{r^k}\partial_\theta^j u \in L^2(\mathrm{D})$ for all $j \in \{1, 2, \cdots, k\}$. To see this, first we note that by the Parseval identity (8.1) and the fact that $\dfrac{1}{|\ell|-k+1} \leqslant \dfrac{k+1}{|\ell|}$ for all $|\ell| \geqslant k$,

$$\begin{aligned}\|\partial_r^k u\|^2_{L^2(\mathrm{D})} &= \int_0^1 \Big(\int_0^{2\pi} \Big|\sum_{|\ell|\geqslant k} |\ell|(|\ell|-1)\cdots(|\ell|-k+1)\widehat{f}_\ell r^{|\ell|-k} e^{i\ell\theta}\Big|^2 d\theta\Big) r dr \\ &= \int_0^1 \sum_{|\ell|\geqslant k} |\ell|^2(|\ell|-1)^2\cdots(|\ell|-k+1)^2 |\widehat{f}_\ell|^2 r^{2|\ell|-2k+1} dr \\ &\leqslant 2\pi \sum_{|\ell|\geqslant k} |\ell|^{2k} |\widehat{f}_\ell|^2 \int_0^1 r^{2|\ell|-2k+1} dr \\ &= \sum_{|\ell|\geqslant k} \frac{|\ell|^{2k}}{|\ell|-k+1} |\widehat{f}_\ell|^2 \\ &\leqslant (k+1)\pi \sum_{|\ell|\geqslant k} |\ell|^{2k-1} |\widehat{f}_\ell|^2 \leqslant (k+1)\pi \|f\|^2_{H^{k-\frac{1}{2}}(\mathbb{S}^1)}.\end{aligned} \tag{6.3}$$

Moreover, since

$$\begin{aligned}\Big\|\frac{1}{r^k}\partial_\theta^j u\Big\|^2_{L^2(\mathrm{D})} &= \int_0^1 \Big(\int_0^{2\pi} \Big|\sum_{\ell\neq 0} (i\ell)^j \widehat{f}_\ell r^{|\ell|-k} e^{i\ell\theta}\Big|^2 d\theta\Big) r dr \\ &= 2\pi \sum_{\ell\neq 0} |\ell|^{2j} |\widehat{f}_\ell|^2 \int_0^1 r^{2|\ell|-2k+1} dr,\end{aligned}$$

it suffices to consider the case $j = k$. Nevertheless,

$$\begin{aligned}
\Big\|\frac{1}{r^k}\partial_\theta^k u\Big\|_{L^2(\mathrm{D})}^2 &= 2\pi \sum_{|\ell|\geqslant 2} |\ell|^{2(k-1)}(|\ell|-1)^2|\widehat{f}_\ell|^2 \int_0^1 r^{2|\ell|-2k+1}dr \\
&\quad + 2\pi \sum_{\ell\neq 0} |\ell|^{2(k-1)}(2|\ell|-1)|\widehat{f}_\ell|^2 \int_0^1 r^{2|\ell|-2k+1}dr \\
&\leqslant 2\pi \sum_{|\ell|\geqslant 2} |\ell|^{2(k-1)}(|\ell|-1)^2|\widehat{f}_\ell|^2 \int_0^1 r^{2|\ell|-2k+1}dr + 2\pi\|f\|_{H^{k-\frac{1}{2}}(\mathbb{S}^1)}^2 \\
&\leqslant 2\pi \sum_{|\ell|\geqslant 3} |\ell|^{2(k-2)}(|\ell|-1)^2(|\ell|-2)^2|\widehat{f}_\ell|^2 \int_0^1 r^{2|\ell|-2k+1}dr \\
&\quad + 2\pi(1+2)\|f\|_{H^{k-\frac{1}{2}}(\mathbb{S}^1)}^2 \\
&\leqslant \cdots\cdots \\
&\leqslant 2\pi \sum_{|\ell|\geqslant k} |\ell|^2(|\ell|-1)^2\cdots(|\ell|-k+1)^2|\widehat{f}_\ell|^2 \int_0^1 r^{2|\ell|-2k+1}dr \\
&\quad + 2\pi(1+2+\cdots+k)\|f\|_{H^{k-\frac{1}{2}}(\mathbb{S}^1)}^2
\end{aligned}$$

which, with the help of (6.3), implies that

$$\Big\|\frac{1}{r^k}\partial_\theta^k u\Big\|_{L^2(\mathrm{D})}^2 \leqslant \pi(k+1)^2\|f\|_{H^{k-\frac{1}{2}}(\mathbb{S}^1)}^2 .$$

As a consequence, we conclude that $\|u\|_{H^k(\mathrm{D})} \leqslant C_k\|f\|_{H^{k-\frac{1}{2}}(\mathbb{S}^1)}$ for some constant $C_k > 0$. ◻

The Hölder spaces on $\bar{\mathrm{D}}$ are defined as follows: if $u : \mathrm{D} \to \mathbb{R}$ is bounded and continuous, we write

$$\|u\|_{\mathscr{C}(\bar{\mathrm{D}})} := \sup_{x\in\mathrm{D}} |u(x)| .$$

For $0 < \alpha \leqslant 1$, the α^{th}-Hölder seminorm of u is

$$[u]_{\mathscr{C}^{0,\alpha}(\bar{\mathrm{D}})} := \sup_{x,y\in\mathrm{D},x\neq y} \frac{|u(x)-u(y)|}{|x-y|^\alpha}$$

and the α^{th}-Hölder norm of u is

$$\|u\|_{\mathscr{C}^{0,\alpha}(\bar{\mathrm{D}})} = \|u\|_{\mathscr{C}(\bar{\mathrm{D}})} + [u]_{\mathscr{C}^{0,\alpha}(\bar{\mathrm{D}})} .$$

We will show that if $f \in H^{3/2}(\mathbb{S}^1)$, then for $0 < \alpha < 1$, $f \in \mathscr{C}^{0,\alpha}(\mathbb{S}^1)$. Next, we will use the result of Theorem 6.8, together with Morrey's inequality and Theorem 2.36 to prove that $u \in \mathscr{C}^{0,\alpha}(\bar{\mathrm{D}})$. Let us explain this. We first prove the following:

$$\boxed{\begin{array}{l} f \in H^{3/2}(\mathbb{S}^1) \text{ implies that } f \in H^{1/2+\alpha}(\mathbb{S}^1) \text{ for } \alpha \in (0,1) \\ \text{which further implies that } f \in \mathscr{C}^{0,\alpha}(\mathbb{S}^1)\,, \end{array}}$$

where the last assertion means that $|f(x+y) - f(x)| \leqslant C|y|^\alpha$.

We start with the identity

$$\begin{aligned} |f(x+y) - f(y)| &= \Big| \sum_{k\in\mathbb{Z}} \widehat{f}_k e^{ikx}(e^{iky}-1) \Big| = \Big| \sum_{k\neq 0} \widehat{f}_k e^{ikx}(e^{iky}-1) \Big| \\ &\leqslant \Big(\sum_{k\neq 0} |\widehat{f}_k|^2 \langle k \rangle^{1+2\alpha} \Big)^{\frac12} \Big(\sum_{k\neq 0} |e^{iky}-1|^2 \langle k\rangle^{-1-2\alpha} \Big)^{\frac12} \\ &= \|f\|_{H^{1/2+\alpha}(\mathbb{S}^1)} \Big(\sum_{k\neq 0} |e^{iky}-1|^2 \langle k\rangle^{-1-2\alpha} \Big)^{\frac12}. \end{aligned}$$

We consider $|y| \leqslant \dfrac{1}{2}$ and break the sum into two parts:

$$\begin{aligned} &\sum_{k\neq 0} |e^{iky}-1|^2 \langle k\rangle^{-1-2\alpha} \\ &\qquad = \sum_{0<|k|\leqslant \frac{1}{|y|}} |e^{iky}-1|^2 \langle k\rangle^{-1-2\alpha} + \sum_{|k|\geqslant \frac{1}{|y|}+1} |e^{iky}-1|^2 \langle k\rangle^{-1-2\alpha}. \end{aligned}$$

For the second sum, we use that $|e^{iky}-1|^2 \leqslant 4$ and employ the integral test to see that

$$\sum_{|k|\geqslant \frac{1}{|y|}+1} \langle k\rangle^{-1-2\alpha} \leqslant 2 \int_{1/|y|}^{\infty} r^{-1-2\alpha} dr \leqslant C|y|^{2\alpha}\,.$$

For the first sum, we note that $|e^{iky}-1| \leqslant k^2|y|^2$ if $|k||y| \leqslant 1$. Once again, we employ the integral test:

$$\begin{aligned} &\sum_{0<|k|\leqslant \frac{1}{|y|}} |e^{iky}-1|^2 \langle k\rangle^{-1-2\alpha} \\ &\qquad \leqslant |e^{iy}-1|^2 + |e^{-iy}-1|^2 + \sum_{2\leqslant |k|\leqslant \frac{1}{|y|}} |e^{iky}-1|^2 \langle k\rangle^{-1-2\alpha} \\ &\qquad \leqslant 2|y|^2 + 2\int_1^{\frac{1}{|y|}} |y|^2 r^2 r^{-1-2\alpha} dr \leqslant C_\alpha\big(|y|^2 + |y|^{2\alpha}\big) \end{aligned}$$

for some constant $C = C_\alpha$. Since $|y| \leqslant 1/2$, we see that

$$\sum_{k \neq 0} |e^{iky} - 1|^2 \langle k \rangle^{-1-2\alpha} \leqslant C_\alpha |y|^\alpha$$

as $\alpha < 1$.

Next, according to Theorem 6.8, if $f \in H^{3/2}(\mathbb{S}^1)$, then $u = \mathrm{PI}(f)$ solves $-\Delta u = 0$ in D with $u = f$ on $\partial \mathrm{D}$, and $\|u\|_{H^2(\mathrm{D})} \leqslant C\|f\|_{H^{3/2}(\mathbb{S}^1)}$. By Theorem 2.36,

$$\|Du\|_{L^q(\mathrm{D})} \leqslant C\sqrt{q}\|u\|_{H^2(\mathrm{D})} \quad \forall\, q \in [1, \infty).$$

Hence, by Morrey's inequality, we see that $u \in \mathscr{C}^{0,1-2/q}(\mathrm{D})$, and thus in $\mathscr{C}^{0,\alpha}(\mathrm{D})$ for $\alpha \in (0,1)$.

6.3 Exercises

Problem 6.1. Given $f \in L^1(\mathbb{S}^1)$, $0 < r < 1$, define

$$P_r f(\theta) = \sum_{n=-\infty}^{\infty} \widehat{f}_n r^{|n|} e^{in\theta}, \quad \widehat{f}_n = \frac{1}{2\pi} \int_0^{2\pi} f(\theta) e^{-in\theta} d\theta \,.$$

Show that

$$P_r f(\theta) = p_r * f(\theta) = \frac{1}{2\pi} \int_0^{2\pi} p_r(\theta - \varphi) f(\varphi) d\varphi \,,$$

where

$$p_r(\theta) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{in\theta} = \frac{1 - r^2}{1 - 2r\cos\theta + r^2} \,.$$

Show that $\displaystyle\frac{1}{2\pi} \int_0^{2\pi} p_r(\theta) d\theta = 1$.

Problem 6.2. If $f \in L^p(\mathbb{S}^1)$, $1 \leqslant p < \infty$, show that

$$P_r f \to f \quad \text{in} \quad L^p(\mathbb{S}^1) \ \text{ as } \ r \nearrow 1 \,.$$

Problem 6.3. Let $\mathrm{D} := B(0,1) \subseteq \mathbb{R}^2$ and let u satisfy the Neumann problem

$$\Delta u = 0 \qquad \text{in} \quad \mathrm{D}\,, \tag{6.4a}$$

$$\frac{\partial u}{\partial r} = g \qquad \text{on} \quad \partial \mathrm{D} := \mathbb{S}^1 \,. \tag{6.4b}$$

If $u = \mathrm{PI}(f) := \sum_{k\in\mathbb{Z}} \widehat{f}_k r^{|k|} e^{ik\theta}$, show that for $f \in H^{3/2}(\mathbb{S}^1)$,

$$g = \Lambda f, \tag{6.5}$$

which is the same as

$$\widehat{g}_k = |k| \widehat{f}_k \,.$$

Λ denotes the Dirichlet to Neumann map given by $\Lambda f(\theta) = \sum_{k\in\mathbb{Z}} \widehat{f}_k |k| e^{ik\theta}$ or $\Lambda f = -i\dfrac{\partial}{\partial\theta} Hf = -iH\dfrac{\partial f}{\partial\theta}$, where H is the Hilbert transform, defined by $Hu(\theta) = \sum_{k\in\mathbb{Z}} (\operatorname{sgn} k)\widehat{g}_k e^{ik\theta}$.

Problem 6.4. Define the function $K(\theta) = \sum_{k\neq 0} |k|^{-1} e^{ik\theta}$. Show that $K \in L^2(\mathbb{S}^1) \subseteq L^1(\mathbb{S}^1)$. Next, show that if $g \in L^2(\mathbb{S}^1)$ and $\int_{\mathbb{S}^1} g(\theta) d\theta = 0$, a solution to (6.5) is given by $f(\theta) = (2\pi)^{-1} \int_{\mathbb{S}^1} K(\theta - \varphi) g(\varphi) d\varphi$.

Problem 6.5. Consider the solution to the Neumann problem (6.4a) and (6.4b). Show that $g \in H^{1/2}(\mathbb{S}^1)$ implies that $u \in H^2(\mathrm{D})$ and that

$$\|u\|^2_{H^2(\mathrm{D})} \leqslant C\big(\|g\|^2_{H^{1/2}(\mathbb{S}^1)} + \|u\|^2_{L^2(\mathrm{D})}\big) \,.$$

Chapter 7

Regularity of the Laplacian on Ω

We have studied the regularity properties of the Laplace operator on $\mathrm{D} = B(0,1) \subseteq \mathbb{R}^2$ using the Poisson integral formula. These properties continue to hold on more general open, bounded, $\mathscr{C}^\infty$ subsets Ω of $\mathbb{R}^{\mathrm{n}}$.

We revisit the Dirichlet problem

$$\Delta u = 0 \qquad \text{in} \quad \Omega\,, \tag{7.1a}$$

$$u = f \qquad \text{on} \quad \partial\Omega\,. \tag{7.1b}$$

Theorem 7.1. *For $k \in \mathbb{N}$, given $f \in H^{k-\frac{1}{2}}(\partial\Omega)$, there exists a unique solution $u \in H^k(\Omega)$ to (7.1) satisfying*

$$\|u\|_{H^k(\Omega)} \leqslant C\|f\|_{H^{k-\frac{1}{2}}(\partial\Omega)}\,, \quad C = C(\Omega)\,.$$

Proof. **Step 1.** $k = 1$**.** We begin by converting (7.1) to a problem with homogeneous boundary conditions. Using the surjectivity of the trace operator provided by Theorem 4.15, there exists $F \in H^1(\Omega)$ such that $\tau F = f$ on $\partial\Omega$, and $\|F\|_{H^1(\Omega)} \leqslant C\|f\|_{H^{\frac{1}{2}}(\partial\Omega)}$. Let $U = u - F$; then $U \in H^1(\Omega)$ and by linearity of the trace operator, $\tau U = 0$ on $\partial\Omega$. It follows from Theorem 2.47 that $U \in H_0^1(\Omega)$ and satisfies $-\Delta U = \Delta F$ in $H_0^1(\Omega)$; that is

$$\langle -\Delta U, v\rangle = \langle \Delta F, v\rangle \qquad \forall\, v \in H_0^1(\Omega)\,.$$

According to Remark 2.64, $-\Delta : H_0^1(\Omega) \to H^{-1}(\Omega)$ is an isomorphism, so that $\Delta F \in H^{-1}(\Omega)$; therefore, by Theorem 2.63, there exists a unique weak solution $U \in H_0^1(\Omega)$, satisfying

$$\int_\Omega DU \cdot Dv\, dx = \langle \Delta F, v\rangle \qquad \forall\, v \in H_0^1(\Omega)\,,$$

with

$$\|U\|_{H^1(\Omega)} \leqslant C\|\Delta F\|_{H^{-1}(\Omega)}\,, \tag{7.2}$$

and hence

$$u = U + F \in H^1(\Omega) \quad \text{and} \quad \|u\|_{H^1(\Omega)} \leqslant \|f\|_{H^{\frac{1}{2}}(\partial\Omega)}\,.$$

Step 2. $k = 2$**.** Next, suppose that $f \in H^{1.5}(\partial\Omega)$. Again employing Theorem 4.15, we obtain $F \in H^2(\Omega)$ such that $\tau F = f$ and $\|F\|_{H^2(\Omega)} \leqslant C\|f\|_{H^{1.5}(\partial\Omega)}$; thus, we see that $\Delta F \in L^2(\Omega)$ and that, in fact,

$$\int_\Omega DU \cdot Dv\, dx = \int_\Omega \Delta F\, v\, dx \qquad \forall\, v \in H^1_0(\Omega)\,. \tag{7.3}$$

We first establish interior regularity. Choose *any* (nonempty) open sets $\Omega_1 \subset\subset \Omega_2 \subset\subset \Omega$ and let $\zeta \in \mathscr{C}^\infty_c(\Omega_2)$ with $0 \leqslant \zeta \leqslant 1$ and $\zeta = 1$ on Ω_1. Let $\epsilon_0 = \min \operatorname{dist}(\operatorname{spt}(\zeta), \partial\Omega_2)/2$. For all $0 < \epsilon < \epsilon_0$, define $U^\epsilon(x) = (\eta_\epsilon * U)(x)$ for all $x \in \Omega_2$, and set

$$v = -\eta_\epsilon * (\zeta^2 U^\epsilon,_j),_j\,.$$

Then $v \in H^1_0(\Omega)$ and can be used as a test function in (7.3); thus,

$$\begin{aligned} -\int_\Omega U,_i\, \eta_\epsilon * (\zeta^2 U^\epsilon,_j),_{ji}\, dx &= -\int_\Omega U,_i\, \eta_\epsilon * [\zeta^2 U^\epsilon,_{ij} + 2\zeta\zeta,_i U^\epsilon,_j],_j\, dx \\ &= \int_{\Omega_2} \zeta^2 U^\epsilon,_{ij} U^\epsilon,_{ij}\, dx - 2\int_\Omega \eta_\epsilon * [\zeta\zeta,_i U^\epsilon,_j],_j\, U,_i\, dx\,, \end{aligned}$$

and

$$\int_\Omega \Delta F\, v\, dx = -\int_{\Omega_2} \Delta F\, \eta_\epsilon * (\zeta^2 U^\epsilon,_j),_j\, dx = -\int_{\Omega_2} \Delta F\, \eta_\epsilon * [\zeta^2 U^\epsilon,_{jj} + 2\zeta\zeta,_j U^\epsilon,_j]\, dx\,.$$

By Young's inequality (Theorem 1.47),

$$\|\eta_\epsilon * [\zeta^2 U^\epsilon,_{jj} + 2\zeta\zeta,_j U^\epsilon,_j]\|_{L^2(\Omega_2)} \leqslant \|\zeta^2 U^\epsilon,_{jj} + 2\zeta\zeta,_j U^\epsilon,_j\|_{L^2(\Omega_2)};$$

hence, by the Cauchy-Young inequality with δ, Lemma 1.46, for $\delta > 0$,

$$\int_\Omega \Delta F\, v\, dx \leqslant \delta\|\zeta D^2 U^\epsilon\|^2_{L^2(\Omega_2)} + C_\delta[\|DU^\epsilon\|^2_{L^2(\Omega_2)} + \|\Delta F\|^2_{L^2(\Omega)}]\,.$$

Similarly,

$$2\int_\Omega \eta_\epsilon * [\zeta\zeta,_i U^\epsilon,_j],_j\, U,_i\, dx \leqslant \delta\|\zeta D^2 U^\epsilon\|^2_{L^2(\Omega_2)} + C_\delta[\|DU^\epsilon\|^2_{L^2(\Omega_2)} + \|\Delta F\|^2_{L^2(\Omega)}]\,.$$

By choosing $\delta < 1$ and readjusting the constant C_δ, we see that

$$\|D^2U^\epsilon\|^2_{L^2(\Omega_1)} \leqslant \|\zeta D^2U^\epsilon\|^2_{L^2(\Omega_2)} \leqslant C_\delta\big[\|DU^\epsilon\|^2_{L^2(\Omega_2)} + \|\Delta F\|^2_{L^2(\Omega)}\big] \leqslant C_\delta\|\Delta F\|^2_{L^2(\Omega)}\,,$$

the last inequality following from (7.2), and Young's inequality.

Since the right-hand side does not depend on $\epsilon > 0$, there exists a subsequence

$$D^2U^{\epsilon'} \rightharpoonup \mathcal{W} \quad \text{in} \quad L^2(\Omega_1)\,.$$

By Theorem 2.21, $U^\epsilon \to U$ in $H^1(\Omega_1)$, so that $\mathcal{W} = D^2U$ on Ω_1. As weak convergence is lower semi-continuous, $\|D^2U\|_{L^2(\Omega_1)} \leqslant C_\epsilon\|\Delta F\|_{L^2(\Omega)}$. As Ω_1 and Ω_2 are arbitrary, we have established that $U \in H^2_{\text{loc}}(\Omega)$ and that

$$\|U\|_{H^2_{\text{loc}}(\Omega)} \leqslant C\|\Delta F\|_{L^2(\Omega)}\,.$$

For any $w \in H^1_0(\Omega)$, set $v = \zeta w$ in (7.3). Since $u \in H^2_{\text{loc}}(\Omega)$, we may integrate by parts to find that

$$\int_\Omega (-\Delta U - \Delta F)\,\zeta w\,dx = 0 \qquad \forall\, w \in H^1_0(\Omega)\,.$$

Since w is arbitrary, and the spt(ζ) can be chosen arbitrarily close to $\partial\Omega$, it follows that for all x in the interior of Ω, we have that

$$-\Delta U(x) = \Delta F(x) \quad \text{for almost every } x \in \Omega\,. \tag{7.4}$$

We proceed to establish the regularity of U all the way to the boundary $\partial\Omega$. Let $\{\mathcal{U}_\ell\}_{\ell=1}^K$ denote an open cover of Ω which intersects the boundary $\partial\Omega$, and let $\{\vartheta_\ell\}_{\ell=1}^K$ denote a collection of charts such that

$$\begin{aligned}
&\vartheta_\ell : B(0, r_\ell) \to \mathcal{U}_\ell \text{ is a } \mathscr{C}^\infty \text{ diffeomorphism}\,,\\
&\det(D\vartheta_\ell) = 1\,,\\
&\vartheta_\ell(B(0, r_\ell) \cap \{x_{\rm n} = 0\}) \to \mathcal{U}_\ell \cap \partial\Omega\,,\\
&\vartheta_\ell(B(0, r_\ell) \cap \{x_{\rm n} > 0\}) \to \mathcal{U}_\ell \cap \Omega\,.
\end{aligned}$$

Let $0 \leqslant \zeta_\ell \leqslant 1$ in $\mathscr{C}^\infty_c(\mathcal{U}_\ell)$ denote a partition of unity subordinate to the open covering $\mathcal{U}_\ell$, and define the horizontal convolution operator, smoothing functions defined on $\mathbb{R}^{\rm n}$ in the first $1, ..., {\rm n}-1$ directions, as follows:

$$\rho_\epsilon *_h F(x_h, x_{\rm n}) = \int_{\mathbb{R}^{{\rm n}-1}} \rho_\epsilon(x_h - y_h) F(y_h, x_{\rm n}) dy_h\,,$$

where $\rho_\epsilon(x_h) = \epsilon^{-(\mathrm{n}-1)}\rho(x_h/\epsilon)$, ρ the standard mollifier on $\mathbb{R}^{\mathrm{n}-1}$, and $x_h = (x_1, ..., x_{\mathrm{n}-1})$. Let α range from 1 to $\mathrm{n}-1$, and substitute the test function

$$v = -\big(\rho_\epsilon *_h[(\zeta_\ell \circ \vartheta_\ell)^2 \rho_\epsilon *_h(U \circ \vartheta_\ell),_\alpha],_\alpha\big) \circ \vartheta_\ell^{-1} \in H^1_0(\Omega)$$

into (7.3), and use the change of variables formula to obtain the identity

$$\int_{B_+(0,r_\ell)} A_i^k (U \circ \vartheta_\ell),_k \, A_i^j (v \circ \vartheta_\ell),_j \; dx = \int_{B_+(0,r_\ell)} (\Delta F) \circ \vartheta_\ell \, v \circ \vartheta_\ell \, dx \,, \tag{7.5}$$

where the $\mathscr{C}^\infty$ matrix $A(x) = [D\vartheta_\ell(x)]^{-1}$ and $B_+(0, r_\ell) = B(0, r_\ell) \cap \{x_{\mathrm{n}} > 0\}$. We define

$$U^\ell = U \circ \vartheta_\ell \,, \text{ and denote the horizontal convolution operator by } H_\epsilon = \rho_\epsilon *_h \,.$$

Then, with $\xi_\ell = \zeta_\ell \circ \vartheta_\ell$, we can rewrite the test function as

$$v \circ \vartheta_\ell = -H_\epsilon[\xi_\ell^2 H_\epsilon U^\ell,_\alpha],_\alpha \;.$$

Since differentiation commutes with convolution, we have that

$$(v \circ \vartheta_\ell),_j = -H_\epsilon(\xi_\ell^2 H_\epsilon U^\ell,_{j\alpha}),_\alpha - 2H_\epsilon(\xi_\ell \xi_\ell,_j \, H_\epsilon U^\ell,_\alpha),_\alpha \;,$$

and we can express the left-hand side of (7.5) as

$$\int_{B_+(0,r_\ell)} A_i^k (U \circ \vartheta_\ell),_k \, A_i^j (v \circ \vartheta_\ell),_j \; dx = \mathcal{I}_1 + \mathcal{I}_2 \,,$$

where

$$\mathcal{I}_1 = -\int_{B_+(0,r_\ell)} A_i^j A_i^k U^\ell,_k \; H_\epsilon(\xi_\ell^2 H_\epsilon U^\ell,_{j\alpha}),_\alpha \; dx \,,$$
$$\mathcal{I}_2 = -2\int_{B_+(0,r_\ell)} A_i^j A_i^k U^\ell,_k \; H_\epsilon(\xi_\ell \xi_\ell,_j \, H_\epsilon U^\ell,_\alpha),_\alpha \; dx \,.$$

Next, we see that

$$\mathcal{I}_1 = \int_{B_+(0,r_\ell)} \big[H_\epsilon(A_i^j A_i^k U^\ell,_k)\big],_\alpha \; (\xi_\ell^2 H_\epsilon U^\ell,_{j\alpha})\, dx = \mathcal{I}_{1a} + \mathcal{I}_{1b} \,,$$

where

$$\mathcal{I}_{1a} = \int_{B_+(0,r_\ell)} (A_i^j A_i^k H_\epsilon U^\ell,_k),_\alpha \; \xi_\ell^2 H_\epsilon U^\ell,_{j\alpha} \; dx \,,$$
$$\mathcal{I}_{1b} = \int_{B_+(0,r_\ell)} \big([\![H_\epsilon, A_i^j A_i^k]\!] U^\ell,_k\big),_\alpha \; \xi_\ell^2 H_\epsilon U^\ell,_{j\alpha} \; dx \,,$$

and where

$$[\![H_\epsilon, A_i^j A_i^k]\!]U^\ell,_k = H_\epsilon(A_i^j A_i^k U^\ell,_k) - A_i^j A_i^k\, H_\epsilon U^\ell,_k \tag{7.6}$$

denotes the commutator of the horizontal convolution operator and multiplication. The integral $\mathcal{I}_{1a}$ produces the positive sign-definite term which will allow us to build the global regularity of U, as well as an error term:

$$\mathcal{I}_{1a} = \int_{B_+(0,r_\ell)} \big[\xi_\ell^2 A_i^j A_i^k H_\epsilon U^\ell,_{k\alpha}\; H_\epsilon U^\ell,_{j\alpha} + (A_i^j A_i^k),_\alpha\, H_\epsilon U^\ell,_k\; \xi_\ell^2 H_\epsilon U^\ell,_{j\alpha}\big]\, dx\,;$$

thus, together with the right hand-side of (7.5), we see that

$$\int_{B_+(0,r_\ell)} \xi_\ell^2 A_i^j A_i^k H_\epsilon U^\ell,_{k\alpha}\; H_\epsilon U^\ell,_{j\alpha}\, dx \leqslant \Big|\int_{B_+(0,r_\ell)} (A_i^j A_i^k),_\alpha\, H_\epsilon U^\ell,_k\; \xi_\ell^2 H_\epsilon U^\ell,_{j\alpha}\,]\, dx\Big|$$
$$+ |\mathcal{I}_{1b}| + |\mathcal{I}_2| + \Big|\int_{B_+(0,r_\ell)} (\Delta F)\circ\vartheta_\ell\, v\circ\vartheta_\ell\, dx\Big|\,.$$

Since each ϑ_ℓ is a $\mathscr{C}^\infty$-diffeomorphism, it follows that the matrix $A\,A^T$ is positive definite: there exists $\lambda > 0$ such that

$$\lambda|Y|^2 \leqslant A_i^j A_i^k Y_j Y_k \quad \forall\, Y \in \mathbb{R}^{\mathrm{n}}\,.$$

It follows that

$$\lambda \int_{B_+(0,r_\ell)} \xi_\ell^2 |\bar\partial D H_\epsilon U^\ell|^2\, dx \leqslant \Big|\int_{B_+(0,r_\ell)} (A_i^j A_i^k),_\alpha\, H_\epsilon U^\ell,_k\; \xi_\ell^2 H_\epsilon U^\ell,_{j\alpha}\,]\, dx\Big|$$
$$+ |\mathcal{I}_{1b}| + |\mathcal{I}_2| + \Big|\int_{B_+(0,r_\ell)} (\Delta F)\circ\vartheta_\ell\, v\circ\vartheta_\ell\, dx\Big|\,,$$

where $D = (\partial_{x_1}, ..., \partial_{x_{\mathrm{n}}})$ and $\bar\partial = (\partial_{x_1}, ..., \partial_{x_{\mathrm{n}-1}})$. Application of the Cauchy-Young inequality with $\delta > 0$ shows that

$$\Big|\int_{B_+(0,r_\ell)} (A_i^j A_i^k),_\alpha\, H_\epsilon U^\ell,_k\; \xi_\ell^2 H_\epsilon U^\ell,_{j\alpha}\,]\, dx\Big| + |\mathcal{I}_2| + \Big|\int_{B_+(0,r_\ell)} (\Delta F)\circ\vartheta_\ell\, v\circ\vartheta_\ell\, dx\Big|$$
$$\leqslant \delta \int_{B_+(0,r_\ell)} \xi_\ell^2 |\bar\partial D H_\epsilon U^\ell|^2\, dx + C_\delta \|\Delta F\|^2_{L^2(\Omega)}\,.$$

It remains to establish such an upper bound for $|\mathcal{I}_{1b}|$.

To do so, we first establish a pointwise bound for (7.6): for $\mathcal{A}^{jk} = A_i^j A_i^k$,

$$[\![H_\epsilon, A_i^j A_i^k]\!]U^\ell,_k(x)$$
$$= \int_{B(x_h,\epsilon)} \rho_\epsilon(x_h - y_h)\big[\mathcal{A}^{jk}(y_h, x_{\mathrm{n}}) - \mathcal{A}^{jk}(x_h, x_{\mathrm{n}})\big]U^\ell,_k(y_h, x_{\mathrm{n}})\, dy_h\,.$$

By Morrey's inequality, $\big|[\mathcal{A}^{jk}(y_h, x_{\rm n}) - \mathcal{A}^{jk}(x_h, x_{\rm n})]\big| \leqslant C\epsilon \|\mathcal{A}\|_{W^{1,\infty}(B_+(0,r_\ell))}$. Since

$$\partial_{x_\alpha} \rho_\epsilon(x_h - y_h) = \frac{1}{\epsilon^2}\rho'\Big(\frac{x-h-y_h}{\epsilon}\Big)\,,$$

we see that

$$\big|\partial_{x_\alpha}\big([\![H_\epsilon, A^j_i A^k_i]\!] U^\ell,_k\big)(x)\big| \leqslant C\int_{B(x_h,\epsilon)} \frac{1}{\epsilon}\rho'\Big(\frac{x-h-y_h}{\epsilon}\Big)\big|U^\ell,_k(y_h, x_{\rm n})\big|\, dy_h$$

and hence by Young's inquality,

$$\big\|\partial_{x_\alpha}\big([\![H_\epsilon, A^j_i A^k_i]\!] U^\ell,_k\big)\big\|_{L^2(B_+(0,r_\ell)} \leqslant C\|U\|_{H^1(\Omega)} \leqslant C\|\Delta F\|_{L^2(\Omega)}\,.$$

It follows from the Cauchy-Young inequality with $\delta > 0$ that

$$|\mathcal{I}_{1b}| \leqslant \delta \int_{B_+(0,r_\ell)} \xi_\ell^2 |\bar\partial D H_\epsilon U^\ell|^2\, dx + C_\delta \|\Delta F\|^2_{L^2(\Omega)}\,.$$

By choosing $2\delta < \lambda$, we obtain the estimate

$$\int_{B_+(0,r_\ell)} \xi_\ell^2 |\bar\partial D H_\epsilon U^\ell|^2\, dx \leqslant C_\delta \|\Delta F\|^2_{L^2(\Omega)}\,.$$

Since the right hand-side is independent of ϵ, we find that

$$\int_{B_+(0,r_\ell)} \xi_\ell^2 |\bar\partial D U^\ell|^2\, dx \leqslant C_\delta \|\Delta F\|^2_{L^2(\Omega)}\,. \tag{7.7}$$

From (7.4), we know that $\Delta U(x) = \Delta F(x)$ for almost every $x \in \mathcal{U}_\ell$. By the chain rule this means that almost everywhere in $B_+(0, r_\ell)$,

$$-\mathcal{A}^{jk} U^\ell,_{kj} = \mathcal{A}^{jk},_j U^\ell,_k + \Delta F \circ \vartheta_\ell\,,$$

or equivalently,

$$-\mathcal{A}^{\rm nn} U^\ell,_{\rm nn} = \mathcal{A}^{j\alpha} U^\ell,_{\alpha j} + \mathcal{A}^{\beta k} U^\ell,_{k\beta} + \mathcal{A}^{jk},_j U^\ell,_k + \Delta F \circ \vartheta_\ell\,. \tag{7.8}$$

Since $\mathcal{A}^{\rm nn} > 0$, it follows from (7.7) that

$$\int_{B_+(0,r_\ell)} \xi_\ell^2 |D^2 U^\ell|^2\, dx \leqslant C_\delta \|\Delta F\|^2_{L^2(\Omega)}\,. \tag{7.9}$$

Summing over ℓ from 1 to K and combining with our interior estimates, we have that

$$\|u\|_{H^2(\Omega)} \leqslant C\|\Delta F\|_{L^2(\Omega)}\,.$$

Step 3. $k \geqslant 3$**.** At this stage, we have obtained a pointwise solution $U \in H^2(\Omega) \cap H^1_0(\Omega)$ to $\Delta u = \Delta F$ in Ω, and $\Delta F \in H^{k-1}(\Omega)$. Next, in each local chart, horizontally differentiate this equation r times until $\bar\partial^r(\Delta F \circ \vartheta_\ell) \in L^2(\Omega)$, and then repeat Step 2 using (7.8). □

Chapter 8

Fourier Series and its Applications

8.1 The Hilbert Space $L^2(\mathbb{T})$

A 2π-periodic function on $\mathbb{R}$ can be identified with a function on the circle, or one-dimensional torus $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$ on which we identify points in $\mathbb{R}$ that differ by $2\pi n$ for some $n \in \mathbb{Z}$. We use $\mathscr{C}(\mathbb{T})$ to denote the space of continuous functions on $\mathbb{R}$ with period 2π. The space $L^2(\mathbb{T})$ is defined as the completion of $\mathscr{C}(\mathbb{T})$ with respect to the L^2-norm

$$\|f\|_{L^2(\mathbb{T})} = \Big[\int_{\mathbb{T}} |f(x)|^2 dx \Big]^{\frac{1}{2}}$$

and we note that the norm is induced by the inner product

$$(f, g)_{L^2(\mathbb{T})} = \int_{\mathbb{T}} f(x)\overline{g(x)}\, dx\,.$$

Quantitatively speaking, the space $L^2(\mathbb{T})$ is the same as $L^2([-\pi, \pi])$; however, when speaking of $L^2(\mathbb{T})$, we are concerned with 2π-periodic L^2-functions, while the L^2-norm is computed only on the interval with length 2π.

Since $L^2(\mathbb{T})$ is a Hilbert space, it is nature to ask if there is an orthonormal basis to $L^2(\mathbb{T})$. The goal of this section is to provide an orthonormal basis to $L^2(\mathbb{T})$.

Definition 8.1. The Fourier basis elements are the functions

$$\mathrm{e}_k(x) = \frac{1}{\sqrt{2\pi}}\, e^{ikx}\,.$$

We note that $\{\mathrm{e}_k\}_{k=-\infty}^{\infty}$ is an orthonormal set in $L^2(\mathbb{T})$. In the following discussion, we will show that $\{\mathrm{e}_k\}_{k=-\infty}^{\infty}$ is maximal; that is, for each $f \in L^2(\mathbb{T})$,

$$f = \sum_{k=-\infty}^{\infty} (f, \mathrm{e}_k)_{L^2(\mathbb{T})}\mathrm{e}_k \quad \text{or} \quad f(x) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \int_{\mathbb{T}} f(y) e^{ik(x-y)} dy\,,$$

where the sum is understood as the L^2-limit.

8.1.1 Trigonometric polynomials

Definition 8.2. A trigonometric polynomial $p(x)$ of degree n is a finite sum of the form

$$p(x) = \frac{c_0}{2} + \sum_{k=1}^{n} (c_k \cos kx + s_k \sin kx) \qquad x \in \mathbb{R}\,.$$

The collection of all trigonometric polynomial of degree n is denoted by $\mathscr{P}_n(\mathbb{T})$, and the collection of all trigonometric polynomials is denoted by $\mathscr{P}(\mathbb{T})$; that is, $\mathscr{P}(\mathbb{T}) = \bigcup\limits_{n=0}^{\infty} \mathscr{P}_n(\mathbb{T})$.

On account of the Euler identity $e^{i\theta} = \cos\theta + i\sin\theta$, a trigonometric polynomial of degree n can also be written as

$$p(x) = \sum_{k=-n}^{n} a_k e^{ikx} \quad \text{with} \quad a_k = \frac{c_{|k|} - is_{|k|}}{2}\,,$$

where s_0 is taken to be 0. Therefore, every trigonometric polynomial of degree n is associated to a unique function of the form $\sum\limits_{k=-n}^{n} a_k e^{ikx}$ and vice versa.

Definition 8.3. The Fourier series associated to a function $f \in L^2(\mathbb{T})$ is the function

$$\sum_{k=-\infty}^{\infty} \widehat{f}(k)\mathrm{e}_k(x) = \frac{c_0}{2} + \sum_{k=1}^{\infty} c_k \cos kx + s_k \sin kx\,,$$

where $\big\{\widehat{f}(k)\big\}_{k=-\infty}^{\infty}$, $\{c_k\}_{k=0}^{\infty}$ and $\{s_k\}_{k=1}^{\infty}$ are called the Fourier coefficients of f given by

$$\widehat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} f(x) e^{-ikx} dx\,, c_k = \frac{1}{\pi} \int_{\mathbb{T}} f(x) \cos kx dx\,, s_k = \frac{1}{\pi} \int_{\mathbb{T}} f(x) \sin kx dx\,.$$

Proposition 8.4. *Let $s_n(f,x)$ denote the partial sum of the Fourier series associated to $f \in L^2(\mathbb{T})$ given by*

$$s_n(f,x) = \frac{1}{2\pi} \sum_{k=-n}^{n} \int_{\mathbb{T}} f(y) e^{ik(x-y)} dy = \frac{c_0}{2} + \sum_{k=1}^{n} c_k \cos kx + s_k \sin kx\,,$$

where $\widehat{f}(k)$, c_k and s_k are defined in Definition 8.3. *Then*

$$\|f - s_n(f,\cdot)\|_{L^2(\mathbb{T})} \leqslant \|f - p\|_{L^2(\mathbb{T})} \qquad \forall\, p \in \mathscr{P}_n(\mathbb{T})\,.$$

Proof. We note that if $p \in \mathscr{P}_n(\mathbb{T})$, then $s_n(p,\cdot) = p$ and

$$\big(f - s_n(f,\cdot), p\big)_{L^2(\mathbb{T})} = 0\,.$$

Therefore, if $p \in \mathscr{P}_n(\mathbb{T})$,

$$\begin{aligned}\|f-p\|^2_{L^2(\mathbb{T})} &= \|f - s_n(f,\cdot) + s_n(f,\cdot) - p\|^2_{L^2(\mathbb{T})}\\ &= \|f - s_n(f,\cdot)\|^2_{L^2(\mathbb{T})} + \|s_n(f-p,\cdot)\|^2_{L^2(\mathbb{T})}\end{aligned}$$

which concludes the proposition. □

8.1.2 Approximations of the identity

Definition 8.5. A family of functions $\big\{\varphi_n \in \mathscr{C}(\mathbb{T}) \,\big|\, n \in \mathbb{N}\big\}$ is an approximation of the identity if

(1) $\varphi_n(x) \geqslant 0$;

(2) $\displaystyle\int_{\mathbb{T}} \varphi_n(x)\,dx = 1$ for every $n \in \mathbb{N}$;

(3) $\displaystyle\lim_{n\to\infty} \int_{\delta \leqslant |x| \leqslant \pi} \varphi_n(x)\,dx = 0$ for every $\delta > 0$, here we identify $\mathbb{T}$ with the interval $[-\pi, \pi]$.

Definition 8.6 (Convolutions on $\mathbb{T}$)**.** The convolution of two continuous function $f, g : \mathbb{T} \to \mathbb{C}$ is the continuous function $f \star g : \mathbb{T} \to \mathbb{C}$ defined by the integral

$$(f \star g)(x) = \int_{\mathbb{T}} f(x-y)g(y)\,dy\,.$$

Note that all the conclusions from Section 1.4 are still valid. In particular, we have

Theorem 8.7. *If $\{\varphi_n\}_{n=1}^{\infty}$ is an approximation of the identity and $f \in \mathscr{C}(\mathbb{T})$, then $\varphi_n \star f$ converges uniformly to f as $n \to \infty$.*

Proof. Without loss of generality, we may assume that $f \not\equiv 0$. By the definition of the convolution,

$$\begin{aligned}\big|(\varphi_n \star f)(x) - f(x)\big| &= \int_{\mathbb{T}} \varphi_n(x-y)f(y)\,dy - f(x)\\ &= \int_{\mathbb{T}} \varphi_n(x-y)\big(f(x) - f(y)\big)dy\,,\end{aligned}$$

where we use (2) of Definition 8.5 to obtain the last equality. Now given $\varepsilon > 0$. Since $f \in \mathscr{C}(\mathbb{T})$, there exists $\delta > 0$ such that $|f(x) - f(y)| < \dfrac{\varepsilon}{2}$ whenever $|x - y| < \delta$. Therefore,

$$\begin{aligned}
&|(\varphi_n \star f)(x) - f(x)| \\
&\quad \leqslant \int_{|x-y|<\delta} \varphi_n(x-y)\big|f(x) - f(y)\big|dy + \int_{\delta \leqslant |x-y|} \varphi_n(x-y)\big|f(x) - f(y)\big|dy \\
&\quad \leqslant \frac{\varepsilon}{2} \int_{\mathbb{T}} \varphi_n(x-y)\, dy + 2 \max_{\mathbb{T}} |f| \int_{\delta \leqslant |z| \leqslant \pi} \varphi_n(z)\, dz\,.
\end{aligned}$$

By (3) of Definition 8.5, there exists $N > 0$ such that if $n \geqslant N$,

$$\int_{\delta \leqslant |z| \leqslant \pi} \varphi_n(z)\, dx < \frac{\varepsilon}{4 \max_{\mathbb{T}} |f|}\,.$$

Therefore, for $n \geqslant N$, $\big|(\varphi_n \star f)(x) - f(x)\big| < \varepsilon$ for all $x \in \mathbb{T}$. □

Theorem 8.8. *The collection of all trigonometric polynomials $\mathscr{P}(\mathbb{T})$ is dense in $\mathscr{C}(\mathbb{T})$ with respect to the uniform norm.*

Proof. Let $\varphi_n(x) = c_n(1 + \cos x)^n$, where c_n is chosen so that $\displaystyle\int_{\mathbb{T}} \varphi_n(x)\, dx = 1$. By the residue theorem,

$$\int_{\mathbb{T}} (1 + \cos x)^n dx = \oint_{\mathbb{S}^1} \Big(1 + \frac{z^2+1}{2z}\Big)^n \frac{dz}{iz} = \frac{1}{i2^n} \oint_{\mathbb{S}^1} \frac{(z+1)^{2n}}{z^{n+1}} dz = \frac{\pi}{2^{n-1}} \binom{2n}{n};$$

thus $c_n = \dfrac{2^{n-1}}{\pi} \dfrac{(n!)^2}{(2n)!}$.

Now $\{\varphi_n\}_{n=1}^\infty$ is clearly non-negative and satisfies (2) of Definition 8.5 for all $n \in \mathbb{N}$. Let $\delta > 0$ be given.

$$\int_{\delta \leqslant |x| \leqslant \pi} \varphi_n(x)\, dx \leqslant \int_{\delta \leqslant |x| \leqslant \pi} c_n(1 + \cos\delta)^n dx \leqslant 2^{2n} \Big(\frac{1 + \cos\delta}{2}\Big)^n \frac{(n!)^2}{(2n)!}\,.$$

By Stirling's formula $\displaystyle\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}\, n^n e^{-n}} = 1$,

$$\begin{aligned}
\lim_{n\to\infty} \int_{\delta \leqslant |x| \leqslant \pi} \varphi_n(x)\, dx &\leqslant \lim_{n\to\infty} 2^{2n} \Big(\frac{1 + \cos\delta}{2}\Big)^n \frac{\big(\sqrt{2\pi n}\, n^n e^{-n}\big)^2}{\sqrt{2\pi(2n)}(2n)^{2n} e^{-2n}} \\
&= \lim_{n\to\infty} \sqrt{\pi n} \Big(\frac{1 + \cos\delta}{2}\Big)^n = 0\,.
\end{aligned}$$

So $\{\varphi_n\}_{n=1}^\infty$ is an approximation of the identity. By Theorem 8.7, $\varphi_k \star f$ converges uniformly to f if $f \in \mathscr{C}(\mathbb{T})$, while $\varphi_n \star f$ is a trigonometric function. □

Corollary 8.9. *For any $f \in L^2(\mathbb{T})$, $\lim\limits_{n\to\infty} \|s_n(f,\cdot) - f\|_{L^2(\mathbb{T})} = 0$, and*

$$\|f\|^2_{L^2(\mathbb{T})} = \sum_{k=-\infty}^{\infty} |\widehat{f}(k)|^2 = \pi\Big[\frac{c_0^2}{2} + \sum_{k=1}^{\infty}(c_k^2 + s_k^2)\Big]. \quad (\textit{Parseval's identity}) \tag{8.1}$$

Proof. By Theorem 8.8, the collection of trigonometric polynomials is dense in $\mathscr{C}(\mathbb{T})$, we know that the space spanned by $\{\mathrm{e}_k\}_{k=-\infty}^{\infty}$ is dense in $\mathscr{C}(\mathbb{T})$. The implication from uniform convergence to $L^2(\mathbb{T})$-convergence then guarantees that the space spanned by $\{\mathrm{e}_k\}_{k=-\infty}^{\infty}$ is dense in $L^2(\mathbb{T})$.

Let $\varepsilon > 0$ be given. By the denseness of the trigonometric polynomials in $L^2(\mathbb{T})$, there exists $h \in \mathscr{P}(\mathbb{T})$ such that $\|f - h\|_{L^2(\mathbb{T})} < \varepsilon$. Suppose that $h \in \mathscr{P}_N(\mathbb{T})$. Then by Proposition 8.4,

$$\|f - s_N(f,\cdot)\|_{L^2(\mathbb{T})} \leqslant \|f - h\|_{L^2(\mathbb{T})} < \varepsilon\,.$$

Since $s_N(f,\cdot) \in \mathscr{P}_n(\mathbb{T})$ if $n \geqslant N$, we must have

$$\|f - s_n(f,\cdot)\|_{L^2(\mathbb{T})} \leqslant \|f - s_N(f,\cdot)\|_{L^2(\mathbb{T})} < \varepsilon \qquad \forall\, n \geqslant N\,.$$

Therefore, $s_n(f,\cdot) \to f$ in $L^2(\mathbb{T})$ as $n \to \infty$, and (8.1) is concluded by the fact that $\|s_n(f,\cdot)\|_{L^2(\mathbb{T})} \to \|f\|_{L^2(\mathbb{T})}$ as $n \to \infty$. □

The proof of the following lemma is left as an exercise.

Lemma 8.10. *Let $f, g \in L^2(\mathbb{T})$. Then*

$$(f,g)_{L^2(\mathbb{T})} = \sum_{k=-\infty}^{\infty} \widehat{f}(k)\overline{\widehat{g}(k)}.$$

By Parseval's identity (8.1), for $f \in L^2(\mathbb{T})$,

$$\lim_{|k|\to\infty} |\widehat{f}(k)| = 0\,.$$

In fact, the Fourier coefficient of a function $f \in L^1(\mathbb{T})$ also converges to 0 which is the Riemann-Lebesgue Lemma.

Lemma 8.11 (Riemann-Lebesgue)**.** *For $f \in L^1(\mathbb{T})$, $\widehat{f}(k) \to 0$ as $|k| \to \infty$.*

Proof. Let $f \in L^1(\mathbb{T})$. Given $\varepsilon > 0$, there exists $f_\varepsilon \in L^2(\mathbb{T})$ such that $\|f - f_\varepsilon\|_{L^1(\mathbb{T})} < \dfrac{\varepsilon}{2}$. For this ε, there exists $N > 0$ such that $|\widehat{f_\varepsilon}(k)| < \dfrac{\varepsilon}{2}$ whenever $k \geqslant N$. Therefore, for

$k \geqslant N$,

$$\begin{aligned}|\widehat{f}(k)| &\leqslant |\widehat{f}(k) - \widehat{f}_\varepsilon(k)| + |\widehat{f}_\varepsilon(k)| = \Big| \int_{\mathbb{T}} \big[f(x) - f_\varepsilon(x)\big] \mathrm{e}_k(x)\, dx \Big| + |\widehat{f}_\varepsilon(k)| \\ &\leqslant \|f - f_\varepsilon\|_{L^1(\mathbb{T})} + |\widehat{f}_\varepsilon(k)| < \varepsilon\end{aligned}$$

which implies that $\lim\limits_{|k|\to\infty} \widehat{f}(k) = 0$. □

Definition 8.12 (Weak convergence)**.** Let $\mathcal{H}$ be a Hilbert space with inner product $(\cdot,\cdot)_{\mathcal{H}}$. A sequence $u_n \in \mathcal{H}$ is said to converge weakly to $u \in \mathcal{H}$ if

$$\lim_{n\to\infty} (u_n, g)_{\mathcal{H}} = (u, g)_{\mathcal{H}} \qquad \forall\, g \in \mathcal{H}\,.$$

We use the notation $u_n \rightharpoonup u$ in $\mathcal{H}$ to denote the weak convergence of u_n to u in $\mathcal{H}$.

With this definition, by the Riemann-Lebesgue Lemma we have the following

Theorem 8.13. *The Fourier basis* e_k *converges weakly to* 0 *in* $L^2(\mathbb{T})$.

8.1.3 Fourier representation of functions on $[0,\pi]$

Any functions defined on $[0,\pi]$ can be viewed as the restriction of an even/odd function defined on $[-\pi,\pi]$ to $[0,\pi]$. An even/odd function f in $L^2([-\pi,\pi])$ can be expressed as

$$f(x) = \frac{c_0}{2} + \sum_{k=1}^{\infty} c_k \cos kx \Big/ \sum_{k=1}^{\infty} s_k \sin kx\,.$$

For a function $f \in L^2(0,2\pi)$ (here we identify $\mathbb{T}$ with $[0,2\pi]$), $g(x) = f(2x)$ is a function in $L^2(0,\pi)$. Since

$$f(x) = \frac{c_0}{2} + \sum_{k=1}^{\infty} (c_k \cos kx + s_k \sin kx)\,,$$

we have

$$g(x) = \frac{c_0}{2} + \sum_{k=1}^{\infty} (c_k \cos 2kx + s_k \sin 2kx)\,,$$

where

$$c_k = \frac{2}{\pi} \int_{\mathbb{T}} g(x) \cos kx dx\,, \qquad s_k = \frac{2}{\pi} \int_{\mathbb{T}} g(x) \sin kx dx\,.$$

As a consequence, $\Big\{\dfrac{1}{\sqrt{\pi}}, \sqrt{\dfrac{2}{\pi}}\cos kx\Big\}_{k=1}^{\infty}$, $\Big\{\sqrt{\dfrac{2}{\pi}}\sin kx\Big\}_{k=1}^{\infty}$ are both maximal orthonormal sets on $L^2(0,\pi)$. So is $\Big\{\dfrac{1}{\sqrt{\pi}}, \sqrt{\dfrac{2}{\pi}}\cos 2kx, \sqrt{\dfrac{2}{\pi}}\sin 2kx\Big\}_{k=1}^{\infty}$.

We note that $\Big\{\pm\dfrac{1}{\sqrt{\pi}}, \pm\sqrt{\dfrac{2}{\pi}}\cos kx\Big\}_{k=1}^{\infty}$ is the collection of all non-trivial functions with unit $L^2(0,\pi)$-norm satisfying

$$\begin{aligned} u_{xx} &= \lambda u \qquad \text{for some } \lambda \in \mathbb{R}\,, \\ u_x(0) &= u_x(\pi) = 0\,, \end{aligned}$$

while $\Big\{\pm\sqrt{\dfrac{2}{\pi}}\sin kx\Big\}_{k=1}^{\infty}$ is the collection of all non-trivial functions with unit $L^2(0,\pi)$-norm satisfying

$$\begin{aligned} u_{xx} &= \lambda u \qquad \text{for some } \lambda \in \mathbb{R}\,, \\ u(0) &= u(\pi) = 0\,. \end{aligned}$$

8.2 Uniform Convergence of the Fourier Series

Given $f \in L^2(\mathbb{T})$, by Corollary 8.9 we know that $s_n(f,\cdot) \to f$ in $L^2(\mathbb{T})$; thus possesses a subsequence $s_{n_j}(f,\cdot)$ which converges to f almost everywhere. In this section, instead of assuming that $f \in L^2(\mathbb{T})$, we consider $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$, and investigate the convergence behavior of the Fourier series of f.

Before proceeding, we define

$$\begin{aligned} D_n(x) &= \frac{1}{2\pi}\sum_{k=-n}^{n} e^{ikx} = \frac{1}{2\pi}\frac{e^{-inx}\big[e^{i(2n+1)x} - 1\big]}{e^{ix} - 1} \\ &= \frac{1}{2\pi}\frac{e^{i(n+1/2)x} - e^{-i(n+1/2)x}}{e^{ix/2} - e^{-ix/2}} = \frac{\sin(n+\frac{1}{2})x}{2\pi\sin\frac{x}{2}}\,. \end{aligned}$$

Then

$$\begin{aligned} s_n(f,x) &= \frac{1}{\sqrt{2\pi}}\sum_{k=-n}^{n}\widehat{f}(k)e^{ikx} = \sum_{k=-n}^{n}\frac{1}{2\pi}\int_{\mathbb{T}} f(y)\, e^{ik(x-y)}dy \\ &= \int_{\mathbb{T}} f(y)D_n(x-y)\,dy = (D_n \star f)(x)\,, \end{aligned}$$

and Corollary 8.9 states that $D_n \star f \to f$ in $L^2(\mathbb{T})$ for all $f \in L^2(\mathbb{T})$.

Definition 8.14. The function

$$D_n(x) = \frac{\sin(n+\frac{1}{2})x}{2\pi \sin\frac{x}{2}} \tag{8.2}$$

is called the *Dirichlet kernel*.

8.2.1 Uniform convergence

In the following, we first consider an easier case $f \in \mathscr{C}^{0,1}(\mathbb{T})$; that is, f is Lipschitz continuous on $\mathbb{T}$. We note that if $f \in \mathscr{C}^{0,1}(\mathbb{T})$, then f is absolutely continuous, differentiable a.e., and satisfies the integration by parts formula

$$\int_a^b f(x)g'(x)\,dx = f(x)g(x)\Big|_{x=a}^{x=b} - \int_a^b f'(x)g(x)\,dx \quad \forall\, g \in \mathscr{C}^1(\mathbb{T})\,.$$

The identity above allows us to prove the uniform convergence much more easily. We have the following

Theorem 8.15. *For any $f \in \mathscr{C}^{0,1}(\mathbb{T})$, $s_n(f,\cdot) = D_n \star f$ converges to f uniformly as $n \to \infty$.*

Proof. Since $\displaystyle\int_{\mathbb{T}} D_n(x-y)\,dy = 1$ for all $x \in \mathbb{T}$,

$$\begin{aligned} s_n(f,x) - f(x) = (D_n \star f - f)(x) &= \int_{\mathbb{T}} D_n(x-y)\big(f(y)-f(x)\big)dy \\ &= \int_{\mathbb{T}} D_n(y)\big(f(x+y)-f(x)\big)dy\,. \end{aligned}$$

We break the integral into two parts: one is the integral over $|y| \leqslant \delta$ and the other is the integral over $\delta < |y| \leqslant \pi$. Since $f \in \mathscr{C}^{0,1}(\mathbb{T})$,

$$|f(y+x) - f(x)| \leqslant \|f'\|_{L^\infty(\mathbb{T})}|y|\,;$$

thus

$$\begin{aligned} \Big|\int_{|y|\leqslant\delta} D_n(y)\big(f(x+y)-f(x)\big)dy\Big| &\leqslant \int_{-\delta}^{\delta} \frac{\big|f(x+y)-f(x)\big|}{\big|\sin\frac{y}{2}\big|}dy \\ &\leqslant \|f'\|_{L^\infty(\mathbb{T})} \int_{-\delta}^{\delta} \frac{y}{\sin\dfrac{y}{2}}dy \leqslant C\delta\,. \end{aligned} \tag{8.3}$$

As for the integral over $\delta < |y| \leqslant \pi$, we have

$$\begin{aligned}\int_\delta^\pi \sin\big(n+\frac{1}{2}\big)y \frac{f(x+y)-f(x)}{\sin\frac{y}{2}}dy &= -\frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{f(x+y)-f(x)}{\sin\frac{y}{2}}\Big|_{y=\delta}^{y=\pi} \\ &\quad + \int_\delta^\pi \frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{d}{dy}\frac{f(x+y)-f(x)}{\sin\frac{y}{2}}dy\,.\end{aligned}$$

Note that the first term on the right-hand side converges to 0 uniformly as $n\to\infty$. For the second term,

$$\begin{aligned}&\Big|\int_\delta^\pi \frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{d}{dy}\frac{f(x+y)-f(x)}{\sin\frac{y}{2}}dy\Big| \\ &\quad\leqslant \Big|\int_\delta^\pi \frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{f'(x+y)}{\sin\frac{y}{2}}dy\Big| + \Big|\int_\delta^\pi \frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{\cos\frac{y}{2}\big(f(x+y)-f(x)\big)}{\sin^2\frac{y}{2}}dy\Big| \\ &\quad\leqslant \|f'\|_{L^\infty(\mathbb{T})}\frac{\pi-\delta}{\big(n+\frac{1}{2}\big)\sin\frac{\delta}{2}} + \|f\|_{L^\infty(\mathbb{T})}\frac{2(\pi-\delta)}{\big(n+\frac{1}{2}\big)\sin^2\frac{\delta}{2}} \leqslant \|f\|_{\mathscr{C}^{0,1}(\mathbb{T})}\frac{2(\pi-\delta)}{\big(n+\frac{1}{2}\big)\sin^2\frac{\delta}{2}}\,.\end{aligned}$$

Therefore, combining the estimate above with (8.3), we find that

$$\limsup_{n\to\infty}\sup_{x\in\mathbb{T}}\Big|\int_{\mathbb{T}} \sin Ly\frac{f(x+y)-f(x)}{\sin\frac{y}{2}}dy\Big| \leqslant C\delta\,;$$

and the conclusion follows from that $\delta > 0$ is chosen arbitrarily. □

The uniform convergence of $s_n(f,\cdot)$ to f for $f\in\mathscr{C}^{0,\alpha}(\mathbb{T})$ with $\alpha\in(0,1)$ requires a lot more work. The idea is to estimate $\big\|f-s_n(f,\cdot)\big\|_{L^\infty(\mathbb{T})}$ in terms of the quantity $\inf\limits_{p\in\mathscr{P}_n(\mathbb{T})}\|f-p\|_{L^\infty(\mathbb{T})}$. Since $s_n(f,\cdot)\in\mathscr{P}_n(\mathbb{T})$, it is obvious that

$$\inf_{p\in\mathscr{P}_n(\mathbb{T})}\|f-p\|_{L^\infty(\mathbb{T})} \leqslant \big\|f-s_n(f,\cdot)\big\|_{L^\infty(\mathbb{T})}\,.$$

The goal is to show the inverse inequality

$$\big\|f-s_n(f,\cdot)\big\|_{L^\infty(\mathbb{T})} \leqslant C_n \inf_{p\in\mathscr{P}_n(\mathbb{T})}\|f-p\|_{L^\infty(\mathbb{T})} \tag{8.4}$$

for some constant C_n, and pick a suitable $p\in\mathscr{P}_n(\mathbb{T})$ which gives a good upper bound for $\big\|f-s_n(f,\cdot)\big\|_{L^\infty(\mathbb{T})}$. The inverse inequality is established via the following

Proposition 8.16. *The Dirichlet kernel D_n satisfies that for all $n\in\mathbb{N}$,*

$$\int_{\mathbb{T}}\big|D_n(x)\big|dx \leqslant 2+\log n\,. \tag{8.5}$$

Proof. The validity of (8.5) for the case $n = 1$ is left to the reader, and we provide the proof for the case $n \geqslant 2$ here. Recall that $D_n(x) = \sum\limits_{k=-n}^{n} \dfrac{e^{ikx}}{2\pi} = \dfrac{\sin(n+\frac{1}{2})x}{2\pi \sin\frac{x}{2}}$. Therefore,

$$\int_{\mathbb{T}} \big|D_n(x)\big| dx = 2\int_0^{\pi} \big|D_n(x)\big| dx = \int_0^{\frac{1}{n}} 2\big|D_n(x)\big| dx + \int_{\frac{1}{n}}^{\pi} \Big|\frac{\sin(n+\frac{1}{2})x}{\pi \sin\frac{x}{2}}\Big| dx \,.$$

The first integral can be estimated by

$$\int_0^{\frac{1}{n}} 2\big|D_n(x)\big| dx \leqslant \frac{1}{\pi}\frac{2n+1}{n} \,. \tag{8.6}$$

Since $\dfrac{2x}{\pi} \leqslant \sin x$ for $0 \leqslant x \leqslant \dfrac{\pi}{2}$, the second integral can be estimated by

$$\int_{\frac{1}{n}}^{\pi} \Big|\frac{\sin(n+\frac{1}{2})x}{\pi \sin\frac{x}{2}}\Big| dx \leqslant \int_{\frac{1}{n}}^{\pi} \frac{1}{x}\, dx = \log \pi + \log n \,. \tag{8.7}$$

We then conclude (8.5) from (8.6) and (8.7) by noting that $\log \pi + \dfrac{2n+1}{n\pi} \leqslant 2$ for all $n \geqslant 2$. □

Remark 8.17. A more subtle estimate can be done to show that

$$\int_{\mathbb{T}} \big|D_n(x)\big| dx \geqslant c_1 + c_2 \log n \qquad \forall\, n \in \mathbb{N}$$

for some positive constants c_1 and c_2.

With the help of Proposition 8.16, we are able to prove the inverse inequality (8.4). The following theorem is a direct consequence of Proposition 8.16.

Theorem 8.18. *Let $f \in \mathscr{C}(\mathbb{T})$; that is, f is a continuous function with period 2π. Then*

$$\big\|f - s_n(f,\cdot)\big\|_{L^\infty(\mathbb{T})} \leqslant (3 + \log n) \inf_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_{L^\infty(\mathbb{T})} \,. \tag{8.8}$$

Proof. For $n \in \mathbb{N}$ and $x \in \mathbb{T}$,

$$\big|s_n(f,x)\big| \leqslant \int_{\mathbb{T}} \big|D_n(y)\big|\big|f(x-y)\big| dy \leqslant (2 + \log n)\|f\|_{L^\infty(\mathbb{T})} \,.$$

Given $\epsilon > 0$, let $p \in \mathscr{P}_n(\mathbb{T})$ such that

$$\|f - p\|_{L^\infty(\mathbb{T})} \leqslant \inf_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_{L^\infty(\mathbb{T})} + \epsilon.$$

Then by the fact that $s_n(p,x) = p(x)$ if $p \in \mathscr{P}_n(\mathbb{T})$, we obtain that

$$\begin{aligned}\|f - s_n(f,\cdot)\|_{L^\infty(\mathbb{T})} &\leqslant \|f-p\|_{L^\infty(\mathbb{T})} + \|p - s_n(f,\cdot)\|_{L^\infty(\mathbb{T})} \\ &\leqslant \|f-p\|_{L^\infty(\mathbb{T})} + \|s_n(f-p,\cdot)\|_{L^\infty(\mathbb{T})} \\ &\leqslant \|f-p\|_{L^\infty(\mathbb{T})} + (2+\log n)\|f-p\|_{L^\infty(\mathbb{T})} \\ &\leqslant (3+\log n)\Big[\inf_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_{L^\infty(\mathbb{T})} + \epsilon\Big],\end{aligned}$$

and (8.8) is obtained by passing to the limit as $\epsilon \to 0$. □

Having established Theorem 8.18, the study of the uniform convergence of $s_n(f,\cdot)$ to f then amounts to the study of the quantity $\inf\limits_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_{L^\infty(\mathbb{T})}$. In Exercise Problem 8.2, the reader is asked to show that

$$\inf_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_{L^\infty(\mathbb{T})} \leqslant \frac{1+2\log n}{2n}\|f\|_{\mathscr{C}^{0,1}(\mathbb{T})};$$

thus by Theorem 8.18, $s_n(f,\cdot)$ converges to f uniformly as $n\to\infty$ if $f \in \mathscr{C}^{0,1}(\mathbb{T})$, a restatement of Theorem 8.15.

The estimate of $\inf\limits_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_{L^\infty(\mathbb{T})}$ for $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$ is more difficult, and requires a clever choice of p. We begin with the following

Lemma 8.19. *If f is a continuous function on $[a,b]$, then for all $\delta_1 > 0$,*

$$\sup_{|x-y|\leqslant\delta_1} |f(x)-f(y)| \leqslant \Big(1 + \frac{\delta_1}{\delta_2}\Big) \sup_{|x-y|\leqslant\delta_2} |f(x)-f(y)|.$$

The proof of Lemma 8.19 is not very difficult, and is left to the readers.

Now we are in the position of prove the theorem due to D Jackson.

Theorem 8.20 (Jackson)**.** *Let f be a 2π-periodic continuous function. Then for some constant $C > 0$,*

$$\inf_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_{L^\infty(\mathbb{T})} \leqslant C \sup_{|x-y|\leqslant\frac{1}{n}} |f(x)-f(y)|.$$

Proof. Let $p(x) = 1 + c_1\cos x + \cdots + c_n \cos nx$ be a positive trigonometric function of degree n with coefficients $\{c_i\}_{i=1}^n$ determined later. Define an operator K on $\mathscr{C}(\mathbb{T})$ by

$$\mathrm{K}f(x) = \frac{1}{2\pi}\int_{\mathbb{T}} p(y) f(x-y)\,dy.$$

Then $\mathrm{K}f \in \mathscr{P}_n(\mathbb{T})$. Lemma 8.19 then implies

$$
\begin{aligned}
\big|\mathrm{K}f(x) - f(x)\big| &\leqslant \frac{1}{2\pi}\int_{\mathbb{T}} p(y)\big|f(x-y) - f(x)\big|dy \\
&\leqslant \frac{1}{2\pi}\int_{-\pi}^{\pi} p(y)\big(1+n|y|\big) \sup_{|x-y|\leqslant \frac{1}{n}} \big|f(x)-f(y)\big|dy \\
&= \Big[1 + \frac{n}{2\pi}\int_{-\pi}^{\pi}|y|p(y)\,dy\Big] \sup_{|x-y|\leqslant \frac{1}{n}} \big|f(x)-f(y)\big|\,.
\end{aligned}
$$

By Hölder's inequality and that $y^2 \leqslant \dfrac{\pi^2}{2}(1-\cos y)$ for $y \in \mathbb{T}$,

$$
\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi}|y|p(y)\,dy &\leqslant \Big[\frac{1}{2\pi}\int_{-\pi}^{\pi} y^2 p(y)\,dy\Big]^{\frac{1}{2}}\Big[\frac{1}{2\pi}\int_{-\pi}^{\pi} p(y)\,dy\Big]^{\frac{1}{2}} \\
&\leqslant \Big[\frac{\pi}{4}\int_{-\pi}^{\pi}(1-\cos y)p(y)\,dy\Big]^{\frac{1}{2}} = \frac{\pi}{2}\sqrt{2-c_1}\,.
\end{aligned}
$$

Therefore,

$$
\|\mathrm{K}f - f\|_{L^\infty(\mathbb{T})} \leqslant \Big(1 + \frac{n\pi}{2}\sqrt{2-c_1}\Big) \sup_{|x-y|\leqslant \frac{1}{n}} \big|f(x)-f(y)\big|\,.
$$

To conclude the theorem, we need to show that the number $n\sqrt{2-c_1}$ can be made bounded by choosing p properly. Nevertheless, let

$$
\begin{aligned}
p(x) &= c\Big|\sum_{k=0}^{n} \sin\frac{(k+1)\pi}{n+2} e^{ikx}\Big|^2 = c\sum_{k=0}^{n}\sum_{\ell=0}^{n} \sin\frac{(k+1)\pi}{n+2}\frac{(\ell+1)\pi}{n+2} e^{i(k-\ell)x} \\
&= c\sum_{k=0}^{n}\sin^2\frac{(k+1)\pi}{n+2} + 2c\sum_{\substack{k,\ell=0\\k>\ell}}^{n} \sin\frac{(k+1)\pi}{n+2}\frac{(\ell+1)\pi}{n+2}\cos(k-\ell)x
\end{aligned}
$$

and choose c so that $p(x) = 1 + c_1\cos x + \cdots + c_n \cos nx$. Then

$$
\begin{aligned}
c^{-1} &= \sum_{k=0}^{n}\sin^2\frac{(k+1)\pi}{n+2} = \frac{1}{2}\sum_{k=0}^{n}\Big[1-\cos\frac{2(k+1)\pi}{n+2}\Big] \\
&= \frac{n+1}{2} - \frac{\sin\frac{(2n+3)\pi}{n+2} - \sin\frac{\pi}{n+2}}{4\sin\frac{\pi}{n+2}} = \frac{n+2}{2}\,,
\end{aligned}
$$

and

$$\begin{aligned}
c_1 &= 2c\sum_{k=1}^{n} \sin\frac{(k+1)\pi}{n+2}\sin\frac{k\pi}{n+2} = c\sum_{k=1}^{n}\Big[\cos\frac{\pi}{n+2} - \cos\frac{(2k+1)\pi}{n+2}\Big] \\
&= c\Big[n\cos\frac{\pi}{n+2} - \frac{\sin\frac{(2n+2)\pi}{n+2} - \sin\frac{2\pi}{n+2}}{2\sin\frac{\pi}{n+2}}\Big] \\
&= c\Big[n\cos\frac{\pi}{n+2} + \frac{\sin\frac{2\pi}{n+2}}{\sin\frac{\pi}{n+2}}\Big] \\
&= c(n+2)\cos\frac{\pi}{n+2} = 2\cos\frac{\pi}{n+2}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
n\sqrt{2-c_1} &= n\Big(2-2\cos\frac{\pi}{n+2}\Big)^{\frac{1}{2}} = 2n\sin\frac{\pi}{2(n+2)} \\
&= 2(n+2)\sin\frac{\pi}{2(n+2)} - 4\sin\frac{\pi}{2(n+2)} \\
&= \pi\frac{2(n+2)}{\pi}\sin\frac{\pi}{2(n+2)} - 4\sin\frac{\pi}{2(n+2)}
\end{aligned}$$

which is bounded by $\pi + 4$. □

Finally, since $\lim\limits_{n\to\infty} n^{-\alpha}\log n = 0$ for all $\alpha \in (0,1]$, we conclude the following

Theorem 8.21. *For all $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$ with $\alpha \in (0,1]$, $s_n(f,\cdot) = D_n \star f$ converges to f uniformly as $n \to \infty$.*

Remark 8.22. The converse of Theorem 8.20 is the Bernstein theorem which states that if f is a 2π-periodic function such that for some constant C (independent of n) and $\alpha \in (0,1)$,

$$\inf_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_{L^\infty(\mathbb{T})} \leqslant Cn^{-\alpha} \tag{8.9}$$

for all $n \in \mathbb{N}$, then $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$. In other words, (8.9) is an equivalent condition to the Hölder continuity with exponent α of 2π-periodic continuous functions. One way of proving the Bernstein theorem can be found in Exercise Problem 8.4.

8.2.2 Jump discontinuity and Gibbs phenomenon

In this section, we study the convergence of the Fourier series of functions with jump discontinuities. We show that the Fourier series evaluated at the jump discontinuity converges to the average of the limits from the left and the right. Moreover, the convergence of the Fourier series is never uniform in the domain excluding these jump discontinuities due to the famous Gibbs phenomenon: near the jump discontinuity the maximum difference between the limit of the Fourier series and the function itself is at least 8% of the jump. To be more precise, we have the following

Theorem 8.23. *Let $f : \mathbb{R} \to \mathbb{R}$ be a piecewise continuously differentiable function which is periodic with some period $L > 0$. Suppose that at some point x_0 the limit from the left $f(x_0^-)$ and the limit from the right $f(x_0^+)$ of the function f exist and differ by a non-zero gap a:*

$$f(x_0^+) - f(x_0^-) = a \neq 0,$$

then there exists a generic constant $c > 0$, independent of f, x_0 and L (in fact, $c = \dfrac{1}{\pi}\displaystyle\int_0^\pi \frac{\sin x}{x}\,dx - \frac{1}{2} \approx 0.089490$), such that

$$\lim_{n\to\infty} s_n\Big(f, x_0 + \frac{L}{2n}\Big) = f(x_0^+) + ca\,, \tag{8.10a}$$

$$\lim_{n\to\infty} s_n\Big(f, x_0 - \frac{L}{2n}\Big) = f(x_0^-) - ca\,. \tag{8.10b}$$

Moreover,

$$\lim_{n\to\infty} s_n(f, x_0) = \frac{f(x_0^+) + f(x_0^-)}{2}\,. \tag{8.11}$$

Proof. Without loss of generality, we may assume that $x_0 = 0$ is the only discontinuity of f, $f(0) = \dfrac{f(0^+) + f(0^-)}{2}$, and $L = 2\pi$. Let g be a discontinuous function defined by

$$g(x) = \begin{cases} \dfrac{a}{2\pi}(x+\pi) & \text{if } -\pi \leqslant x < 0\,, \\ \quad 0 & \text{if } x = 0\,, \\ \dfrac{a}{2\pi}(x-\pi) & \text{if } 0 < x \leqslant \pi\,. \end{cases}$$

Then $F = f + g$ is Lipchitz continuous on $\mathbb{T}$, thus by Theorem 8.15,

$$\begin{aligned} \frac{f(0^+) + f(0^-)}{2} &= F(0) = \lim_{n\to\infty} s_n(F, 0) = \lim_{n\to\infty} s_n(f, 0) + \lim_{n\to\infty} s_n(g, 0) \\ &= \lim_{n\to\infty} s_n(f, 0)\,. \end{aligned}$$

This proves (8.11).

By $\widehat{g}(k) = \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\pi}^{\pi} g(x)\, e^{-ikx} dx = \frac{ia}{\sqrt{2\pi}k}$ if $k \neq 0$, and $\widehat{g}(0) = 0$, we find that

$s_n(g,x) = \sum_{k=-n}^{n} \frac{\widehat{g}(k)}{\sqrt{2\pi}} e^{ikx} = -\sum_{k=1}^{n} \frac{a}{\pi k} \sin(kx)$; thus

$$s_n\big(g, \frac{\pi}{n}\big) = -\sum_{k=1}^{n} \frac{a}{\pi k} \sin\frac{k\pi}{n} = -\frac{a}{\pi} \sum_{k=1}^{n} \frac{n}{k\pi} \sin\frac{k\pi}{n}\frac{\pi}{n} \to -\frac{a}{\pi} \int_0^{\pi} \frac{\sin x}{x} dx\,.$$

Therefore, by the continuity of F,

$$\begin{aligned}\frac{f(0^+)+f(0^-)}{2} &= \lim_{n\to\infty} F\big(\frac{\pi}{n}\big) = \lim_{n\to\infty} s_n\big(f, \frac{\pi}{n}\big) + \lim_{n\to\infty} s_n\big(g, \frac{\pi}{n}\big)\\ &= \lim_{n\to\infty} s_n\big(f, \frac{\pi}{n}\big) - \frac{a}{\pi}\int_0^{\pi} \frac{\sin x}{x} dx\,,\end{aligned}$$

and (8.10a) follows from $\dfrac{f(0^+)+f(0^-)}{2} + \dfrac{a}{2} = f(0^+)$. (8.10b) can be proved in the same fashion, and is left as an exercise. □

8.3 The Sobolev Space $H^s(\mathbb{T})$

Definition 8.24. For $s > 0$ (not necessary an integer), the Sobolev space $H^s(\mathbb{T})$ consists of all functions $f \in L^2(\mathbb{T})$ such that

$$\sum_{k=-\infty}^{\infty} |k|^{2s} |\widehat{f}(k)|^2 < \infty\,.$$

If $f, g \in H^s(\mathbb{T})$, then the $H^s(\mathbb{T})$-inner product of f and g is defined by

$$(f,g)_{H^s(\mathbb{T})} = \sum_{n=-\infty}^{\infty} (1+|k|^2)^s \widehat{f}(k)\, \overline{\widehat{g}(k)}$$

which induces the $H^s(\mathbb{T})$-norm as

$$\|f\|^2_{H^s(\mathbb{T})} = \sum_{k=-\infty}^{\infty} (1+|k|^2)^s |\widehat{f}(k)|^2\,.$$

Example 8.25. Consider the heavyside function H defined by

$$H(x) = \begin{cases} 1 & \text{if } 0 \leqslant x \leqslant \pi\,, \\ 0 & \text{if } -\pi \leqslant x < 0\,. \end{cases}$$

The Fourier coefficients for H is

$$\widehat{H}(k) = \frac{1}{\sqrt{2\pi}} \int_0^{\pi} e^{-ikx} dx = \begin{cases} \sqrt{\dfrac{\pi}{2}} & \text{if } k = 0, \\ 0 & \text{if } k \text{ is even, } k \neq 0, \\ \sqrt{\dfrac{2}{\pi}}\,\dfrac{1}{ik} & \text{if } k \text{ is odd,} \end{cases}$$

hence $H \in H^s(\mathbb{T})$ if $s < \dfrac{1}{2}$.

Proposition 8.26. *If $0 < s < r$, then $H^r(\mathbb{T}) \subseteq H^s(\mathbb{T})$.*

Proposition 8.27. *If $k \in \mathbb{N}$, then $\mathscr{C}^k(\mathbb{T}) \subseteq H^k(\mathbb{T})$, where $\mathscr{C}^k(\mathbb{T})$ consists of all k-times continuously differentiable (2π-periodic) functions.*

Theorem 8.28. *Let $0 < r < t < \infty$, and $s = \alpha r + (1-\alpha)t$ for some $\alpha \in (0,1)$. Then*

$$\|u\|_{H^s(\mathbb{T})} \leqslant \|u\|^{\alpha}_{H^r(\mathbb{T})}\|u\|^{1-\alpha}_{H^t(\mathbb{T})}\,. \tag{8.12}$$

Proof. By definition,

$$\begin{aligned}\|u\|^2_{H^s(\mathbb{T})} &= \sum_{k=-\infty}^{\infty}(1+|k|^2)^s|\widehat{u}(k)|^2 \\ &= \sum_{k=-\infty}^{\infty}(1+|k|^2)^{\alpha r}|\widehat{u}(k)|^{2\alpha}(1+|k|^2)^{(1-\alpha)t}|\widehat{u}(k)|^{2(1-\alpha)}\,.\end{aligned}$$

Noting that $\dfrac{1}{\alpha^{-1}} + \dfrac{1}{(1-\alpha)^{-1}} = 1$, by the Hölder inequality we find that

$$\begin{aligned}&\sum_{k=-\infty}^{\infty}(1+|k|^2)^{\alpha r}|\widehat{u}(k)|^{2\alpha}(1+|k|^2)^{(1-\alpha)t}|\widehat{u}(k)|^{2(1-\alpha)} \\ &\qquad \leqslant \Big[\sum_{k=-\infty}^{\infty}(1+|k|^2)^r|\widehat{u}(k)|^2\Big]^{\alpha}\Big[\sum_{k=-\infty}^{\infty}(1+|k|^2)^t|\widehat{u}(k)|^2\Big]^{1-\alpha}\end{aligned}$$

which leads to (8.12). □

Theorem 8.29 (Sobolev embedding, the simplest version)**.** *If $f \in H^s(\mathbb{T})$ for some $s > \dfrac{1}{2}$, then there exists $\widetilde{f} \in \mathscr{C}(\mathbb{T})$ so that $f = \widetilde{f}$ almost everywhere. Moreover, there exists a constant $C_s > 0$ such that*

$$\|f\|_{L^\infty(\mathbb{T})} \leqslant C_s\|f\|_{H^s(\mathbb{T})} \qquad \forall\, f \in H^s(\mathbb{T})\,. \tag{8.13}$$

Proof. Let $s_n(f,x)$ be the partial sum of the Fourier series of f defined as before. Then for $n \geqslant m$,

$$\begin{aligned}\big|s_n(f,x) - s_m(f,x)\big| &= \frac{1}{\sqrt{2\pi}}\Big|\sum_{m<|k|\leqslant n}\widehat{f}(k)e^{ikx}\Big| \leqslant \frac{1}{\sqrt{2\pi}}\sum_{m<|k|\leqslant n}|\widehat{f}(k)| \\ &\leqslant \frac{1}{\sqrt{2\pi}}\Big[\sum_{m<|k|\leqslant n}(1+|k|^2)^s|\widehat{f}(k)|^2\Big]^{1/2}\Big[\sum_{m<|k|\leqslant n}\frac{1}{(1+|k|^2)^s}\Big]^{1/2}.\end{aligned}$$

Therefore, $\left\|s_n(f,\cdot) - s_m(f,\cdot)\right\|_{L^\infty(\mathbb{T})} \to 0$ as $n, m \to \infty$, which implies that $s_n(f,\cdot)$ converges uniformly; hence $\widetilde{f} \equiv \lim\limits_{n\to\infty} s_n(f,\cdot)$ is continuous.

The constant C_s in (8.13) can be chosen as $\dfrac{1}{\sqrt{2\pi}}\Big[\sum\limits_{k=-\infty}^{\infty} \dfrac{1}{(1+|k|^2)^s}\Big]^{\frac{1}{2}}$. ▫

8.3.1 Characterization of $H^1(\mathbb{T})$

Definition 8.24 gives a quantitative way of describing functions in $H^s(\mathbb{T})$. In this section, a qualitative point of view of $H^1(\mathbb{T})$ is provided based on the Hahn-Banach theorem from functional analysis. Roughly speaking, a function $f \in H^1(\mathbb{T})$ has *weak derivatives* belonging to $L^2(\mathbb{T})$ and satisfies the integration by parts formula. We start from stating the following

Theorem 8.30 (Hahn-Banach)**.** *If Y is a linear subspace of a normed linear space X and $T : Y \to \mathbb{R}$ is a bounded linear functional on Y with $\|T\| = M$, then there is a bounded linear functional $\widetilde{T} : X \to \mathbb{R}$ on X such that $\widetilde{T}$ restricted to Y is equal to T and $\|\widetilde{T}\| = M$.*

In other words, a bounded linear functional on a normed linear space can be extended to a bounded linear functional on a larger space without changing the size of its norm.

Let $f \in H^1(\mathbb{T})$ and $\varphi \in \mathscr{C}^1(\mathbb{T})$. We define a (bounded) linear functional T_f on $\mathscr{C}^1(\mathbb{T})$ by

$$T_f(\varphi) = \int_{\mathbb{T}} f(x)\varphi'(x)\,dx\,.$$

The goal is to extend T_f to a bounded linear functional $\widetilde{T}_f$ defined on $L^2(\mathbb{T})$. Since the application of the Hahn-Banach theorem requires that the range of the linear function to be real, in the following discussion we will always assume that f and φ are real-valued functions.

Since $\varphi' \in \mathscr{C}(\mathbb{T}) \subseteq L^2(\mathbb{T})$, we can compute $\widehat{\varphi}'$ and obtain that $\widehat{\varphi}'(k) = ik\widehat{\varphi}(k)$. Therefore,

$$\int_{\mathbb{T}} f(x)\varphi'(x)\,dx = (f,\varphi')_{L^2(\mathbb{T})} = \sum_{k=-\infty}^{\infty} \widehat{f}(k)\overline{ik\widehat{\varphi}(k)}\,;$$

hence by Hölder's inequality,

$$\Big|\int_{\mathbb{T}} f(x)\varphi'(x)\,dx\Big| \leqslant \sum_{k=-\infty}^{\infty} |k||\widehat{f}(k)||\widehat{\varphi}(k)| \leqslant \|f\|_{H^1(\mathbb{T})}\|\varphi\|_{L^2(\mathbb{T})}\,.$$

The computation above shows that if $f \in H^1(\mathbb{T})$, T_f is a bounded linear functional (on a subspace of $L^2(\mathbb{T})$). By the Hahn-Banach theorem, T_f can be extended to a bounded linear functional $\widetilde{T}_f : L^2(\mathbb{T}) \to \mathbb{R}$. By the Riesz representation theorem, there is a function $g \in L^2(\mathbb{T})$ such that

$$\widetilde{T}_f(\varphi) = (\varphi, g)_{L^2(\mathbb{T})} = \int_{\mathbb{T}} \varphi(x) g(x)\, dx \qquad \forall\, \varphi \in L^2(\mathbb{T})\,.$$

In particular, for $\varphi \in \mathscr{C}^1(\mathbb{T})$,

$$\int_{\mathbb{T}} \varphi(x) g(x)\, dx = \widetilde{T}_f(\varphi) = T_f(\varphi) = \int_{\mathbb{T}} f(x)\varphi'(x)\, dx\,.$$

The function $h = -g$ is called the *weak derivative* of f, and usually is denoted by f' as well. The reason for calling h the weak derivative of f is that if $f \in \mathscr{C}^1(\mathbb{T})$, then

$$-\int_{\mathbb{T}} f'(x)\varphi(x)\, dx = \int_{\mathbb{T}} f(x)\varphi'(x)\, dx\,, \tag{8.14}$$

so h is indeed the derivative of f. Note that $g \in L^2(\mathbb{T})$ is "the same as" saying that $f' \in L^2(\mathbb{T})$. In fact, we have the following

Theorem 8.31. *A function f belongs to $H^1(\mathbb{T})$ if and only if $f \in L^2(\mathbb{T})$ and there exists a function $g \in L^2(\mathbb{T})$, called the weak derivative of f, such that*

$$-\int_{\mathbb{T}} g(x)\varphi(x)dx = \int_{\mathbb{T}} f(x)\varphi'(x)dx \qquad \forall\, \varphi \in \mathscr{C}^1(\mathbb{T})\,. \tag{8.15}$$

In other words, the space $H^1(\mathbb{T})$ consists of all functions in $L^2(\mathbb{T})$ possessing weak derivatives in $L^2(\mathbb{T})$.

Proof. It remains to show that $f \in H^1(\mathbb{T})$ is a necessary condition. Suppose that $f \in L^2(\mathbb{T})$ and there exists $g \in L^2(\mathbb{T})$ satisfying (8.15). Then

$$\begin{aligned}\sum_{k=-\infty}^{\infty} \widehat{g}(k)\overline{\widehat{\varphi}(k)} &= (g,\varphi)_{L^2(\mathbb{T})} = \int_{\mathbb{T}} g(x)\overline{\varphi(x)}\, dx = \int_{\mathbb{T}} f(x)\overline{\varphi'(x)}\, dx \\ &= \sum_{k=-\infty}^{\infty} \widehat{f}(k)\overline{\widehat{\varphi'}(k)} = -\sum_{k=-\infty}^{\infty} ik\widehat{f}(k)\overline{\widehat{\varphi}(k)}\,.\end{aligned}$$

This implies that $\widehat{g}(k) = -ik\widehat{f}(k)$; thus

$$\sum_{k=-\infty}^{\infty} |k|^2|\widehat{f}(k)|^2 = \sum_{k=-\infty}^{\infty} |\widehat{g}(k)|^2 = \|g\|^2_{L^2(\mathbb{T})} < \infty\,.$$

□

Corollary 8.32. *Let $f \in H^1(\mathbb{T})$. Then $\|f\|^2_{H^1(\mathbb{T})} = \|f\|^2_{L^2(\mathbb{T})} + \|f'\|^2_{L^2(\mathbb{T})}$, where f' is the weak derivative of f.*

Remark 8.33. The proof of Theorem 8.31 also implies that the Fourier coefficients $\widehat{f'}(k)$ of the weak derivative f' is $ik\widehat{f}(k)$ since $f' = -g$. Therefore, if $\dfrac{d}{dx}$ denotes the weak differentiation operator

$$\frac{d}{dx}\Big[\frac{1}{\sqrt{2\pi}}\sum_{k=-\infty}^{\infty}\widehat{f}(k)e^{ikx}\Big] = \frac{1}{\sqrt{2\pi}}\sum_{k=-\infty}^{\infty} ik\widehat{f}(k)e^{ikx} = \frac{1}{\sqrt{2\pi}}\sum_{k=-\infty}^{\infty}\frac{d}{dx}\Big[\widehat{f}(k)e^{ikx}\Big];$$

thus $\dfrac{d}{dx}$ commutes with the infinite sum (in which the convergence of the infinite sum is understood in the L^2-sense).

Remark 8.34. Let $f \in H^1(\mathbb{T})$, then for any given $\epsilon > 0$, there is a function $f_\epsilon \in \mathscr{C}^1(\mathbb{T})$ such that

$$\|f - f_\epsilon\|_{H^1(\mathbb{T})} < \epsilon\,;$$

that is, $H^1(\mathbb{T})$ is the completion of the normed space $\big(\mathscr{C}^1(\mathbb{T}), \|\cdot\|_{H^1(\mathbb{T})}\big)$.

Remark 8.35. The Hahn-Banach theorem does not guarantee the uniqueness of the extension $\widetilde{T}_f$. Therefore, there might be two extensions $\widetilde{T}_{f_1}$ and $\widetilde{T}_{f_2}$ mapping from $L^2(\mathbb{T})$ to $\mathbb{R}$ that equal T_f on $\mathscr{C}^1(\mathbb{T})$. Suppose that g_1 and g_2 are the corresponding representations of $\widetilde{T}_{f_1}$ and $\widetilde{T}_{f_2}$. By definition,

$$\int_{\mathbb{T}} \varphi(x)g_1(x)\,dx = \widetilde{T}_{f_1}(\varphi) = T_f(\varphi) = \widetilde{T}_{f_2}(\varphi) = \int_{\mathbb{T}} \varphi(x)g_2(x)\,dx$$

for all $\varphi \in \mathscr{C}^1(\mathbb{T})$. Therefore, $g_1 = g_2$ a.e. in $L^2(\mathbb{T})$; thus the extension $\widetilde{T}_f$ is indeed unique. The key here is that $\mathscr{C}^1(\mathbb{T})$ is dense in $L^2(\mathbb{T})$.

Similarly, we have the following

Theorem 8.36. *A function $f \in H^k(\mathbb{T})$ if and only if for each $0 \leqslant j < k$, $f^{(j)} \equiv \dfrac{d^j f}{dx^j}$ is weakly differentiable with weak derivative $f^{(j+1)}$ belonging to $L^2(\mathbb{T})$. Moreover, there are positive constants C_1 and C_2 such that*

$$C_1\|f\|_{H^k(\mathbb{T})} \leqslant \sum_{j=0}^{k}\|f^{(j)}\|_{L^2(\mathbb{T})} \leqslant C_2\|f\|_{H^k(\mathbb{T})}\,.$$

8.3.2 The space $H^k(0,\pi)$

Motivated by Theorem 8.31 and Corollary 8.32, we look for a qualitative description of the H^k-space using the language of weak derivatives.

Definition 8.37. A function $u \in L^1_{\rm loc}(0,\pi)$ is said to be weakly differentiable if there exists a function $g \in L^1_{\rm loc}(0,\pi)$ such that

$$-\int_0^\pi g(x)\varphi(x)dx = \int_0^\pi f(x)\varphi'(x)dx \qquad \forall\, \varphi \in \mathscr{C}^1_c((0,\pi))\,. \tag{8.16}$$

The function g is called the weak derivative of f, and is denoted by f'.

We note that in the definition above, the functional framework $L^1_{\rm loc}(0,\pi)$ is chosen so that the integrals in (8.16) make sense. Moreover, the test function φ in (8.16) is compactly supported in $(0,\pi)$; that is, $\mathrm{spt}(\varphi) \subseteq (0,\pi)$.

Definition 8.38. The space $H^k(0,\pi)$ consists of all functions $f \in L^2(\mathbb{T})$ possessing square integrable weak derivatives $f^{(j)} \equiv \dfrac{d^j f}{dx^j}$ for all $0 \leqslant j \leqslant k$; that is,

$$H^k(0,\pi) \equiv \Big\{ f \in L^2(0,\pi) \,\Big|\, \int_0^\pi |f^{(j)}(x)|^2 dx < \infty \quad \forall\, j = 0,1,\cdots,k \Big\}\,.$$

The space $H^k(0,\pi)$ is equipped with a norm given by

$$\|f\|_{H^k(0,\pi)} = \Big[\sum_{j=0}^k \|f^{(j)}\|^2_{L^2(0,\pi)}\Big]^{1/2}$$

which is induced by the inner product

$$(f,g)_{H^k(0,\pi)} = \sum_{j=0}^k \big(f^{(j)}, g^{(j)}\big)_{L^2(0,\pi)} \qquad \forall\, f,g \in H^k(0,\pi)\,.$$

Remark 8.39. As mentioned in Section 8.1, $\Big\{\sqrt{\dfrac{1}{\pi}}, \sqrt{\dfrac{2}{\pi}}\cos kx\Big\}_{k=1}^\infty$ is an orthonormal basis of $L^2(0,\pi)$. Let $w_0 = \sqrt{\dfrac{1}{\pi}}$ and $w_k = \sqrt{\dfrac{2}{\pi}}\dfrac{\cos kx}{\sqrt{1+k^2}}$. Then $\{w_k\}_{k=0}^\infty$ is an orthonormal basis of $H^1(0,\pi)$ (see Exercise Problem 8.5). Expand $\sin x$ in terms of this H^1-basis, we obtain that

$$\sin x = \frac{2}{\pi} - \sum_{k=1}^\infty \frac{4}{\pi(4k^2-1)}\cos 2kx = \frac{2}{\pi} - \lim_{n\to\infty}\sum_{k=1}^n \frac{4}{\pi(4k^2-1)}\cos 2kx\,,$$

where the limit is taken in the H^1-topology, or equivalently,

$$\lim_{n\to\infty} \left\| \sin x - \frac{2}{\pi} + \sum_{k=1}^{n} \frac{4}{\pi(4k^2-1)} \cos 2kx \right\|_{H^1(0,\pi)} = 0\,.$$

Note that w_k has the property that the derivative of w_k, $\dfrac{\partial w_k}{\partial x}$, vanishes at the boundary points $x = 0$ and $x = \pi$ for all k, but the derivative of $\sin x$ at the boundary points does not vanish.

8.4 1-Dimensional Heat Equations with Periodic Boundary Condition

In this section, we consider the heat equation:

$$u_t(x,t) - u_{xx}(x,t) = f(x,t) \qquad \text{for all } (x,t) \in (0,2\pi) \times (0,T), \tag{8.17a}$$

$$u(0,t) = u(2\pi,t) \qquad \text{for all } t \in (0,T), \tag{8.17b}$$

$$u(x,0) = g(x) \qquad \text{for all } x \in (0,2\pi). \tag{8.17c}$$

Condition (8.17b) is called *the periodic boundary condition*, which enables us to treat solutions $u(\cdot,t)$ as a periodic function defined on $\mathbb{R}$ for all $t \in [0,T]$. We assume that $g \in H^2(\mathbb{T})$, $\max\limits_{t\in[0,T]} \|f(\cdot,t)\|_{L^2(\mathbb{T})} < \infty$, and

$$\int_0^T \|f(\cdot,t)\|^2_{H^1(\mathbb{T})} dt \equiv \int_0^T \int_{\mathbb{T}} \big(|f(x,t)|^2 + |f_x(x,t)|^2\big) dxdt < \infty\,.$$

8.4.1 Formal approaches

Assume that for all $t \in [0,T]$, $u(\cdot,t) \in L^2(\mathbb{T})$. Therefore, if $d_n(t)$ is the Fourier coefficient of $u(\cdot,t)$, we can express $u(x,t)$ as

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} d_k(t) e^{ikx} = \sum_{k=-\infty}^{\infty} d_k(t)\, \mathrm{e}_k(x)\,.$$

Because of (8.17c), we must have $d_k(0) = \widehat{g}(k)$. Moreover, for almost all $t \in [0,T]$, $f(\cdot,t) \in L^2(\mathbb{T})$. Therefore,

$$f(x,t) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \widehat{f}_k(t) e^{ikx} \qquad \text{for almost all } t \in (0,T)\,,$$

where $\widehat{f}_k(t)$ is the Fourier coefficients defined by

$$\widehat{f}_k(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} f(x,t) e^{-ikx} dx\,.$$

Suppose that we can switch the order of the differentiation and the summation, then

$$u_t(x,t) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} d_k'(t) e^{ikx}\,, \quad u_{xx}(x,t) = -\frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} k^2 d_k(t) e^{ikx}\,;$$

thus by (8.17a), for almost all $t \in [0,T]$,

$$\frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \Big[d_k'(t) + k^2 d_k(t) - \widehat{f}_k(t) \Big] e^{ikx} = 0\,. \tag{8.18}$$

Since $\{\mathrm{e}_k\}_{k=-\infty}^{\infty}$ is maximal, we find that $d_k(t)$ solves the ODE

$$d_k'(t) + k^2 d_k(t) = \widehat{f}_k(t)\,. \tag{8.19}$$

Together with the initial condition $d_k(0) = \widehat{g}(k)$, we find that

$$d_k(t) = e^{-k^2 t} \widehat{g}(k) + \int_0^t \widehat{f}_k(s) e^{-k^2(t-s)} ds$$

which implies that a solution $u(x,t)$ can be written as

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \Big[e^{-k^2 t} \widehat{g}(k) + \int_0^t \widehat{f}_k(s) e^{-k^2(t-s)} ds \Big] e^{ikx}\,. \tag{8.20}$$

8.4.2 Rigorous approaches

Before proceeding, we state a very important theorem in the study of differential equations.

Theorem 8.40 (The Gronwall inequality)**.** *Let $x(t)$ be a non-negative, continuous function on the interval $[0,T]$. If $x(t)$ satisfies $x'(t) \leqslant M + Cx(t)$ for all $t \in [0,T]$, then*

$$x(t) \leqslant e^{Ct} x(0) + \frac{M}{C}(e^{Ct} - 1) \qquad \forall\, t \in [0,T]\,. \tag{8.21}$$

Proof. Multiplying both sides of the differential inequality by the integrating factor e^{-Ct}, we find that

$$\frac{d}{dt}\Big[e^{-Ct} x(t) \Big] \leqslant M e^{-Ct}\,.$$

The desiblack inequality is then obtained by integrating the inequality above in time from 0 to t for some $t \in [0,T]$, and the detail is left to the readers. □

Corollary 8.41. *Let $y(t)$ be a non-negative, integrable function on the interval $[0,T]$. If $y(t)$ satisfies*

$$y(t) \leqslant M + C\int_0^t y(s)\,ds \qquad \forall\, t \in [0,T]\,, \tag{8.22}$$

then

$$y(t) \leqslant Me^{Ct} \qquad \forall\, t \in [0,T]\,.$$

Proof. Let $x(t) = \displaystyle\int_0^t y(s)\,ds$ and apply Theorem 8.40. □

Let $f_n(x,t) = \dfrac{1}{\sqrt{2\pi}}\displaystyle\sum_{k=-n}^{n}\widehat{f}_k(t)e^{ikx}$ and $g_n(x) = \dfrac{1}{\sqrt{2\pi}}\displaystyle\sum_{k=-n}^{n}\widehat{g}(k)e^{ikx}$. We look for a solution $u_n(x,t)$ to

$$u_{nt}(x,t) - u_{nxx}(x,t) = f_n(x,t) \qquad \text{for all } (x,t) \in (0,2\pi)\times(0,T), \tag{8.23a}$$

$$u_n(0,t) = u_n(2\pi,t) \qquad \text{for all } t \in (0,T), \tag{8.23b}$$

$$u_n(x,0) = g_n(x) \qquad \text{for all } x \in (0,2\pi). \tag{8.23c}$$

The same procedure as the formal approach implies that

$$u_n(x,t) = \frac{1}{\sqrt{2\pi}}\sum_{k=-n}^{n}\Big[e^{-k^2t}\widehat{g}(k) + \int_0^t \widehat{f}_k(s)e^{-k^2(t-s)}ds\Big]e^{ikx}$$

is a solution to (8.23). Our goal is to show that u_n converges to the solution of (8.17).

Energy estimates

In order to show that u_n converges (in certain sense), we need to show that it is a Cauchy sequence. Define $v^{n,m} = u_n - u_m$, $g^{n,m} = g_n - g_m$ and $f^{n,m} = f_n - f_m$. Then $v^{n,m}$ satisfies

$$v_t^{n,m}(x,t) - v_{xx}^{n,m}(x,t) = f^{n,m}(x,t) \qquad \text{for all } (x,t) \in (0,2\pi)\times(0,T), \tag{8.24a}$$

$$v^{n,m}(0,t) = v^{n,m}(2\pi,t) \qquad \text{for all } t \in (0,T), \tag{8.24b}$$

$$v^{n,m}(x,0) = g^{n,m}(x) \qquad \text{for all } x \in (0,2\pi). \tag{8.24c}$$

Multiplying (8.24a) by $v_{xxxx}^{n,m}(x,t)$ and integrating over $\mathbb{T}$,

$$\int_{\mathbb{T}}\Big[v_t^{n,m}(x,t) - v_{xx}^{n,m}(x,t)\Big]v_{xxxx}^{n,m}(x,t)\,dx = \int_{\mathbb{T}} f^{n,m}(x,t)v_{xxxx}^{n,m}(x,t)\,dx\,. \tag{8.25}$$

Integrating by parts in x, we find that

$$\begin{aligned}\int_{\mathbb{T}} v_t^{n,m}(x,t)v_{xxxx}^{n,m}(x,t)\,dx &= \int_{\mathbb{T}} v_{xxt}^{n,m}(x,t)v_{xx}^{n,m}(x,t)\,dx \\ &= \int_{\mathbb{T}} \frac{1}{2}\frac{\partial}{\partial t}|v_{xx}^{n,m}(x,t)|^2 dx = \frac{1}{2}\frac{d}{dt}\|v_{xx}^{n,m}(\cdot,t)\|_{L^2(\mathbb{T})}^2\end{aligned}$$

and

$$-\int_{\mathbb{T}} v_{xx}^{n,m}(x,t)u_{nxxxx}(x,t)\,dx = \int_{\mathbb{T}} v_{xxx}^{n,m}(x,t)v_{xxx}^{n,m}(x,t)\,dx = \|v_{xxx}^{n,m}(\cdot,t)\|_{L^2(\mathbb{T})}^2 .$$

Moreover, by Hölder's and Young's inequality,

$$\begin{aligned}&\int_{\mathbb{T}} f^{n,m}(x,s)v_{xxxx}^{n,m}(x,s)\,dx = -\int_{\mathbb{T}} f_x^{n,m}(x,s)v_{xxx}^{n,m}(x,s)\,dx \\ &\quad\leqslant \|f_x^{n,m}(\cdot,s)\|_{L^2(\mathbb{T})}\|v_{xxx}^{n,m}(\cdot,s)\|_{L^2(\mathbb{T})} \leqslant \frac{1}{2}\Big[\|f_x^{n,m}(\cdot,s)\|_{L^2(\mathbb{T})}^2 + \|v_{xxx}^{n,m}(\cdot,s)\|_{L^2(\mathbb{T})}^2\Big] .\end{aligned}$$

As a consequence, (8.25) implies that

$$\frac{d}{dt}\|v_{xx}^{n,m}(\cdot,t)\|_{L^2(\mathbb{T})}^2 + \|v_{xxx}^{n,m}(\cdot,t)\|_{L^2(\mathbb{T})}^2 \leqslant \|f_x^{n,m}(\cdot,t)\|_{L^2(\mathbb{T})}^2$$

and integrating in t over the time interval $(0,t)$ further implies that

$$\begin{aligned}&\|v_{xx}^{n,m}(\cdot,t)\|_{L^2(\mathbb{T})}^2 + \int_0^t \|v_{xxx}^{n,m}(\cdot,t)\|_{L^2(\mathbb{T})}^2 \\ &\qquad\qquad \leqslant \|g_{xx}^{n,m}\|_{L^2(\mathbb{T})}^2 + \int_0^t \|f_x^{n,m}(\cdot,s)\|_{L^2(\mathbb{T})}^2\,ds .\end{aligned} \tag{8.26}$$

Similarly, multiplying (8.24a) by $v^{n,m}$ or $v_{xx}^{n,m}$ and then integrating over $\mathbb{T}$, we obtain that

$$\begin{aligned}&\|v_x^{n,m}(\cdot,t)\|_{L^2(\mathbb{T})}^2 + \int_0^t \|v_{xx}^{n,m}(\cdot,t)\|_{L^2(\mathbb{T})}^2 \\ &\qquad\qquad \leqslant \|g_x^{n,m}\|_{L^2(\mathbb{T})}^2 + \int_0^t \|f^{n,m}(\cdot,s)\|_{L^2(\mathbb{T})}^2\,ds .\end{aligned} \tag{8.27}$$

and

$$\begin{aligned}&\|v^{n,m}(\cdot,t)\|_{L^2(\mathbb{T})}^2 + 2\int_0^t \|v_x^{n,m}(\cdot,t)\|_{L^2(\mathbb{T})}^2 \leqslant \|g^{n,m}\|_{H^2(\mathbb{T})}^2 \\ &\qquad + \int_0^t \|f^{n,m}(\cdot,s)\|_{H^1(\mathbb{T})}^2\,ds + \int_0^t \|v^{n,m}(\cdot,s)\|_{L^2(\mathbb{T})}^2 ds .\end{aligned} \tag{8.28}$$

Summing (8.26), (8.27) and (8.28),

$$\|v^{n,m}(\cdot,t)\|^2_{H^2(\mathbb{T})} \leqslant 2\Big[\|g^{n,m}\|^2_{H^2(\mathbb{T})} + \int_0^t \|f^{n,m}(\cdot,s)\|^2_{H^1(\mathbb{T})}\,ds\Big] + \int_0^t \|v^{n,m}(\cdot,s)\|^2_{H^2(\mathbb{T})}\,ds\,;$$

thus the Gronwall inequality suggests that

$$\max_{t\in[0,T]} \|v^{n,m}(t)\|^2_{H^2(\mathbb{T})} \leqslant 2\Big[\|g^{n,m}\|^2_{H^2(\mathbb{T})} + \int_0^T \|f^{n,m}(\cdot,s)\|^2_{H^1(\mathbb{T})}\,ds\Big]e^T\,. \tag{8.29}$$

Since $g \in H^2(\mathbb{T})$ and $\displaystyle\int_0^T \|f(\cdot,t)\|^2_{H^1(\mathbb{T})}dt < \infty$,

$$\lim_{n,m\to\infty} \|g^{n,m}\|_{H^2(\mathbb{T})} = 0 \quad \text{and} \quad \lim_{n,m\to\infty} \int_0^T \|f^{n,m}(\cdot,s)\|^2_{H^1(\mathbb{T})}ds = 0\,.$$

As a consequence, u_n converges uniformly in $H^2(\mathbb{T})$; that is, there exists $u \in \mathscr{C}([0,T];H^2(\mathbb{T}))$ such that

$$\lim_{n,m\to\infty} \max_{t\in[0,T]} \|u_n(\cdot,t) - u(\cdot,t)\|_{H^2(\mathbb{T})} = 0\,.$$

We note that the equality above also suggests that $g_n = u_n(\cdot,0) \to u(\cdot,0)$ in $H^2(\mathbb{T})$; thus $u(x,0) = g(x)$. Moreover, because of the assumption that $\max\limits_{t\in[0,T]}\|f(\cdot,t)\|_{L^2(\Gamma)} < \infty$, $f_n \to f \in \mathscr{C}([0,T];H^2(\mathbb{T}))$ as well. Therefore,

$$\lim_{n,m\to\infty} \max_{t\in[0,T]} \|v^{n,m}_t(\cdot,t)\|_{L^2(\mathbb{T})} = \lim_{n,m\to\infty} \max_{t\in[0,T]} \|f^{n,m}(\cdot,t) - v^{n,m}_{xx}(\cdot,t)\|_{L^2(\mathbb{T})} = 0$$

which implies that u_{nt} converges uniformly in $L^2(\mathbb{T})$. Assume that $u_{nt} \to w$ in $\mathscr{C}([0,T];L^2(\mathbb{T}))$, we must have $u_t = w$ due to the uniform convergence. Moreover, similar to (8.29) we obtain that $u \in \mathscr{C}([0,T];H^2(\mathbb{T}))$ satisfies

$$\max_{t\in[0,T]} \|u(t)\|^2_{H^2(\mathbb{T})} \leqslant C\Big[\|g\|^2_{H^2(\mathbb{T})} + \int_0^T \|f(t)\|^2_{H^1(\mathbb{T})}dt\Big]\,. \tag{8.30}$$

So we conclude the following

Theorem 8.42. *Suppose that $f \in L^2(0,T;H^1(\mathbb{T})) \cap \mathscr{C}([0,T];L^2(\mathbb{T}))$, and $g \in H^2(\mathbb{T})$. Then*

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \Big[e^{-k^2t}\widehat{g}(k) + \int_0^t \widehat{f}_k(s)e^{-k^2(t-s)}ds\Big]e^{ikx} \tag{8.31}$$

solves (8.17) (*in the sense of weak spatial derivatives*). *Moreover, u belongs to the space $\mathscr{C}([0,T];H^2(\mathbb{T}))$ and satisfies* (8.30).

Remark 8.43. Let $f(x,t) = \big(\frac{2x}{\pi} - \frac{x^2}{\pi^2}\big) + \frac{2t}{\pi^2}$. Then $f \in L^2(0,T;H^1(\mathbb{T})) \cap \mathscr{C}([0,T];L^2(\mathbb{T}))$, and the function $u(x,t) = \big(\frac{2x}{\pi} - \frac{x^2}{\pi^2}\big)t$ satisfies

$$\begin{aligned} u_t(x,t) - u_{xx}(x,t) &= f(x,t) && \text{for all } (x,t) \in (0,2\pi)\times(0,\infty),\\ u(0,t) &= u(2\pi,t) && \text{for all } t>0,\\ u(x,0) &= 0 && \text{for all } x \in (0,2\pi) \end{aligned}$$

in the poinwise sense; however, $u(\cdot,t) \notin H^2(\mathbb{T})$ for all $t>0$. In fact, extending u periodically with period 2π, we have $u(0^+,t) = u(0^-,t) = 0$, and

$$u_x(0^+,t) = \frac{2}{\pi}t = -u_x(0^-,t)$$

which suggests that the "temperature" (the physical quantity that u presents) near by the origin increases in t while the "temperature" at the origin is always zero. Since we expect that the heat will flow into the origin so that the temperature at the origin also increases, this particular u is not a reasonable solution.

The solution given by (8.31) is

$$u(x,t) = \frac{1}{2\pi}\Big[\big(\frac{4\pi}{3}t + \frac{2}{\pi}t^2\big) + \frac{4}{\pi}\sum_{k\neq 0}\frac{(e^{-k^2t}-1)}{k^4}e^{ikx}\Big].$$

8.4.3 The special case $f = 0$

There are some good properties for the solution u to the heat equation (with periodic boundary condition) when there is no external forcing. We study these properties in this sub-section.

Maximum principle

Multiplying the heat equation $u_t - u_{xx} = 0$ by $pu|u|^{p-2}$ and then integrating over $\mathbb{T}$, we find that

$$\frac{d}{dt}\|u(\cdot,t)\|^p_{L^p(\mathbb{T})} + p(p-1)\int_{\mathbb{T}} u_x^2(x,t)u^{p-2}(x,t)\,dx = 0\,.$$

Integrating in time over the time interval $(0,t)$ then implies that

$$\big\|u(\cdot,t)\big\|_{L^p(\mathbb{T})} \leqslant \|g\|_{L^p(\mathbb{T})}\,.$$

Passing to the limit as $p \to \infty$, we find that

$$\big\|u(\cdot,t)\big\|_{L^\infty(\mathbb{T})} \leqslant \|g\|_{L^\infty(\mathbb{T})} \qquad \forall\, t > 0\,. \tag{8.32}$$

This inequality reads that the magnitude of the solution never exceeds the magnitude of the initial state, and is called the maximum principle for the heat equation (with periodic boundary condition).

8.4.4 Decay estimates

Under the assumption $f = 0$; that is, we are in the situation that there is no heat source in the environment, we expect that the solution/temperature will converges to the average $\bar{g} \equiv \fint_{\mathbb{T}} g(x)\,dx \equiv \dfrac{1}{2\pi}\displaystyle\int_{\mathbb{T}} g(x)\,dx$ as $t \to \infty$. We would like to study the convergence rate of $u - \bar{g}$.

By (8.31) with $f = 0$ and the Parseval identity,

$$\|u(\cdot,t) - \bar{g}\|^2_{L^2(\mathbb{T})} = \sum_{k\neq 0} e^{-2k^2 t}|\widehat{g}(k)|^2 \leqslant e^{-2t}\sum_{k\neq 0}|\widehat{g}(k)|^2 \leqslant e^{-2t}\|g\|^2_{L^2(\mathbb{T})}$$

which implies that

$$\big\|u(\cdot,t) - \bar{g}\big\|_{L^2(\mathbb{T})} \leqslant e^{-t}\|g\|_{L^2(\mathbb{T})} \qquad \forall\, t > 0\,.$$

Usually we are more interested in the case of $t \gg 1$. In such a case, we may evaluate $u - \bar{g}$ in $L^\infty(\mathbb{T})$ and obtain that

$$\begin{aligned}\big\|u(\cdot,t) - \bar{g}\big\|_{L^\infty(\mathbb{T})} &\leqslant \frac{1}{\sqrt{2\pi}}\sum_{k\neq 0} e^{-k^2 t}\big|\widehat{g}(k)\big| \leqslant \frac{1}{2\pi}\|g\|_{L^1(\mathbb{T})}\sum_{n\neq 0} e^{-k^2(t-1)}e^{-k^2} \\ &\overset{(t\,\geqslant\,1)}{\leqslant} \frac{1}{2\pi}e^{-(t-1)}\|g\|_{L^1(\mathbb{T})}\sum_{k\neq 0} e^{-k^2} \leqslant Ce^{-t}\|g\|_{L^1(\mathbb{T})}\,,\end{aligned}$$

where we use the fact that $\sup\limits_{k\in\mathbb{Z}}\big|\widehat{g}(k)\big| \leqslant \dfrac{1}{\sqrt{2\pi}}\|g\|_{L^1(\mathbb{T})}$ to conclude the inequality. Moreover, suppose that g is smooth so that u is smooth, then

$$\frac{\partial^\ell u}{\partial x^\ell}(x,t) = \sum_{k\neq 0} e^{-k^2 t}\widehat{g}(k)(ik)^\ell e^{ikx}\,;$$

thus for all $k \in \mathbb{N}$, similar argument implies that

$$\Big\|\frac{\partial^\ell u}{\partial x^\ell}(\cdot,t)\Big\|_{L^\infty(\mathbb{T})} \leqslant e^{-(t-1)}\|g\|_{L^1(\mathbb{T})}\sum_{k\neq 0} e^{-k^2}|k|^\ell \leqslant C_\ell e^{-t}\|g\|_{L^1(\mathbb{T})} \qquad \forall\, t \geqslant 1\,.$$

This proves the following

Theorem 8.44. *Let u be the solution to the heat equation* (8.17) *with $f = 0$ and $g \in L^1(\mathbb{T})$. Then the ℓ-th partial derivatives of $u - \bar{g}$ with respect to x decays exponentially to zero in the uniform sense.*

8.5 1-Dimensional Heat Equation with Dirichlet Boundary Condition

In this section, we consider the following initial-boundary value problem for the heat equation

$$u_t(x,t) - u_{xx}(x,t) = f(x,t) \qquad \text{for all } (x,t) \in (0,L) \times (0,T), \tag{8.33a}$$

$$u(0,t) = u(L,t) = 0 \qquad \text{for all } t \in (0,T)\,, \tag{8.33b}$$

$$u(x,0) = g(x) \qquad \text{for all } x \in [0,L]. \tag{8.33c}$$

Because of the boundary condition (8.33b), we use the orthonormal basis $\Big\{\sqrt{\frac{2}{L}} \sin \frac{k\pi x}{L}\Big\}_{k=1}^{\infty}$. Assume that $u(x,t) = \sum\limits_{k=1}^{\infty} d_k(t) \sin \frac{k\pi x}{L}$. Then

$$d_k'(t) + \frac{\pi^2 k^2}{L^2} d_k(t) = f_k(t) \equiv \frac{2}{L} \int_0^L f(x,t) \sin \frac{k\pi x}{L} dx \quad \forall\, t > 0$$

with initial condition

$$d_k(0) = \frac{2}{L} \int_0^L g(x) \sin \frac{k\pi x}{L} dx\,.$$

Therefore, by solving the ODE for $d_k(t)$, we expect that the solution u to (8.33) can be expressed by

$$u(x,t) = \sum_{k=1}^{\infty} \Big[d_k(0) e^{-\frac{\pi^2 k^2}{L^2} t} + \int_0^t f_k(s) e^{-\frac{\pi^2 k^2}{L^2}(t-s)} ds \Big] \sin \frac{k\pi x}{L}\,. \tag{8.34}$$

Following the procedure in the previous section, let u_n, f_n and g_n be the partial sums

$$u_n(x,t) = \sum_{k=1}^{n} d_k(t) \sin \frac{k\pi x}{L},\ f_n(x,t) = \sum_{k=1}^{n} f_k(t) \sin \frac{k\pi x}{L},\ g_n(x) = \sum_{k=1}^{n} d_k(0) \sin \frac{k\pi x}{L},$$

and define $v^{n,m} = u_n - u_m$ and $g^{n,m} = g_n - g_m$, $f^{n,m} = f_n - f_m$. Then $v^{n,m}$ satisfies

$$
\begin{aligned}
v^{n,m}_t(x,t) - v^{n,m}_{xx}(x,t) &= f^{n,m}(x,t) &&\text{for all } (x,t) \in (0,L)\times(0,T), && (8.35a)\\
v^{n,m}(0,t) = v^{n,m}(L,t) &= 0 &&\text{for all } t \in (0,T), && (8.35b)\\
v^{n,m}(x,0) &= g^{n,m}(x) &&\text{for all } x \in (0,L). && (8.35c)
\end{aligned}
$$

Unlike the case in Section 8.4.2, this time we cannot multiply (8.35a) by $v^{n,m}_{xxxx}(x,t)$ then integrating over $(0,L)$ since non-vanishing uncontrollable boundary terms pop out after integrating by parts. To overcome this, we differentiate (8.35a) with respect to t and then multiply the resulting equation with $v^{n,m}_t(x,t)$ and obtain that

$$\big(v^{n,m}_{tt}(t), v^{n,m}_t(t)\big)_{L^2(0,L)} - \big(v^{n,m}_{xxt}(t), v^{n,m}_t(t)\big)_{L^2(0,L)} = \big(f^{n,m}_t(t), v^{n,m}_t(t)\big)_{L^2(0,L)} .$$

It is easy to see that

$$\big(v^{n,m}_{tt}(t), v^{n,m}_t(t)\big)_{L^2(0,L)} = \frac{1}{2}\frac{d}{dt}\|v^{n,m}_t(\cdot,t)\|^2_{L^2(0,L)} .$$

Since $v^{n,m}_t(0,t) = v^{n,m}_t(L,t) = 0$, integrating by parts we have

$$-\big(v^{n,m}_{xxt}(t), v^{n,m}_t(t)\big)_{L^2(0,L)} = \|v^{n,m}_{xt}(\cdot,t)\|^2_{L^2(0,L)} .$$

As a consequence,

$$
\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|v^{n,m}_t(\cdot,t)\|^2_{L^2(0,L)} + \|v^{n,m}_{xt}(\cdot,t)\|^2_{L^2(0,L)}\\
&\qquad\qquad \leqslant \frac{1}{2}\|f^{n,m}_t(\cdot,t)\|^2_{L^2(0,L)} + \frac{1}{2}\|v^{n,m}_t(\cdot,t)\|^2_{L^2(0,L)} ;
\end{aligned}
$$

thus the Gronwall inequality implies that

$$\max_{t\in[0,T]}\|v^{n,m}_t(\cdot,t)\|^2_{L^2(0,L)} \leqslant \Big[\|v^{n,m}_t(\cdot,0)\|^2_{L^2(0,L)} + \int_0^T \|f^{n,m}_t(\cdot,t)\|^2_{L^2(0,L)}dt\Big]e^T .$$

Using (8.35a,c), we further conclude that

$$
\begin{aligned}
&\max_{t\in[0,T]}\|v^{n,m}_t(\cdot,t)\|^2_{L^2(0,L)}\\
&\quad \leqslant C\Big[\|v^{n,m}_{xx}(\cdot,0)\|^2_{L^2(0,T)} + \|f^{n,m}(\cdot,0)\|^2_{L^2(0,T)} + \int_0^T \|f^{n,m}_t(\cdot,t)\|^2_{L^2(0,L)}dt\Big]e^T\\
&\quad \leqslant C\Big[\|g^{n,m}_{xx}\|^2_{L^2(0,T)} + \|f^{n,m}(\cdot,0)\|^2_{L^2(0,T)} + \int_0^T \|f^{n,m}_t(\cdot,t)\|^2_{L^2(0,L)}dt\Big]e^T , \qquad (8.36)
\end{aligned}
$$

where we emphasize that $v^{n,m}_{xx} = (v^{n,m})_{xx}$ and $g^{n,m}_{xx} = (g^{n,m})_{xx}$.

Suppose that $f \in \mathscr{C}([0,T]; L^2(0,L))$ with $f_t \in L^2(0,T; L^2(0,L))$. Then

$$\|f^{n,m}(\cdot,0)\|^2_{L^2(0,L)} + \int_0^T \|f^{n,m}_t(\cdot,t)\|^2_{L^2(0,L)} dt \to 0 \quad \text{as } n,m \to \infty .$$

The convergence of $\|g^{n,m}_{xx}\|_{L^2(0,L)}$ to 0 as $n,m \to \infty$; however, is a bit trickier. We first note that $g \in H^2(0,L)$ does not guarantee $\|g^{n,m}_{xx}\|_{L^2(0,L)} \to 0$. For example, if $g = 1$ is a constant function, then the weak derivatives $g' = g'' = 0$ which suggests that $g \in H^2(0,L)$, but for $m > n$,

$$g^{n,m}(x) = \sum_{k=n+1}^{m} \frac{2}{k\pi}(1-(-1)^k) \sin \frac{k\pi x}{L}$$

which clearly suggests that $\|g^{n,m}_{xx}\|_{L^2(0,L)} \to \infty$ as $m \to \infty$. The key here is that the basis $\Big\{\sqrt{\frac{2}{L}} \sin \frac{k\pi x}{L}\Big\}_{k=1}^{\infty}$ we use does not always satisfy the property that

$$\Big(\frac{d}{dx}\Big)^2 \sum_{k=1}^{\infty} a_k \sin \frac{k\pi x}{L} = \sum_{k=1}^{\infty} a_k \Big(\frac{d}{dx}\Big)^2 \sin \frac{k\pi x}{L} , \tag{8.37}$$

where $\frac{d}{dx}$ is the weak differential operator, and $\{a_k\}_{k=1}^{\infty}$ is a sequence that decays very fast (so that the sum make senses). For (8.37) to hold, the function $\sum_{k=1}^{\infty} a_k \sin \frac{k\pi x}{L}$ and its weak derivative has to vanish at $x = 0$ and $x = L$. In fact, we have the following

LEMMA 8.45. *If $g \in H^2(0,L)$ and $g(0) = g'(0) = g(L) = g'(L) = 0$, then the partial sum*

$$g_n(x) = \sqrt{\frac{2}{L}} \sum_{k=1}^{n} \widehat{g}(k) \sin \frac{k\pi x}{L} , \quad \text{where} \;\; \widehat{g}(k) = \sqrt{\frac{2}{L}} \int_0^L g(x) \sin \frac{k\pi x}{L} dx ,$$

has the property that

$$\lim_{n,m\to\infty} \big\|(g_n - g_m)''\big\|_{L^2(0,L)} = 0 .$$

Proof. If $g \in H^2(0,L)$ and $g(0) = g'(0) = g(L) = g'(L) = 0$, then

$$\begin{aligned}\widehat{g''}(k) &= \sqrt{\frac{2}{L}} \int_0^L g''(x) \sin \frac{k\pi x}{L} dx = -\frac{k\pi}{L}\sqrt{\frac{2}{L}} \int_0^L g'(x) \cos \frac{k\pi x}{L} dx \\ &= -\frac{k^2\pi^2}{L^2}\sqrt{\frac{2}{L}} \int_0^L \sin \frac{k\pi x}{L} dx = -\frac{k^2\pi^2}{L^2} \widehat{g}(k) .\end{aligned}$$

As a consequence,

1. $g_n''(x) = -\sqrt{\dfrac{2}{L}} \sum\limits_{k=1}^{n} \dfrac{k^2\pi^2}{L^2} \widehat{g}(k) \sin \dfrac{k\pi x}{L}$;

2. $g'' \in L^2(0,L)$ if and only if $\sum\limits_{k=1}^{\infty} |k|^4 \big|\widehat{g}(k)\big|^2 < \infty$;

thus for $m > n$,

$$\big\|(g_n - g_m)''\big\|^2_{L^2(0,L)} = \sum_{k=n+1}^{m} \Big|\frac{k^2\pi^2}{L^2}\widehat{g}(k)\Big|^2$$

which converges to 0 as $n, m \to \infty$. □

In addition to $f \in \mathscr{C}([0,T]; L^2(0,L))$ with $f_t \in L^2(0,T; L^2(0,L))$, we now assume further that $g \in H^2(0,L)$ with $g(0) = g'(0) = g(L) = g'(L) = 0$. Then Lemma 8.45 suggests that $v_t^{n,m}$ is a Cauchy sequence in $\mathscr{C}([0,T]; L^2(0,L))$. Similarly, multiplying (8.35a) by $v^{n,m}$ and then integrating over the interval $(0,L)$, with the help of the Gronwall inequality, provides the estimate

$$\max_{t\in[0,T]} \|v^{n,m}(\cdot,t)\|^2_{L^2(0,L)} \leqslant C\Big[\|g^{n,m}\|^2_{L^2(0,L)} + \int_0^T \|f^{n,m}(\cdot,t)\|^2_{L^2(0,L)} dt\Big] e^T .$$

Moreover, using (8.35a) we can also conclude that

$$\max_{t\in[0,T]} \|v_{xx}^{n,m}(\cdot,t)\|_{L^2(0,L)} \leqslant \max_{t\in[0,T]} \Big[\|v_t^{n,m}(\cdot,t)\|_{L^2(0,L)} + \|f^{n,m}(\cdot,t)\|_{L^2(0,L)}\Big];$$

thus $v^{n,m}$ is a Cauchy sequence in $\mathscr{C}([0,T]; H^2(0,L))$. Therefore, u_n converges uniformly to some function $u \in \mathscr{C}([0,T]; H^2(0,L))$ (which also implies that u_{nxx} converges uniformly to $u_{xx} \in \mathscr{C}([0,T]; L^2(0,L))$) and u_{nt} converges uniformly to $u_t \in \mathscr{C}([0,T]; L^2(0,L))$, and u satisfies

$$\begin{aligned} &\max_{t\in[0,T]} \Big[\|u_t(t)\|_{L^2(0,L)} + \|u(t)\|_{H^2(0,L)}\Big] \\ &\qquad \leqslant C\Big[\|g\|_{H^2(0,L)} + \max_{t\in[0,T]} \|f(t)\|^2_{L^2(0,L)} + \int_0^T \|f_t(t)\|^2_{L^2(0,L)} dt\Big]. \end{aligned} \tag{8.38}$$

Therefore, we establish the following

Theorem 8.46. *Let $f \in \mathscr{C}([0,T]; L^2(0,L))$ with $f_t \in L^2(0,T; L^2(0,L))$ and $g \in H^2(0,L)$ with $g(0) = g'(0) = g(L) = g'(L) = 0$, then u defined in* (8.34) *is a solution to* (8.33). *Moreover, u satisfies* (8.38).

8.6 Exercises

Problem 8.1. Prove Lemma 8.10.

Problem 8.2. Let f be a 2π-periodic Lipchitz function. Show that for $n \geqslant 2$,

$$\|f - F_{n+1} \star f\|_{L^\infty(\mathbb{T})} \leqslant \frac{1 + 2\log n}{2n} \|f\|_{\mathscr{C}^{0,1}(\mathbb{T})} \tag{8.39}$$

and

$$\big\|f - s_n(f, \cdot)\big\|_{L^\infty(\mathbb{T})} \leqslant \frac{2\pi(1 + \log n)^2}{n} \|f\|_{\mathscr{C}^{0,1}(\mathbb{T})} . \tag{8.40}$$

Hint: For (8.39), apply the estimate

$$F_n(x) \leqslant \min\Big\{\frac{n+1}{2\pi}, \frac{\pi}{2(n+1)x^2}\Big\}$$

in the following inequality:

$$\big|f(x) - F_{n+1} \star f(x)\big| \leqslant \Big[\int_{-\delta}^{\delta} + \int_{-\pi}^{-\delta} + \int_{\delta}^{\pi}\Big]\big|f(x+y) - f(x)\big|F_{n+1}(y)\,dy$$

with $\delta = \dfrac{\pi}{n+1}$. For (8.40), use (8.8) and note that

$$\inf_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_{L^\infty(\mathbb{T})} \leqslant \|f - F_n \star f\|_{L^\infty(\mathbb{T})} .$$

Problem 8.3. A function $f : \mathbb{T} \to \mathbb{R}$ is said to be piecewise $\mathscr{C}^1$ if there are finitely many disjoint open intervals I_i so that $f \in \mathscr{C}^1(I_i)$ for all i and $\bigcup\limits_i \bar{I}_i = \mathbb{T}$. Show that $D_n \star f$ converges to f uniformly as $n \to \infty$ on any compact subset of I_i.

Problem 8.4. In this problem, we are concerned with the following

Theorem 8.47 (Bernstein)**.** *Suppose that f is a 2π-periodic function such that for some constant C and $\alpha \in (0,1)$,*

$$\inf_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_{L^\infty(\mathbb{T})} \leqslant Cn^{-\alpha}$$

for all $n \in \mathbb{N}$. Then $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$.

Complete the following to prove the theorem.

1. Suppose that there is $p \in \mathscr{P}_n(\mathbb{T})$ such that
$$\|p'\|_{L^\infty(\mathbb{T})} > n\,, \quad \|p\|_{L^\infty(\mathbb{T})} < 1\,, \quad \text{and} \quad p'(0) = \|p'\|_{L^\infty(\mathbb{T})}.$$
Choose $\gamma \in \big[-\frac{\pi}{n}, \frac{\pi}{n}\big]$ such that $\sin(n\gamma) = -p(0)$ and $\cos(n\gamma) > 0$, and define $\alpha_k = \gamma + \frac{\pi}{n}\big(k + \frac{1}{2}\big)$ for $-n \leqslant k \leqslant n$. Show that the function $r(x) = \sin n(x-\gamma) - p(x)$ has at least one zeros in each interval (α_k, α_{k+1}).

2. Let $s \in \mathbb{N}$ be such that such that $0 \in (\alpha_s, \alpha_{s+1})$. Show that r has at least 3 distinct zeros in (α_s, α_{s+1}) by noting that $r'(0) < 0$ and $r(0) = 0$.

3. Combining 1 and 2, show that
$$\|p'\|_{L^\infty(\mathbb{T})} \leqslant n\|p\|_{L^\infty(\mathbb{T})} \qquad \forall\, p \in \mathscr{P}_n(\mathbb{T}). \tag{8.41}$$

4. Choose $p_n \in \mathscr{P}_n(\mathbb{T})$ such that $\|f - p_n\| \leqslant 2Cn^{-\alpha}$ for $n \in \mathbb{N}$. Define $q_0 = p_1$, and $q_n = p_{2^n} - p_{2^{n-1}}$ for $n \in \mathbb{N}$. Show that $\sum\limits_{n=0}^{\infty} q_n = f$ and the convergence is uniform.

5. Show that $\|q_n\|_{L^\infty(\mathbb{T})} \leqslant 6C2^{-n\alpha}$. As a consequence, show that
$$\big|q_n(x) - q_n(y)\big| \leqslant 6Cn2^{n(1-\alpha)}|x-y| \;\text{ and }\; \big|q_n(x) - q_n(y)\big| \leqslant 12C2^{-n\alpha}.$$

6. For any $x, y \in \mathbb{T}$ with $|x-y| \leqslant 1$, choose $m \in \mathbb{N}$ such that $2^{-m} \leqslant |x-y| \leqslant 2^{1-m}$. Then use the inequality
$$\big|f(x) - f(y)\big| \leqslant \sum_{n=0}^{m-1} \big|q_n(x) - q_n(y)\big| + \sum_{n=m}^{\infty} \big|q_n(x) - q_n(y)\big|$$
to show that $\big|f(x) - f(y)\big| \leqslant B|x-y|^\alpha$ for some constant $B > 0$.

Problem 8.5. Show that $\{w_k\}_{k=0}^\infty$ defined in Remark 8.39 is an orthonormal basis of $H^1(0,\pi)$.

Hint: Use the Parseval identity to show that $\{w_k\}_{k=0}^\infty$ is a maximal orthonormal set of $H^1(0,\pi)$; that is, show that for all $f \in H^1(0,\pi)$,
$$\|f\|^2_{H^1(0,\pi)} = \int_0^\pi \big(|f(x)|^2 + |f'(x)|^2\big)\,dx = \sum_{k=0}^{\infty} \big|(f, w_k)_{H^1(0,\pi)}\big|^2\,.$$
You might need the fact that $\Big\{\sqrt{\frac{2}{\pi}}\sin kx\Big\}_{k=1}^\infty$ is an orthonormal basis of $L^2(0,\pi)$.

Problem 8.6. Let $f(x) = x$ on $[-\pi, \pi]$. Then $f'(x) = 1$ is certainly a $L^2(-\pi,\pi)$-function. However, you may want to check that $\sum\limits_{n=-\infty}^{\infty} |n|^2 |\widehat{f}(n)|^2 = \infty$, so by "definition", it does not seem to be a function in $H^1(-\pi,\pi)$. What is wrong with the argument?

Problem 8.7. Show that $H^1(\mathbb{T})$ is the completion of the normed space $\big(\mathscr{C}(\mathbb{T}), \|\cdot\|_{H^1(\mathbb{T})}\big)$.

Problem 8.8 (Generalized Gronwall inequality)**.** Show that if $a \in L^1(0,T)$ is a non-negative function , and $x(t)$ satisfies the following integral integral inequality

$$x(t) \leqslant M + \int_0^t a(s)x(s)\,ds\,.$$

Then $x(t) \leqslant M \exp\Big(\int_0^t a(s)\,ds\Big)$ for all $t \in [0,T]$. In particular, if x satisfies

$$x'(t) \leqslant b(t) + a(t)x(t)$$

for some $a, b \in L^1(0,T)$ and $a \geqslant 0$, then

$$x(t) \leqslant \big[x(0) + \|b\|_{L^1(0,T)}\big] \exp\Big(\int_0^t a(s)\,ds\Big) \quad \forall\, t \in [0,T]\,.$$

Problem 8.9. Use Fourier series to formally solve the following initial-boundary value problem for the wave equation

$$\begin{aligned} u_{tt}(x,t) &= c^2 u_{xx}(x,t) && \text{in } (0,1)\times\mathbb{R}\,,\\ u(0,t) &= u(1,t) = 0 && \text{for all } t\,,\\ u(x,0) &= f(x)\,, \quad u_t(x,0) = g(x) && \forall\, x \in [0,1]\,. \end{aligned}$$

Derive the following two conservation laws from your Fourier series solution and directly from the PDE:

$$\frac{d}{dt}\int_0^1 \Big[|u_t(x,t)|^2 + c^2|u_x(x,t)|^2\Big]dx = 0\,.$$

Problem 8.10. Use Fourier series to formally solve the following initial-boundary value problem for the Schrödinger equation

$$\begin{aligned} iu_t(x,t) &= -u_{xx}(x,t) && \text{in } (0,1)\times\mathbb{R}\,,\\ u(0,t) &= u(1,t) = 0 && \text{for all } t\,,\\ u(x,0) &= f(x) && \forall\, x \in [0,1]\,. \end{aligned}$$

Derive the following two conservation laws from your Fourier series solution and directly from the PDE:

$$\frac{d}{dt}\int_0^1 |u(x,t)|^2 dx = 0\,, \qquad \frac{d}{dt}\int_0^1 |u_x(x,t)|^2 dx = 0\,.$$

Problem 8.11. Try using the Fourier series to solve

$$\begin{aligned} u_t(x,t) &= u_{xx}(x,t) && \text{in } (0,\pi)\times(0,\infty)\,,\\ u(0,t) &= u_x(\pi,t) = 0 && \text{for all } t\,,\\ u(x,0) &= f(x) && \forall\, x\in[0,\pi]\,. \end{aligned}$$

The most important task is to look for a suitable basis that fits the boundary condition.

Problem 8.12. Let (r,θ) be the polar coordinate on $\mathbb{R}^2$.

(1) Show that a harmonic function u on $\Omega\subset\mathbb{R}^2$ satisfies

$$\frac{1}{r}(ru_r)_r + \frac{1}{r^2}u_{\theta\theta} = 0 \qquad r>0\,.$$

(2) For $\alpha>0$, let Ω_α be the wedge given in polar coordinates (r,θ) by

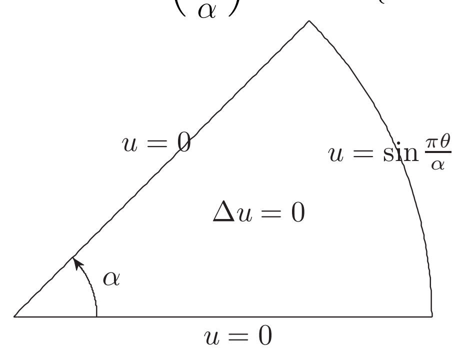
$$\Omega_\alpha = \{(r,\theta)\mid 0<r<1\,,\ 0<\theta<\alpha\}\,.$$

Based on the fact that the general solution to

$$r^2R''(r) + rR'(r) - s^2R(r) = 0$$

is of the form $R(r) = C_1r^s + C_2r^{-s}$, use the Fourier series to find a bounded solution to the following boundary value problem

$$\begin{aligned} \Delta u &= 0 && \text{in} \quad \Omega_\alpha\,,\\ u &= 0 && \text{on} \quad \{\theta = 0,\alpha\}\,,\\ u &= \sin\Big(\frac{\pi\theta}{\alpha}\Big) && \text{on} \quad \{r=1\}\,. \end{aligned}$$



Hint: Suppose that $u(r,\theta) = \sum_k R_k(r)\mathrm{e}_k(\theta)$, where $\{\mathrm{e}_k\}$ forms an orthonormal basis of $L^2(0,\alpha)$ satisfying certain boundary conditions (you have to figure out what these boundary conditions are). Solve R_k by finding an ODE for R_k.

(3) Find all $\alpha > 0$ so that $u \in \mathscr{C}^2(\overline{\Omega_\alpha})$.

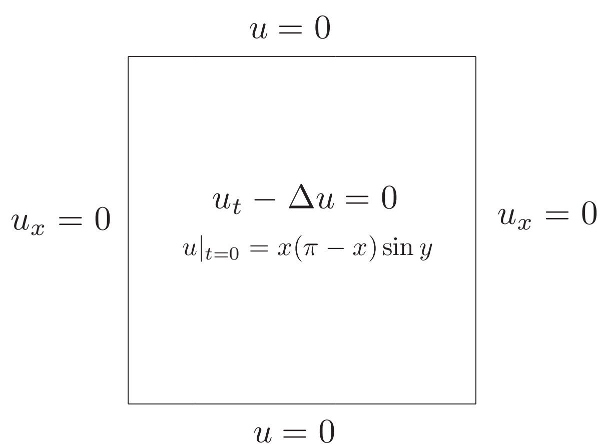
PROBLEM 8.13. Complete the following.

(1) Suppose that $\{\mathrm{e}_n\}_{n=1}^\infty$ is an orthonormal basis of $L^2(0,\ell_1)$ and $\{\widetilde{e}_m\}_{m=1}^\infty$ is an orthonormal basis of $L^2(0,\ell_2)$. Show that $\{\mathrm{e}_n(x)\widetilde{e}_m(y)\}_{n,m=1}^\infty$ forms an orthonormal basis of $L^2([0,\ell_1]\times[0,\ell_2])$.

Hint: Check the orthonormality and the maximality. For the maximality, check the Parseval identity.

(2) Solve the following PDE:

$$\begin{aligned}
u_t(x,y,t) - \Delta u(x,y,t) &= 0 && (x,y) \in (0,\pi)\times(0,\pi)\,, t > 0\,,\\
u(x,y,0) &= x(\pi - x)\sin y && (x,y) \in (0,\pi)\times(0,\pi)\,,\\
u_x(0,y,t) = u_x(\pi,y,t) &= 0 && y \in (0,\pi)\,, t > 0\,,\\
u(x,0,t) = u(x,\pi,t) &= 0 && x \in (0,\pi)\,, t > 0\,.
\end{aligned}$$



(3) Show that for all $t \geqslant 0$, u from (b) satisfies

$$\int_0^\pi\int_0^\pi |u(x,y,t)|^2 dxdy + 2\int_0^t\int_0^\pi\int_0^\pi |\nabla u(x,y,s)|^2 dxdyds = \frac{\pi^6}{60}\,.$$

Appendix A

Review of Elementary Analysis

In this chapter of appendix, we review some of the most important contents from elementary analysis.

A.1 Differential Calculus

A.1.1 Bounded Linear Maps

Before defining the differentiability of functions of several variables, we introduce the notion of a bounded linear map.

Definition A.1. A map L from a vector space X into a vector space Y is said to be ***linear*** if $L(cx_1 + x_2) = cL(x_1) + L(x_2)$ for all $x_1, x_2 \in X$ and $c \in \mathbb{R}$. We often write Lx instead of $L(x)$, and the collection of all linear maps from X to Y is denoted by $\mathscr{L}(X, Y)$.

Suppose further that X and Y are normed spaces equipped with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively. A linear map $L : X \to Y$ is said to be bounded if

$$\sup_{\|x\|_X = 1} \|Lx\|_Y < \infty\,.$$

The collection of all bounded linear maps from X to Y is denoted by $\mathscr{B}(X, Y)$, and the number $\sup\limits_{\|x\|_X = 1} \|Lx\|_Y$ is often denoted by $\|L\|_{\mathscr{B}(X,Y)}$.

Example A.2. Let $\mathcal{M}_{n\times m} \equiv \big\{n \times m \text{ matrix with entries in } \mathbb{R}\big\}$, and we remind the readers that if $A \in \mathcal{M}_{n\times m}$, then $A : \begin{cases} \mathbb{R}^m & \to & \mathbb{R}^n \\ x & \mapsto & Ax \end{cases}$. Define

$$\|A\|_p = \sup_{\|x\|_p=1} \|Ax\|_p = \sup_{x\neq 0} \frac{\|Ax\|_p}{\|x\|_p} \quad \forall\, A \in \mathcal{M}_{n\times m}\,;$$

that is, $\|A\|_p$ is the least upper bound of the set $\left\{\frac{\|Ax\|_p}{\|x\|_p} \,\middle|\, x \neq 0, x \in \mathbb{R}^m\right\}$. Therefore, $\frac{\|Ax\|_p}{\|x\|_p} \leqslant \|A\|_p\ \forall\, x \neq 0$; thus

$$\|Ax\|_p \leqslant \|A\|_p\|x\|_p \quad \forall\, x \in \mathbb{R}^m\,.$$

Consider the case $p = 1, p = 2$ and $p = \infty$ respectively.

1. $p = 2$: Let $(\cdot,\cdot)_{\mathbb{R}^k}$ denote the inner product in Euclidean space $\mathbb{R}^k$. Then

$$\|Ax\|_2^2 = (Ax, Ax)_{\mathbb{R}^n} = (x, A^{\mathrm{T}}Ax)_{\mathbb{R}^m} = (x, P\Lambda P^{\mathrm{T}}x)_{\mathbb{R}^m} = (P^{\mathrm{T}}x, \Lambda P^{\mathrm{T}}x)_{\mathbb{R}^n}\,,$$

in which we use the fact that $A^{\mathrm{T}}A$ is symmetric; thus diagonalizable by an orthonormal matrix P (that is, $A^{\mathrm{T}}A = P\Lambda P^{\mathrm{T}}$, $P^{\mathrm{T}}P = I$, Λ is a diagonal matrix). Therefore,

$$\begin{aligned}\sup_{\|x\|_2=1} \|Ax\|_2^2 &= \sup_{\|x\|_2=1} (P^{\mathrm{T}}x, \Lambda P^{\mathrm{T}}x) = \sup_{\|y\|_2=1} (y, \Lambda y) \quad (\text{Let } y = P^{\mathrm{T}}x, \text{ then } \|y\|_2 = 1)\\ &= \sup_{\|y\|_2=1} (\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2)\\ &= \max\big\{\lambda_1, \cdots, \lambda_n\big\} = \text{maximum eigenvalue of } A^{\mathrm{T}}A\end{aligned}$$

which implies that $\|A\|_2 = \sqrt{\text{maximum eigenvalue of } A^{\mathrm{T}}A}$.

2. $p = \infty$: $\|A\|_\infty = \sup\limits_{\|x\|_\infty=1} \|Ax\|_\infty = \max\left\{\sum\limits_{j=1}^m |a_{1j}|, \sum\limits_{j=1}^m |a_{2j}|, \cdots \sum\limits_{j=1}^m |a_{nj}|\right\}$.

Reason: Let $x = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$ and $A = \big[a_{ij}\big]_{n\times m}$. Then

$$Ax = \begin{bmatrix} a_{11}x_1 + \cdots + a_{1m}x_m \\ a_{21}x_1 + \cdots + a_{2m}x_m \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nm}x_m \end{bmatrix}$$

Assume $\max\limits_{1\leqslant i\leqslant n}\sum\limits_{j=1}^{m}|a_{ij}| = \sum\limits_{j=1}^{m}|a_{kj}|$ for some $1\leqslant k\leqslant n$. Let

$$x = (\mathrm{sgn}(a_{k1}), \mathrm{sgn}(a_{k2}), \cdots, \mathrm{sgn}(a_{kn}))\ .$$

Then $\|x\|_\infty = 1$, and $\|Ax\|_\infty = \sum\limits_{j=1}^{m}|a_{kj}|$.

On the other hand, if $\|x\|_\infty = 1$, then

$$|a_{i1}x_1 + a_{i2}x_2 + \cdots a_{im}x_m| \leqslant \sum_{j=1}^{m}|a_{ij}| \leqslant \max\left\{\sum_{j=1}^{m}|a_{1j}|, \sum_{j=1}^{m}|a_{2j}|, \cdots \sum_{j=1}^{m}|a_{nj}|\right\};$$

thus $\|A\|_\infty = \max\left\{\sum\limits_{j=1}^{m}|a_{1j}|, \sum\limits_{j=1}^{m}|a_{2j}|, \cdots \sum\limits_{j=1}^{m}|a_{nj}|\right\}$. In other words, $\|A\|_\infty$ is the largest sum of the absolute value of row entries.

3. $p = 1$: $\|A\|_1 = \max\left\{\sum\limits_{i=1}^{n}|a_{i1}|, \sum\limits_{i=1}^{n}|a_{i2}|, \cdots, \sum\limits_{i=1}^{n}|a_{im}|\right\}$.

Example A.3. Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be given by $Lx = Ax$, where A is an $m\times n$ matrix. Then Example A.2 shows that $\|L\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)}$ is the square root of the largest eigenvalue of $A^{\mathrm{T}}A$ which is certainly a finite number. Therefore, any linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is bounded.

Example A.4. Let $\mathscr{C}$ be the collection of all continuous real-valued functions on the interval $[0,1]$; that is,

$$\mathscr{C} = \left\{f : [0,1] \to \mathbb{R}\,\middle|\, f \text{ is continuous on } [0,1]\right\}.$$

For each $f \in \mathscr{C}$, we define

$$\|f\|_p = \begin{cases} \left[\displaystyle\int_0^1 |f(x)|^p dx\right]^{\frac{1}{p}} & \text{if } 1\leqslant p<\infty\,, \\ \max\limits_{x\in[0,1]}|f(x)| & \text{if } p=\infty\,. \end{cases}$$

The function $\|\cdot\|_p : \mathscr{C} \to \mathbb{R}$ is a norm on $\mathscr{C}$ (Minkowski's inequality).

Proposition A.5. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces, and $L \in \mathscr{B}(X,Y)$. Then*

$$\|L\|_{\mathscr{B}(X,Y)} = \sup_{x\neq 0}\frac{\|Lx\|_Y}{\|x\|_X} = \inf\left\{M>0\,\middle|\, \|Lx\|_Y \leqslant M\|x\|_X\right\}.$$

In particular, the first equality implies that

$$\|Lx\|_Y \leqslant \|L\|_{\mathscr{B}(X,Y)}\|x\|_X \qquad \forall\, x \in X\,.$$

Proposition A.6. *Let $(X,\|\cdot\|_X)$ and $(Y,\|\cdot\|_Y)$ be normed spaces, and $L \in \mathscr{L}(X,Y)$. Then L is continuous on X if and only if $L \in \mathscr{B}(X,Y)$.*

Proposition A.7. *Let $(X,\|\cdot\|_X)$ and $(Y,\|\cdot\|_Y)$ be normed vector spaces. Then $\big(\mathscr{B}(X,Y),\|\cdot\|_{\mathscr{B}(X,Y)}\big)$ is a normed space. Moreover, if $(Y,\|\cdot\|_Y)$ is a Banach space, so is $\big(\mathscr{B}(X,Y),\|\cdot\|_{\mathscr{B}(X,Y)}\big)$.*

Proposition A.8. *Let $(X,\|\cdot\|_X)$, $(Y,\|\cdot\|_Y)$, $(Z,\|\cdot\|_Z)$ be normed spaces, and $L \in \mathscr{B}(X,Y)$, $K \in \mathscr{B}(Y,Z)$. Then $K\circ L \in \mathscr{B}(X,Z)$, and*

$$\|K\circ L\|_{\mathscr{B}(X,Z)} \leqslant \|K\|_{\mathscr{B}(Y,Z)}\|L\|_{\mathscr{B}(X,Y)}\,.$$

We often write $K\circ L$ as KL if K and L are linear.

From now on, when the domain X and the target Y of a linear map L is clear, we use $\|L\|$ instead of $\|L\|_{\mathscr{B}(X,Y)}$ to simplify the notation.

Theorem A.9. *Let $(X,\|\cdot\|_X)$ and $(Y,\|\cdot\|_Y)$ be normed spaces, and X be finite dimensional. Then every linear map from X to Y is bounded; that is, $\mathscr{L}(X,Y) = \mathscr{B}(X,Y)$.*

Theorem A.10. *Let $\mathrm{GL}(n)$ be the set of all invertible linear maps on $\mathbb{R}^n$; that is,*

$$\mathrm{GL}(n) = \big\{L \in \mathscr{L}(\mathbb{R}^n,\mathbb{R}^n)\,\big|\,L \text{ is one-to-one (and onto)}\big\}\,.$$

1. *If $L \in \mathrm{GL}(n)$ and $K \in \mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)$ satisfying $\|K-L\|\|L^{-1}\| < 1$, then $K \in \mathrm{GL}(n)$.*
2. *$\mathrm{GL}(n)$ is an open set of $\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)$.*
3. *The mapping $L \mapsto L^{-1}$ is continuous on $\mathrm{GL}(n)$.*

Remark A.11. Even though 2 is a direct consequence of 1 in Theorem A.10, there is another way to see that $\mathrm{GL}(n)$ is open in $\mathscr{B}(\mathbb{R}^n,\mathbb{R}^n)$. Let $\mathcal{M}(n)$ be the collection of $n\times n$ real matrices, and $\|\cdot\|_2$ be the matrix norm. Also define $\|\cdot\| : \mathcal{M}(n) \to \mathbb{R}$ by

$$\|A\| = \max\big\{|a_{ij}|\,\big|\,A = [a_{ij}]1 \leqslant i,j \leqslant n\big\}\,.$$

Then $\|\cdot\|$ is also a norm on $\mathcal{M}(n)$. Since $\mathcal{M}(n)$ is finite dimensional (in fact, $\dim\mathcal{M}(n)=n^2$), $\|\cdot\|$ and $\|\cdot\|_2$ are equivalent norms on $\mathcal{M}(n)$; that is, there exists $C,c>0$ such that

$$c\|A\|\leqslant\|A\|_2\leqslant C\|A\| \qquad \forall\, A\in\mathcal{M}(n)\,.$$

Let $\{A_k\}_{k=1}^{\infty}\subseteq\mathcal{M}(n)$ be a sequence of $n\times n$ real matrices. The equivalence between $\|\cdot\|$ and $\|\cdot\|_2$ implies that $A_k\to A$ in $\mathcal{M}(n)$ if and only if each entry of A_k converges to corresponding entry of A. Therefore, the determinant function is continuous on $\mathcal{M}(n)$. In other words,

$$\lim_{A_k\to A}\det(A_k)=\det(A) \qquad \forall\, A\in\mathcal{M}(n)\,.$$

Since $\mathrm{GL}(n)$ can be viewed as the collection of $n\times n$ matrices with non-zero determinant; that is,

$$\mathrm{GL}(n)=\big\{A\in\mathcal{M}(n)\,\big|\,\det(A)\neq 0\big\}\,,$$

by the continuity of the determinant function and the fact that the pre-image of open sets under continuous functions are open, we conclude that $\mathrm{GL}(n)$ is open in $\mathcal{M}(n)$.

A.1.2 Definition of Derivatives

Definition A.12. Let $(X,\|\cdot\|_X)$ and $(Y,\|\cdot\|_Y)$ be two normed spaces. A map $f:A\subseteq X\to Y$ is said to be ***differentiable*** at $x_0\in A$ if there is a bounded linear map, denoted by $(Df)(x_0):X\to Y$ and called the ***derivative*** of f at x_0, such that

$$\lim_{\substack{x\to x_0\\ x\in A}}\frac{\big\|f(x)-f(x_0)-(Df)(x_0)(x-x_0)\big\|_Y}{\|x-x_0\|_X}=0\,,$$

where $(Df)(x_0)(x-x_0)$ denotes the value of the linear map $(Df)(x_0)$ applied to the vector $x-x_0\in X$ (so $(Df)(x_0)(x-x_0)\in Y$). In other words, f is differentiable at $x_0\in A$ if there exists $L\in\mathscr{B}(X,Y)$ such that

$$\forall\,\varepsilon>0,\exists\,\delta>0 \ni \|f(x)-f(x_0)-L(x-x_0)\|_Y\leqslant\varepsilon\|x-x_0\|_X \text{ whenever } x\in B(x_0,\delta)\cap A\,.$$

If f is differentiable at each point of A, we say that f is differentiable on A.

Example A.13. Let $f:\mathrm{GL}(n)\to\mathrm{GL}(n)$ be given by $f(L)=L^{-1}$, where $\mathrm{GL}(n)$ is defined in Theorem A.10. Then f is differentiable at any "point" $L\in\mathrm{GL}(n)$ with derivative $(Df)(K)\in\mathscr{B}(\mathrm{GL}(n),\mathrm{GL}(n))$ given by $(Df)(L)(K)=-L^{-1}KL^{-1}$ for all $K\in\mathrm{GL}(n)$. The proof is left as an exercise.

Theorem A.14. *Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be normed vector spaces, $\mathcal{U} \subseteq X$ be an open set, and $f : \mathcal{U} \to Y$ be differentiable at $x_0 \in \mathcal{U}$. Then $(Df)(x_0)$ is uniquely determined by f.*

Remark A.15. Let $\mathcal{U} \subseteq \mathbb{R}^n$ be an open set and suppose that $f : \mathcal{U} \to \mathbb{R}^m$ is differentiable on $\mathcal{U}$. Then $Df : \mathcal{U} \to \mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)$. Treating Df as a map from $\mathcal{U}$ to the normed space $\big(\mathscr{B}(\mathbb{R}^n, \mathbb{R}^m), \|\cdot\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)}\big)$, and suppose that Df is also differentiable on $\mathcal{U}$. Then the derivative of Df, denoted by $D^2 f$, is a map from $\mathcal{U}$ to $\mathscr{B}(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n, \mathbb{R}^m))$. In other words, for each $a \in \mathcal{U}$, $(D^2 f)(a) \in \mathscr{B}(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n, \mathbb{R}^m))$ satisfying

$$\lim_{x\to a} \frac{\big\|(Df)(x) - (Df)(a) - (D^2 f)(a)(x-a)\big\|_{\mathscr{B}(\mathbb{R}^n,\mathbb{R}^m)}}{\|x-a\|_{\mathbb{R}^n}} = 0\,,$$

here $(D^2 f)(a)$ is bounded linear map from $\mathbb{R}^n$ to $\mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)$; thus $(D^2 f)(a)(x-a) \in \mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)$.

Definition A.16. Let $\{\mathrm{e}_k\}_{k=1}^n$ be the standard basis of $\mathbb{R}^n$, $\mathcal{U} \subseteq \mathbb{R}^n$ be an open set, $a \in \mathcal{U}$ and $f : \mathcal{U} \to \mathbb{R}$ be a function. The partial derivative of f at a in the direction e_j, denoted by $\dfrac{\partial f}{\partial x_j}(a)$, is the limit

$$\lim_{h\to 0} \frac{f(a + h\mathrm{e}_j) - f(a)}{h}$$

if it exists. In other words, if $a = (a_1, \cdots, a_n)$, then

$$\frac{\partial f}{\partial x_j}(a) = \lim_{h\to 0} \frac{f(a_1, \cdots, a_{j-1}, a_j + h, a_{j+1}, \cdots, a_n) - f(a_1, \cdots, a_n)}{h}\,.$$

Theorem A.17. *Suppose $\mathcal{U} \subseteq \mathbb{R}^n$ is an open set and $f : \mathcal{U} \to \mathbb{R}^m$ is differentiable at $a \in \mathcal{U}$. Then the partial derivatives $\dfrac{\partial f_i}{\partial x_j}(a)$ exists for all $i = 1, \cdots m$ and $j = 1, \cdots n$, and the matrix representation of the linear map $Df(a)$ with respect to the standard basis of $\mathbb{R}^n$ and $\mathbb{R}^m$ is given by*

$$\big[(Df)(a)\big] = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \cdots & \dfrac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(a) & \cdots & \dfrac{\partial f_m}{\partial x_n}(a) \end{bmatrix} \quad \text{or} \quad \big[(Df)(a)\big]_{ij} = \frac{\partial f_i}{\partial x_j}(a)\,.$$

DEFINITION A.18. Let $\mathcal{U} \subseteq \mathbb{R}^n$ be an open set, and $f : \mathcal{U} \to \mathbb{R}^m$. The matrix

$$(Jf)(x) \equiv \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix} (x) \equiv \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(x) & \cdots & \dfrac{\partial f_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(x) & \cdots & \dfrac{\partial f_m}{\partial x_n}(x) \end{bmatrix}$$

is called the ***Jacobian matrix*** of f at x (if each entry exists). If $n = m$, the determinant of $(Jf)(x)$ is called the ***Jacobian*** of f at x.

REMARK A.19. A function f might not be differential even if the Jacobian matrix Jf exists; however, if f is differentiable at x_0, then $(Df)(x)$ can be represented by $(Jf)(x)$; that is, $\big[(Df)(x)\big] = (Jf)(x)$.

REMARK A.20. For each $x \in A$, $Df(x)$ is a linear map, but Df in general is not linear in x.

DEFINITION A.21. Let $\mathcal{U} \subseteq \mathbb{R}^n$ be an open set. The derivative of a scalar function $f : \mathcal{U} \to \mathbb{R}$ is called the ***gradient*** of f and is denoted by grad$f$ or ∇f.

A.1.3 Properties of Differentiable Functions

Continuity of Differentiable Maps

THEOREM A.22. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces, $\mathcal{U} \subseteq X$ be open, and $f : \mathcal{U} \to Y$ be differentiable at $x_0 \in \mathcal{U}$. Then f is continuous at x_0.*

Proof. Since f is differentiable at x_0, there exists $L \in \mathscr{B}(X, Y)$ such that

$$\exists\, \delta_1 > 0 \ni \big\|f(x) - f(x_0) - L(x - x_0)\big\|_Y \leqslant \|x - x_0\|_X \quad \forall\, x \in B(x_0, \delta_1)\,.$$

As a consequence,

$$\big\|f(x) - f(x_0)\big\|_Y \leqslant \big(\|L\| + 1\big)\|x - x_0\|_X \quad \forall\, x \in B(x_0, \delta_1)\,. \tag{A.1}$$

For a given $\varepsilon > 0$, let $\delta = \min\Big\{\delta_1, \dfrac{\varepsilon}{2(\|L\| + 1)}\Big\}$. Then $\delta > 0$, and if $x \in B(x_0, \delta)$,

$$\big\|f(x) - f(x_0)\big\|_Y \leqslant \frac{\varepsilon}{2} < \varepsilon\,.$$

□

REMARK A.23. In fact, if f is differentiable at x_0, then f satisfies the "local Lipschitz property"; that is, there exists $M = M(x_0) > 0$ and $\delta = \delta(x_0) > 0$ such that if $\|x - x_0\|_X < \delta$ then $f(x) - f(x_0)\|_Y \leqslant M\|x - x_0\|_X$ since we can choose $M = \|L\| + 1$ and $\delta = \delta_1$ (see (A.1)).

The Product Rules and Gradients

Proposition A.24. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, and $f : \mathcal{U} \to \mathbb{R}^m$ and $g : \mathcal{U} \to \mathbb{R}$ be differentiable at $x_0 \in \mathcal{U}$. Then $gf : \mathcal{U} \to \mathbb{R}^m$ is differentiable at x_0, and*

$$D(gf)(x_0)(v) = g(x_0)(Df)(x_0)(v) + (Dg)(x_0)(v)f(x_0)\,. \tag{A.2}$$

Moreover, if $g(x_0) \neq 0$, then $\frac{f}{g} : \mathcal{U} \to \mathbb{R}^m$ is also differentiable at x_0, and $D(\frac{f}{g})(x_0) : \mathbb{R}^n \to \mathbb{R}^m$ is given by

$$D\big(\frac{f}{g}\big)(x_0)(v) = \frac{g(x_0)\big((Df)(x_0)(v)\big) - (Dg)(x_0)(v)f(x_0)}{g^2(x_0)}\,. \tag{A.3}$$

The Chain Rule

Theorem A.25. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ and $\mathcal{V} \subseteq \mathbb{R}^m$ be open sets. Suppose that $f : \mathcal{U} \to \mathbb{R}^m$ is differentiable at $x_0 \in \mathcal{U}$, $f(\mathcal{U}) \subseteq \mathcal{V}$, and $g : \mathcal{V} \to \mathbb{R}^\ell$ is differentiable at $f(x_0)$. Then the map $F = g \circ f : \mathcal{U} \to \mathbb{R}^\ell$ defined by*

$$F(x) = g\big(f(x)\big) \qquad \forall\, x \in \mathcal{U}$$

is differentiable at x_0, and

$$(DF)(x_0)(h) = (Dg)\big(f(x_0)\big)\big((Df)(x_0)(h)\big)$$

or equivalently,

$$\big((DF)(x_0)\big)_{ij} = \sum_{k=1}^{m} \frac{\partial g_i}{\partial y_k}\big(f(x_0)\big)\frac{\partial f_k}{\partial x_j}(x_0)\,.$$

The Mean Value Theorem

Theorem A.26. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, and $f : \mathcal{U} \to \mathbb{R}^m$ with $f = (f_1, \cdots, f_m)$. Suppose that f is differentiable on $\mathcal{U}$ and the line segment joining x and y lies in $\mathcal{U}$. Then there exist points $c_1, \cdots, c_m$ on that segment such that*

$$f_i(y) - f_i(x) = (Df_i)(c_i)(y - x) \qquad \forall\, i = 1, \cdots, m.$$

Corollary A.27. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open and convex, and $f : \mathcal{U} \to \mathbb{R}^m$ be differentiable. Then for all $x, y \in \mathcal{U}$, there exists $c_1, \cdots, c_m$ on $\overline{xy}$ such that*

$$f_i(y) - f_i(x) = (Df_i)(c_i)(y - x).$$

THEOREM A.28. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, $K \subseteq \mathcal{U}$ be compact, and $f : \mathcal{U} \to \mathbb{R}$ be of class $\mathscr{C}^1$. Then for each $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$\big|f(y) - f(x) - (Df)(x)(y-x)\big| \leqslant \varepsilon\|y-x\|_{\mathbb{R}^n} \quad \text{if} \;\; \|y-x\|_{\mathbb{R}^n} < \delta \text{ and } x, y \in K\,.$$

COROLLARY A.29. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, $K \subseteq \mathcal{U}$ be compact, and $f : \mathcal{U} \to \mathbb{R}^m$ be of class $\mathscr{C}^1$. Then for each $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$\big\|f(y) - f(x) - (Df)(x)(y-x)\big\|_{\mathbb{R}^m} \leqslant \varepsilon\|y-x\|_{\mathbb{R}^n} \quad \text{if} \;\; \|y-x\|_{\mathbb{R}^n} < \delta \text{ and } x, y \in K\,.$$

A.1.4 Conditions for Differentiability

PROPOSITION A.30. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, $a \in \mathcal{U}$, and $f = (f_1, \cdots, f_m) : \mathcal{U} \to \mathbb{R}^m$. Then f is differentiable at a if and only if f_i is differentiable at a for all $i = 1, \cdots, m$. In other words, for vector-valued functions defined on an open subset of $\mathbb{R}^n$,*

$$\textit{Componentwise differentiable} \;\Leftrightarrow\; \textit{Differentiable.}$$

THEOREM A.31. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, $a \in \mathcal{U}$, and $f : \mathcal{U} \to \mathbb{R}$. If each entry of the Jacobian matrix $\left[\dfrac{\partial f}{\partial x_1} \cdots \dfrac{\partial f}{\partial x_n}\right]$ of f*

1. *exists in a neighborhood of a, and*

2. *is continuous at a except perhaps one entry.*

Then f is differentiable at a.

DEFINITION A.32. Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, and $f : \mathcal{U} \to \mathbb{R}^m$ be differentiable on $\mathcal{U}$. f is said to be ***continuously differentiable*** on $\mathcal{U}$ if $Df : \mathcal{U} \to \mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)$ is continuous on $\mathcal{U}$. The collection of all continuously differentiable mappings from $\mathcal{U}$ to $\mathbb{R}^m$ is denoted by $\mathscr{C}^1(\mathcal{U}; \mathbb{R}^m)$. The collection of all bounded differentiable functions from $\mathcal{U}$ to $\mathbb{R}^m$ whose derivative is continuous and bounded is denoted by $\mathscr{C}_b^1(\mathcal{U}; \mathbb{R}^m)$. In other words,

$$\mathscr{C}^1(\mathcal{U}; \mathbb{R}^m) = \big\{f : \mathcal{U} \to \mathbb{R}^m \text{ is differentiable} \,\big|\, Df : \mathcal{U} \to \mathscr{B}(\mathbb{R}^n, \mathbb{R}^m) \text{ is continuous}\big\}$$

and

$$\mathscr{C}_b^1(\mathcal{U}; \mathbb{R}^m) = \Big\{f \in \mathscr{C}^1(\mathcal{U}; \mathbb{R}^m) \,\Big|\, \sup_{x \in \mathcal{U}} |f(x)| + \sup_{x \in \mathcal{U}} \|Df(x)\|_{\mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)} < \infty\Big\}\,.$$

Corollary A.33. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, and $f : \mathcal{U} \to \mathbb{R}^m$. Then $f \in \mathscr{C}^1(\mathcal{U};\mathbb{R}^m)$ if and only if the partial derivatives $\dfrac{\partial f_i}{\partial x_j}$ exist and are continuous on $\mathcal{U}$ for $i = 1, \cdots, m$ and $j = 1, \cdots, n$.*

Proposition A.34. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open. Given $f \in \mathscr{C}_b^1(\mathcal{U};\mathbb{R}^m)$, define*

$$\|f\|_{\mathscr{C}_b^1(\mathcal{U};\mathbb{R}^m)} = \sup_{x \in \mathcal{U}} \Big[|f(x)| + \sum_{i=1}^{m} \sum_{j=1}^{n} \Big| \frac{\partial f_i}{\partial x_j}(x) \Big| \Big] .$$

Then $\big(\mathscr{C}_b^1(\mathcal{U};\mathbb{R}^m), \|\cdot\|_{\mathscr{C}_b^1(\mathcal{U};\mathbb{R}^m)}\big)$ is a Banach space.

Proof. Left as an exercise. ◻

Definition A.35. Let f be real-valued and defined on a neighborhood of $x_0 \in \mathbb{R}^n$, and let $\mathrm{v} \in \mathbb{R}^n$ be a unit vector. Then

$$(D_{\mathrm{v}} f)(x_0) \equiv \frac{d}{dt}\Big|_{t=0} f(x_0 + t\mathrm{v}) = \lim_{t \to 0} \frac{f(x_0 + t\mathrm{v}) - f(x_0)}{t}$$

is called the ***directional derivative*** of f at x_0 in the direction v.

Remark A.36. Let $\{\mathrm{e}_j\}_{j=1}^n$ be the standard basis of $\mathbb{R}^n$. Then the partial derivative $\dfrac{\partial f}{\partial x_j}(x_0)$ (if it exists) is the directional derivative of f at x_0 in the direction e_j.

Theorem A.37. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, and $f : \mathcal{U} \to \mathbb{R}$ be differentiable at x_0. Then the directional derivative of f at x_0 in the direction v is $(Df)(x_0)(\mathrm{v})$.*

Proof. Let $\varepsilon > 0$ be given. Since f is differentiable at x_0, there exists $\delta > 0$ such that

$$\big|f(x) - f(x_0) - (Df)(x_0)(x - x_0)\big| \leqslant \frac{\varepsilon}{2} \|x - x_0\|_{\mathbb{R}^n} \text{ whenever } \|x - x_0\|_{\mathbb{R}^n} < \delta \,.$$

In particular, if $x = x_0 + t\mathrm{v}$ with v being a unit vector in $\mathbb{R}^n$ and $0 < |t| < \delta$, then

$$\begin{aligned} \Big| \frac{f(x_0 + t\mathrm{v}) - f(x_0)}{t} - (Df)(x_0)(\mathrm{v}) \Big| &= \frac{\big|f(x_0 + t\mathrm{v}) - f(x_0) - (Df)(x_0)(t\mathrm{v})\big|}{|t|} \\ &= \frac{\big|f(x) - f(x_0) - (Df)(x_0)(x - x_0)\big|}{|t|} \leqslant \frac{\varepsilon}{2} < \varepsilon \,; \end{aligned}$$

thus $(D_{\mathrm{v}} f)(x_0) = (Df)(x_0)(\mathrm{v})$. ◻

Remark A.38. When $v \in \mathbb{R}^n$ but $0 < \|v\|_{\mathbb{R}^n} \neq 1$, we let $\mathrm{v} = \dfrac{v}{\|v\|_{\mathbb{R}^n}}$. Then the direction derivatives of a function $f : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$ at $a \in \mathcal{U}$ in the direction v is

$$(D_{\mathrm{v}} f)(a) = \lim_{t \to 0} \frac{f(a + t\mathrm{v}) - f(a)}{t}\,.$$

Making a change of variable $s = \dfrac{t}{\|v\|_{\mathbb{R}^n}}$. Then

$$(Df)(x_0)(v) = \|v\|_{\mathbb{R}^n}(Df)(x_0)(\mathrm{v}) = \|v\|_{\mathbb{R}^n} \lim_{t \to 0} \frac{f(a + t\mathrm{v}) - f(a)}{t} = \lim_{s \to 0} \frac{f(a + sv) - f(a)}{s}\,.$$

We sometimes also call the value $(Df)(x_0)(v)$ the "directional derivative" of f in the "direction" v.

A.1.5 Higher Derivatives of Functions

Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, and $f : \mathcal{U} \to \mathbb{R}^m$ is differentiable. By Proposition A.7, the space $\big(\mathscr{B}(\mathbb{R}^n, \mathbb{R}^m), \|\cdot\|_{\mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)}\big)$ is a normed space (in fact, it is a Banach space), so it is legitimate to ask if $Df : \mathcal{U} \to \mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)$ is differentiable or not. If Df is differentiable at x_0, we call f twice differentiable at x_0, and denote the twice derivative of f at x_0 as $(D^2 f)(x_0)$. If Df is differentiable on $\mathcal{U}$, then $D^2 f : \mathcal{U} \to \mathscr{B}\big(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n, \mathbb{R}^m)\big)$. Similar, we can introduce three times differentiability of a function if it is twice differentiable. In general, we have the following

Definition A.39. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces, and $\mathcal{U} \subseteq X$ be open. A function $f : \mathcal{U} \to Y$ is said to be ***twice differentiable*** at $a \in \mathcal{U}$ if

1. f is (once) differentiable in a neighborhood of a;

2. there exists $L_2 \in \mathscr{B}\big(X, \mathscr{B}(X, Y)\big)$, usually denoted by $(D^2 f)(a)$ and called the ***second derivative*** of f at a, such that

$$\lim_{x \to a} \frac{\big\|(Df)(x) - (Df)(a) - L_2(x - a)\big\|_{\mathscr{B}(X,Y)}}{\|x - a\|_X} = 0\,.$$

For any two vectors $u, v \in X$, $(D^2 f)(a)(v) \in \mathscr{B}(X, Y)$ and $(D^2 f)(a)(v)(u) \in Y$. The vector $(D^2 f)(a)(v)(u)$ is usually denoted by $(D^2 f)(a)(u, v)$.

In general, a function f is said to be ***k-times differentiable*** at $a \in \mathcal{U}$ if

1. f is $(k-1)$-times differentiable in a neighborhood of a;
2. there exists $L_k \in \mathscr{B}(\underbrace{X, \mathscr{B}(X, \cdots, \mathscr{B}(X}_{k \text{ copies of "}X\text{"}}, Y\underbrace{) \cdots))}_{k \text{ copies of ")"}}$, usually denoted by $(D^k f)(a)$ and called the ***k-th derivative*** of f at a, such that
$$\lim_{x \to a} \frac{\big\|(D^{k-1}f)(x) - (D^{k-1}f)(a) - L_k(x-a)\big\|_{\mathscr{B}(X, \mathscr{B}(X, \cdots, \mathscr{B}(X,Y)\cdots))}}{\|x-a\|_X} = 0\,.$$

For any k vectors $u^{(1)}, \cdots u^{(k)} \in X$, the vector $(D^k f)(a)(u^{(1)}, \cdots, u^{(k)})$ is defined as the vector
$$(D^k f)(a)(u^{(k)})(u^{(k-1)}) \cdots (u^{(1)})\,.$$

Remark A.40. We focus on what $(D^k f)(a)(u_k)(\cdots)(u_1)$ means in this remark. We first look at the case that f is twice differentiable at a. With $x = a + tv$ for $v \in X$ with $\|v\|_X = 1$ in the definition, we find that
$$\lim_{t \to 0} \frac{\big\|(Df)(a+tv) - (Df)(a) - t(D^2 f)(a)(v)\big\|_{\mathscr{B}(X,Y)}}{|t|} = 0\,.$$
Since $(Df)(a+tv) - (Df)(a) - t(D^2 f)(a)(v) \in \mathscr{B}(X,Y)$, for all $u \in X$ with $\|u\|_X = 1$ we have
$$\begin{aligned}
&\lim_{t \to 0} \frac{\big\|(Df)(a+tv)(u) - (Df)(a)(u) - t(D^2 f)(a)(v)(u)\big\|_Y}{|t|} \\
&\qquad = \lim_{t \to 0} \frac{\big\|\big[(Df)(a+tv) - (Df)(a) - t(D^2 f)(a)(v)\big](u)\big\|_Y}{|t|} \\
&\qquad \leqslant \lim_{t \to 0} \frac{\big\|(Df)(a+tv) - (Df)(a) - t(D^2 f)(a)(v)\big\|_{\mathscr{B}(X,Y)}}{|t|} = 0\,.
\end{aligned}$$
On the other hand, by the definition of the direction derivative,
$$(Df)(a+tv)(u) - (Df)(a)(u) = \lim_{s \to 0} \Big[\frac{f(a+tv+su) - f(a+tv)}{s} - \frac{f(a+su) - f(a)}{s}\Big];$$
thus the limit above implies that
$$\begin{aligned}
(D^2 f)(a)(v)(u) &= \lim_{t \to 0} \lim_{s \to 0} \frac{f(a+tv+su) - f(a+tv) - f(a+su) + f(a)}{st} \\
&= \lim_{t \to 0} \frac{\displaystyle\lim_{s \to 0} \frac{f(a+tv+su) - f(a+tv)}{s} - \lim_{s \to 0} \frac{f(a+su) - f(a)}{s}}{t} \\
&= D_v(D_u f)(a)\,.
\end{aligned}$$

Therefore, $(D^2f)(a)(v)(u)$ is obtained by first differentiating f near a in the u-direction, then differentiating (Df) at a in the v-direction.

In general, $(D^kf)(a)(u_k)\cdots(u_1)$ is obtained by first differentiating f near a in the u_1-direction, then differentiating (Df) near a in the u_2-direction, and so on, and finally differentiating $(D^{k-1}f)$ at a in the u_k-direction.

Remark A.41. Since $(D^2f)(a) \in \mathscr{B}(X, \mathscr{B}(X, Y))$, if $v_1, v_2 \in X$ and $c \in \mathbb{R}$, we have $(D^2f)(a)(cv_1 + v_2) = c(D^2f)(a)(v_1) + (D^2f)(a)(v_2)$ (treated as "vectors" in $\mathscr{B}(X, Y)$); thus

$$(D^2f)(a)(cv_1 + v_2)(u) = c(D^2f)(a)(v_1)(u) + (D^2f)(a)(v_2)(u) \qquad \forall\, u, v_1, v_2 \in X\,.$$

On the other hand, since $(D^2f)(a)(v) \in \mathscr{B}(X, Y)$,

$$(D^2f)(a)(v)(cu_1 + u_2) = c(D^2f)(a)(v)(u_1) + (D^2f)(a)(v)(u_2) \qquad \forall\, u_1, u_2, v \in X\,.$$

Therefore, $(D^2f)(a)(v)(u)$ is linear in both u and v variables. A map with such kind of property is called a ***bilinear*** map (meaning 2-linear). In particular, $(D^2f)(a) : X \times X \to Y$ is a bilinear map.

In general, the vector $(D^kf)(a)(u^{(1)}, \cdots, u^{(k)})$ is linear in $u^{(1)}, \cdots, u^{(k)}$; that is,

$$\begin{aligned}&(D^kf)(a)(u^{(1)}, \cdots, u^{(i-1)}, \alpha v + \beta w, u^{(i+1)}, \cdots, u^{(k)})\\ &\qquad = \alpha(D^kf)(a)(u^{(1)}, \cdots, u^{(i-1)}, v, u^{(i+1)}, \cdots, u^{(k)})\\ &\qquad\quad + \beta(D^kf)(a)(u^{(1)}, \cdots, u^{(i-1)}, w, u^{(i+1)}, \cdots, u^{(k)})\end{aligned}$$

for all $v, w \in X$, $\alpha, \beta \in \mathbb{R}$, and $i = 1, \cdots, n$. Such kind of map which is linear in each component when the other $k-1$ components are fixed is called k-***linear***.

Consider the case that X is finite dimensional with $\dim(X) = n$, $\{e_1, e_2, \ldots, e_n\}$ is a basis of X, and $Y = \mathbb{R}$. Then $(D^2f)(a) : X \times X \to Y$ is a bilinear form (here the term "form" means that $Y = \mathbb{R}$). A bilinear form $B : X \times X \to \mathbb{R}$ can be represented as follows: Let $a_{ij} = B(e_i, e_j) \in \mathbb{R}$ for $i, j = 1, 2, \cdots, n$. Given $x, y \in \mathbb{R}^n$, write $u = \sum\limits_{i=1}^{n} u_i e_i$ and $v = \sum\limits_{j=1}^{n} v_j e_j$. Then by the bilinearity of B,

$$B(u, v) = B\Big(\sum_{i=1}^{n} u_i e_i, \sum_{j=1}^{n} v_j e_j\Big) = \sum_{i,j=1}^{n} u_i v_j a_{ij} = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}.$$

Therefore, if $f : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}$ is twice differentiable at a, then the bilinear form $(D^2 f)(a)$ can be represented as

$$(D^2 f)(a)(u, v) = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \begin{bmatrix} (D^2 f)(\mathrm{e}_1, \mathrm{e}_1) & \cdots & (D^2 f)(a)(\mathrm{e}_1, \mathrm{e}_n) \\ \vdots & \ddots & \vdots \\ (D^2 f)(\mathrm{e}_n, \mathrm{e}_1) & \cdots & (D^2 f)(a)(\mathrm{e}_n, \mathrm{e}_n) \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}.$$

The following proposition is an analogue of Proposition A.30. The proof is similar to the one of Proposition A.30, and is left as an exercise.

Proposition A.42. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, $x_0 \in \mathcal{U}$, and $f = (f_1, \cdots, f_m) : \mathcal{U} \to \mathbb{R}^m$. Then f is k-times differentiable at x_0 if and only if f_i is k-times differentiable at x_0 for all $i = 1, \cdots, m$.*

Due to the proposition above, when talking about the higher-order differentiability of $f : \mathcal{U} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and a point $x_0 \in \mathcal{U}$, from now on we only focus on the case $m = 1$.

Proposition A.43. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, and $f : \mathcal{U} \to \mathbb{R}$. Suppose that f is k-times differentiable at a. Then for k vectors $u^{(1)}, \cdots, u^{(k)} \in \mathbb{R}^n$,*

$$(D^k f)(a)(u^{(1)}, \cdots, u^{(k)}) = \sum_{j_1, \cdots, j_k = 1}^{n} \frac{\partial^k f}{\partial x_{j_k} \partial x_{j_{k-1}} \cdots \partial x_{j_1}}(a) u_{j_1}^{(1)} u_{j_2}^{(2)} \cdots u_{j_k}^{(k)},$$

where $u^{(i)} = (u_1^{(i)}, u_2^{(i)}, \cdots, u_n^{(i)})$ for all $i = 1, \cdots, k$.

Proof. Let $\{\mathrm{e}_j\}_{j=1}^n$ be the standard basis of $\mathbb{R}^n$. By Remark A.41 (on multi-linearity), it suffices to show that

$$(D^k f)(a)(\mathrm{e}_{j_1}, \cdots, \mathrm{e}_{j_k}) = \frac{\partial^k f}{\partial x_{j_k} \partial x_{j_{k-1}} \cdots \partial x_{j_1}}(a) \tag{A.4}$$

since if so, we must have

$$\begin{aligned} (D^k f)(a)(u^{(1)}, \cdots, u^{(k)}) &= (D^k f)(a)\Big(\sum_{j_1=1}^{n} u_{j_1}^{(1)} \mathrm{e}_{j_1}, \cdots, \sum_{j_k=1}^{n} u_{j_k}^{(k)} \mathrm{e}_{j_k} \Big) \\ &= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_k=1}^{n} (D^k f)(a)(\mathrm{e}_{j_1}, \cdots, \mathrm{e}_{j_k}) u_{j_1}^{(1)} u_{j_2}^{(2)} \cdots u_{j_k}^{(k)} \\ &= \sum_{j_1, \cdots, j_k = 1}^{n} \frac{\partial^k f}{\partial x_{j_k} \partial x_{j_{k-1}} \cdots \partial x_{j_1}}(a) u_{j_1}^{(1)} u_{j_2}^{(2)} \cdots u_{j_k}^{(k)}. \end{aligned}$$

We prove the proposition by induction. Note that the case $k = 1$ is true because of Theorem A.17. Next we assume that (A.4) holds true for $k = \ell$ if f is $(\ell - 1)$-times differentiable in a neighborhood of a and f is ℓ-times differentiable at a. Now we show that (A.4) also holds true for $k = \ell + 1$ if f is ℓ-times differentiable in a neighborhood of a, and f is $(\ell + 1)$-times differentiable at a. By the definition of $(\ell + 1)$-times differentiability at a,

$$\lim_{x\to a} \frac{\left\|(D^\ell f)(x) - (D^\ell f)(a) - (D^{\ell+1} f)(a)(x-a)\right\|_{\mathscr{B}(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n, \cdots, \mathscr{B}(\mathbb{R}^n, \mathbb{R})\cdots))}}{\|x-a\|_{\mathbb{R}^n}} = 0\,.$$

Since

$$\begin{aligned}
&\Big|\big[(D^\ell f)(x) - (D^\ell f)(a) - (D^{\ell+1} f)(a)(x-a)\big](\mathrm{e}_{j_\ell})\cdots(\mathrm{e}_{j_2})(\mathrm{e}_{j_1})\Big| \\
&\qquad \leqslant \Big\|\big[(D^\ell f)(x) - (D^\ell f)(a) - (D^{\ell+1} f)(a)(x-a)\big](\mathrm{e}_{j_\ell})\cdots(\mathrm{e}_{j_2})\Big\|_{\mathscr{B}(\mathbb{R}^n, \mathbb{R})} \\
&\qquad \leqslant \cdots\cdots \\
&\qquad \leqslant \left\|(D^\ell f)(x) - (D^\ell f)(a) - (D^{\ell+1} f)(a)(x-a)\right\|_{\mathscr{B}(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n, \cdots, \mathscr{B}(\mathbb{R}^n, \mathbb{R})\cdots))},
\end{aligned}$$

using (A.4) (for the case $k = \ell$) we conclude that

$$\begin{aligned}
&\lim_{x\to a} \frac{\left|\dfrac{\partial^\ell f}{\partial x_{j_\ell}\partial x_{j_{k-1}}\cdots\partial x_{j_1}}(x) - \dfrac{\partial^\ell f}{\partial x_{j_\ell}\partial x_{j_{k-1}}\cdots\partial x_{j_1}}(a) - (D^{\ell+1} f)(a)(\mathrm{e}_{j_1}, \cdots, \mathrm{e}_{j_\ell}, x-a)\right|}{\|x-a\|_{\mathbb{R}^n}} \\
&\quad = \lim_{x\to a} \frac{\left|(D^\ell f)(x)(\mathrm{e}_{j_1}, \cdots, \mathrm{e}_{j_\ell}) - (D^\ell f)(a)(\mathrm{e}_{j_1}, \cdots, \mathrm{e}_{j_\ell}) - (D^{\ell+1} f)(a)(x-a)(\mathrm{e}_{j_1}, \cdots, \mathrm{e}_{j_\ell})\right|}{\|x-a\|_{\mathbb{R}^n}} \\
&\quad \leqslant \lim_{x\to a} \frac{\left\|(D^\ell f)(x) - (D^\ell f)(a) - (D^{\ell+1} f)(a)(x-a)\right\|_{\mathscr{B}(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n, \cdots, \mathscr{B}(\mathbb{R}^n, \mathbb{R})\cdots))}}{\|x-a\|_{\mathbb{R}^n}} = 0\,.
\end{aligned}$$

In particular, if $x = a + t\mathrm{e}_{j_{\ell+1}}$ for some $j_{\ell+1} = 1, \cdots, n$, by the definition of partial derivatives we conclude that

$$\begin{aligned}
&(D^{\ell+1} f)(a)(\mathrm{e}_{j_1}, \cdots, \mathrm{e}_{j_\ell}, \mathrm{e}_{j_{\ell+1}}) \\
&\qquad = \lim_{t\to 0} \frac{\dfrac{\partial^\ell f}{\partial x_{j_\ell}\partial x_{j_{k-1}}\cdots\partial x_{j_1}}(a + t\mathrm{e}_{j_{\ell+1}}) - \dfrac{\partial^\ell f}{\partial x_{j_\ell}\partial x_{j_{k-1}}\cdots\partial x_{j_1}}(a)}{t} \\
&\qquad = \frac{\partial^{\ell+1} f}{\partial x_{j_{\ell+1}}\partial x_{j_\ell}\partial x_{j_{k-1}}\cdots\partial x_{j_1}}(a)
\end{aligned}$$

which is (A.4) for the case $k = \ell + 1$. □

Example A.44. Let $f:\mathbb{R}^2\to\mathbb{R}$ be given by $f(x_1,x_2)=x_1^2\cos x_2$, and $u^{(1)}=(2,0)$, $u^{(2)}=(1,1)$, $u^{(3)}=(0,-1)$. Suppose that f is three-times differentiable at $a=(0,0)$ (in fact it is, but we have not talked about this yet). Then

$$\begin{aligned}
&(D^3f)(a)(u^{(1)},u^{(2)},u^{(3)})\\
&\quad=\sum_{i,j,k=1}^{2}\frac{\partial^3 f}{\partial x_k\partial x_j\partial x_i}(a)u_i^{(1)}u_j^{(2)}u_k^{(3)}=\sum_{j=1}^{2}\frac{\partial^3 f}{\partial x_2\partial x_j\partial x_1}(a)\cdot 2\cdot u_j^{(2)}\cdot(-1)\\
&\quad=\frac{\partial^3 f}{\partial x_2\partial x_1^2}(0,0)\cdot 2\cdot 1\cdot(-1)+\frac{\partial^3 f}{\partial x_2^2\partial x_1}(0,0)\cdot 2\cdot 1\cdot(-1)=0\,.
\end{aligned}$$

Example A.45. Let $f:\mathbb{R}^2\to\mathbb{R}$ be twice differentiable at $a=(a_1,a_2)\in\mathbb{R}^2$. Then the proposition above suggests that for $u=(u_1,u_2),v=(v_1,v_2)\in\mathbb{R}^2$,

$$\begin{aligned}
(D^2f)(a)(v)(u)=(D^2f)(a)(u,v)&=\sum_{i,j=1}^{2}\frac{\partial^2 f}{\partial x_j\partial x_i}(a)u_iv_j\\
&=\frac{\partial^2 f}{\partial x_1^2}(a)u_1v_1+\frac{\partial^2 f}{\partial x_2\partial x_1}(a)u_1v_2+\frac{\partial^2 f}{\partial x_1\partial x_2}(a)u_2v_1+\frac{\partial^2 f}{\partial x_2^2}(a)u_2v_2\\
&=\begin{bmatrix}u_1 & u_2\end{bmatrix}\begin{bmatrix}\dfrac{\partial^2 f}{\partial x_1^2}(a) & \dfrac{\partial^2 f}{\partial x_2\partial x_1}(a)\\[2ex] \dfrac{\partial^2 f}{\partial x_1\partial x_2}(a) & \dfrac{\partial^2 f}{\partial x_2^2}(a)\end{bmatrix}\begin{bmatrix}v_1\\ v_2\end{bmatrix}.
\end{aligned}$$

In general, if $f:\mathbb{R}^n\to\mathbb{R}$ be twice differentiable at $a=(a_1,\cdots,a_n)\in\mathbb{R}^n$. Then for $u=(u_1,\cdots,u_n),v=(v_1,\cdots,v_n)\in\mathbb{R}^2$

$$(D^2f)(a)(v)(u)=\begin{bmatrix}u_1 & \cdots & u_n\end{bmatrix}\begin{bmatrix}\dfrac{\partial^2 f}{\partial x_1^2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_n\partial x_1}(a)\\ \vdots & \ddots & \vdots\\ \dfrac{\partial^2 f}{\partial x_1\partial x_n}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}(a)\end{bmatrix}\begin{bmatrix}v_1\\ \vdots\\ v_n\end{bmatrix}.$$

The bilinear form $B:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ given by

$$B(u,v)=(D^2f)(a)(v)(u)\qquad\forall\,u,v\in\mathbb{R}^n$$

is called the ***Hessian*** of f, and is represented (in the matrix form) as an $n\times n$ matrix by

$$\begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_1}(a) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}(a) \end{bmatrix}.$$

If the second partial derivatives $\dfrac{\partial^2 f}{\partial x_j \partial x_i}(a)$ of f at a exists for all $i, j = 1, \cdots, n$ (here the twice differentiability of f at a is ignored), the matrix (on the right-hand side of equality) above is also called the ***Hessian matrix*** of f at a.

Even though there is no reason to believe that $(D^2 f)(a)(u, v) = (D^2 f)(a)(v, u)$ (since the left-hand side means first differentiating f in u-direction and then differentiating Df in v-direction, while the right-hand side means first differentiating f in v-direction then differentiating Df in u-direction), it is still reasonable to ask whether $(D^2 f)(a)$ is symmetric or not; that is, could it be true that $(D^2 f)(a)(u, v) = (D^2 f)(a)(v, u)$ for all $u, v \in \mathbb{R}^n$? When f is twice differentiable at a, this is equivalent of asking (by plugging in $u = \mathrm{e}_i$ and $v = \mathrm{e}_j$) that whether or not

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(a) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\,. \tag{A.5}$$

The following example provides a function $f : \mathbb{R}^2 \to \mathbb{R}$ such that (A.5) does not hold at $a = (0, 0)$. We remark that the function in the following example is not twice differentiable at a even though the Hessian matrix of f at a can still be computed.

Example A.46. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

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

Then

$$f_x(x, y) = \begin{cases} \dfrac{x^4 y + 4x^2 y^3 - y^5}{(x^2 + y^2)^2} & \text{if } (x, y) \neq (0, 0)\,, \\ 0 & \text{if } (x, y) = (0, 0)\,, \end{cases}$$

and

$$f_y(x, y) = \begin{cases} \dfrac{x^5 - 4x^3 y^2 - xy^4}{(x^2 + y^2)^2} & \text{if } (x, y) \neq (0, 0)\,, \\ 0 & \text{if } (x, y) = (0, 0)\,, \end{cases}$$

It is clear that f_x and f_y are continuous on $\mathbb{R}^2$; thus f is differentiable on $\mathbb{R}^2$. However,

$$f_{xy}(0,0) = \lim_{k\to 0} \frac{f_x(0,k) - f_x(0,0)}{k} = -1\,,$$

while

$$f_{yx}(0,0) = \lim_{h\to 0} \frac{f_y(h,0) - f_y(0,0)}{h} = 1\,;$$

thus the Hessian matrix of f at the origin is not symmetric.

Definition A.47. A function is said to be ***of class*** $\mathscr{C}^r$ if the first r derivatives exist and are continuous. A function is said to be ***smooth*** or ***of class*** $\mathscr{C}^\infty$ if it is of class $\mathscr{C}^r$ for all positive integer r.

The following theorem is an analogue of Corollary A.33.

Theorem A.48. *Let* $\mathcal{U} \to \mathbb{R}^n$ *and* $f : \mathcal{U} \to \mathbb{R}$. *Suppose that the partial derivative* $\dfrac{\partial^k f}{\partial x_{j_k}\partial x_{j_{k-1}}\cdots\partial x_{j_1}}$ *exists in a neighborhood of* $a \in \mathcal{U}$ *and is continuous at* a *for all* $j_1, \cdots, j_k = 1, \cdots, n$. *Then* f *is k-times differentiable at* a. *Moreover, if* $\dfrac{\partial^k f}{\partial x_{j_k}\partial x_{j_{k-1}}\cdots\partial x_{j_1}}$ *is continuous on* $\mathcal{U}$, *then* f *is of class* $\mathscr{C}^k$.

Theorem A.49. *Let* $\mathcal{U} \subseteq \mathbb{R}^n$ *be open, and* $f : \mathcal{U} \to \mathbb{R}$. *Suppose that the mixed partial derivatives* $\dfrac{\partial f}{\partial x_i}, \dfrac{\partial f}{\partial x_j}, \dfrac{\partial^2 f}{\partial x_j \partial x_i}, \dfrac{\partial^2 f}{\partial x_j \partial x_i}$ *exist in a neighborhood of* a, *and are continuous at* a. *Then*

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(a) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\,. \tag{A.6}$$

Proof. Let $S(a,h,k) = f(a + h\mathrm{e}_i + k\mathrm{e}_j) - f(a + h\mathrm{e}_i) - f(a + k\mathrm{e}_j) + f(a)$, and define $\varphi(x) = f(x + h\mathrm{e}_i) - f(x)$ as well as $\psi(x) = f(x + k\mathrm{e}_j) - f(x)$ for x in a neighborhood of a. Then $S(a,h,k) = \varphi(a + k\mathrm{e}_j) - \varphi(a) = \psi(a + h\mathrm{e}_i) - \psi(a)$; thus the mean value theorem implies that there exists c on the line segment joining a and $a + k\mathrm{e}_j$ and d on the line segment joining a and $a + h\mathrm{e}_i$ such that

$$S(a,h,k) = \varphi(a + k\mathrm{e}_j) - \varphi(a) = k\frac{\partial \varphi}{\partial x_j}(c) = k\Big(\frac{\partial f}{\partial x_j}(c + h\mathrm{e}_i) - \frac{\partial f}{\partial x_j}(c)\Big)\,,$$

$$S(a,h,k) = \psi(a + h\mathrm{e}_i) - \psi(a) = h\frac{\partial \psi}{\partial x_i}(d) = h\Big(\frac{\partial f}{\partial x_i}(d + k\mathrm{e}_j) - \frac{\partial f}{\partial x_i}(d)\Big)\,.$$

As a consequence, if $h \neq 0 \neq k$,

$$\frac{1}{k}\Big(\frac{\partial f}{\partial x_i}(d + k\mathrm{e}_j) - \frac{\partial f}{\partial x_i}(d)\Big) = \frac{S(a,h,k)}{hk} = \frac{1}{h}\Big(\frac{\partial f}{\partial x_j}(c + h\mathrm{e}_i) - \frac{\partial f}{\partial x_j}(c)\Big)$$

By the mean value theorem again, there exists c_1 and d_1 on the line segment joining c, $c + h\mathrm{e}_i$ and d, $d + k\mathrm{e}_j$, respectively, such that

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(d_1) = \frac{\partial^2 f}{\partial x_i \partial x_j}(c_1)\,.$$

The theorem is then concluded by the continuity of $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ and $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ at a, and $c_1 \to a$ and $d_1 \to a$ as $(h,k) \to (0,0)$. □

COROLLARY A.50. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, and f is of class $\mathscr{C}^2$. Then*

$$(D^2 f)(a)(u,v) = (D^2 f)(a)(v,u) \qquad \forall\, a \in \mathcal{U} \text{ and } u, v \in \mathbb{R}^n\,.$$

REMARK A.51. In view of Remark A.40, (A.6) is the same as the following identity

$$\begin{aligned}
&\lim_{h\to 0}\lim_{k\to 0} \frac{f(a + h\mathrm{e}_i + k\mathrm{e}_j) - f(a + h\mathrm{e}_i) - f(a + k\mathrm{e}_j) + f(a)}{hk} \\
&\qquad = \lim_{k\to 0}\lim_{h\to 0} \frac{f(a + h\mathrm{e}_i + k\mathrm{e}_j) - f(a + h\mathrm{e}_i) - f(a + k\mathrm{e}_j) + f(a)}{hk}
\end{aligned}$$

which implies that the order of the two limits $\lim\limits_{h\to 0}$ and $\lim\limits_{k\to 0}$ can be interchanged without changing the value of the limit (under certain conditions).

EXAMPLE A.52. Let $f(x,y) = yx^2\cos y^2$. Then

$$\begin{aligned}
f_{xy}(x,y) &= (2xy\cos y^2)_y = 2x\cos y^2 - 2xy(2y)\sin y^2 = 2x\cos y^2 - 4xy^2\sin y^2\,,\\
f_{yx}(x,y) &= (x^2\cos y^2 - yx^2(2y)\sin y^2)_x = (x^2\cos y^2 - 2x^2y^2\sin y^2)_x \\
&= 2x\cos y^2 - 4xy^2\sin y^2 = f_{xy}(x,y)\,.
\end{aligned}$$

A.1.6 The differentiation of the determinant and the Piola identity

THEOREM A.53. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, and for each $1 \leqslant i,j \leqslant n$, $a_{ij} : \mathcal{U} \to \mathbb{R}$ be differentiable functions. Define $A = [a_{ij}]$ and $J = \det(A)$. Then*

$$\frac{\partial J}{\partial x_k} = \mathrm{tr}\Big(\mathrm{Adj}(A)\frac{\partial A}{\partial x_k}\Big) \qquad \forall\, 1 \leqslant k \leqslant n\,, \tag{A.7}$$

where for a square matrix $M = [m_{ij}]$, $\mathrm{tr}(M)$ denotes the trace of M, $\mathrm{Adj}(M)$ denotes the adjoint matrix of M, and $\dfrac{\partial M}{\partial x_k}$ denotes the matrix whose (i,j)-th entry is given by $\dfrac{\partial m_{ij}}{\partial x_k}$.

Proof. By the property of determinant,

$$\frac{\partial J}{\partial x_k} = \begin{vmatrix} \frac{\partial a_{11}}{\partial x_k} & a_{12} & \cdots & a_{1n} \\ \frac{\partial a_{21}}{\partial x_k} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ \frac{\partial a_{n1}}{\partial x_k} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \frac{\partial a_{12}}{\partial x_k} & a_{13} & \cdots & a_{1n} \\ a_{21} & \frac{\partial a_{22}}{\partial x_k} & a_{23} & \cdots & a_{2n} \\ \vdots & & & & \vdots \\ a_{n1} & \frac{\partial a_{n2}}{\partial x_k} & a_{n3} & \cdots & a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} a_{11} & \cdots & a_{(n-1)1} & \frac{\partial a_{1n}}{\partial x_k} \\ a_{21} & \cdots & a_{(n-1)2} & \frac{\partial a_{2n}}{\partial x_k} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & a_{(n-1)n} & \frac{\partial a_{n1}}{\partial x_k} \end{vmatrix},$$

and (A.7) follows from the expansion of the determinant using the reduction-of-size recursive formula. □

Remark A.54. In general, let A be an n × n matrix-valued function, and δ be an operator satisfying $\delta(fg) = f\delta g + (\delta f)g$ whenever the product makes sense. Then

$$\delta \det(\mathrm{A}) = \mathrm{tr}\big(\mathrm{Adj}(\mathrm{A})\delta\mathrm{A}\big) = \det(\mathrm{A})\mathrm{tr}\big(\mathrm{A}^{-1}\delta\mathrm{A}\big), \tag{A.8}$$

where $\delta\mathrm{A} \equiv [\delta a_{ij}]_{\mathrm{n}\times\mathrm{n}}$ if $\mathrm{A} = [a_{ij}]_{\mathrm{n}\times\mathrm{n}}$.

Suppose that $\psi : \Omega \subseteq \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^{\mathrm{n}}$ is a given twice differentiable diffeomorphism (thus $\det(\nabla\psi) \neq 0$). Let $\mathrm{M} = [\nabla\psi]$, and $\mathrm{J} = \det(\mathrm{M})$. Then the adjoint matrix of M is JM^{-1}. With A denoting the inverse of $[\nabla\psi]$, Theorem A.53 implies that

$$\frac{\partial J}{\partial x_k} = \mathrm{tr}\Big(\mathrm{JM}^{-1}\frac{\partial \mathrm{M}}{\partial x_k}\Big) = \sum_{i,j=1}^{\mathrm{n}} \mathrm{JA}^j_i \frac{\partial \psi^i}{\partial x_j}, \tag{A.9}$$

Theorem A.55 (Piola's identity)**.** *Let $\psi : \Omega \subseteq \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^{\mathrm{n}}$ be of class $\mathscr{C}^2$ such that $\det(\nabla\psi) \neq 0$ in Ω, and $[a_{ij}]_{\mathrm{n}\times\mathrm{n}}$ be the adjoint matrix of $[\nabla\psi]$; that is, $a = \det([\nabla\psi])[\nabla\psi]^{-1}$. Then*

$$\sum_{j=1}^{\mathrm{n}} \frac{\partial}{\partial x_j} a_{ji} = 0. \tag{A.10}$$

Proof. Let $\mathrm{J} = \det([\nabla\psi])$ and $\mathrm{A} = [\nabla\psi]^{-1}$. Then $a_{ji} = \mathrm{JA}^j_i$. Moreover, since $\mathrm{A}\nabla\psi = \mathrm{I}_n$, $\displaystyle\sum_{r=1}^{\mathrm{n}} \mathrm{A}^j_r \frac{\partial \psi^r}{\partial x_s} = \delta_{js}$; thus

$$0 = \frac{\partial}{\partial x_k}\Big[\sum_{r=1}^{\mathrm{n}} \mathrm{A}^j_r \frac{\partial \psi^r}{\partial x_s}\Big] = \sum_{r=1}^{\mathrm{n}} \Big[\frac{\partial \mathrm{A}^j_r}{\partial x_k}\frac{\partial \psi^r}{\partial x_s} + \mathrm{A}^j_r \frac{\partial^2 \psi^r}{\partial x_k \partial x_s}\Big]$$

which, after multiplying the equality above by A^s_i and then summing over s, implies that

$$\frac{\partial \mathrm{A}^j_i}{\partial x_k} = -\sum_{r,s=1}^{\mathrm{n}} \mathrm{A}^j_r \frac{\partial^2 \psi^r}{\partial x_k \partial x_s} \mathrm{A}^s_i. \tag{A.11}$$

As a consequence,

$$\sum_{j=1}^{\mathrm{n}} \frac{\partial}{\partial x_j}(\mathrm{JA}_i^j) = \sum_{j=1}^{\mathrm{n}}\sum_{r,s=1}^{\mathrm{n}}\Big[\mathrm{JA}_s^r\frac{\partial^2\psi^s}{\partial x_j\partial x_r}\mathrm{A}_i^j - \mathrm{JA}_r^j\frac{\partial^2\psi^r}{\partial x_j\partial x_s}\mathrm{A}_i^s\Big] = 0\,,$$

where Theorem A.49 is applied to conclude the last equality. □

A.2 Integral Calculus

A.2.1 Integrable Functions

Definition A.56. Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a bounded function. For any partition

$$\mathcal{P} = \Big\{\Delta_{i_1 i_2\cdots i_n}\,\Big|\,\Delta_{i_1 i_2\cdots i_n} = [x_{i_1}^{(1)}, x_{i_1+1}^{(1)}]\times[x_{i_2}^{(2)}, x_{i_2+1}^{(2)}]\times\cdots\times[x_{i_n}^{(n)}, x_{i_{n+1}}^{(n+1)}],$$
$$i_k = 0, 1, \cdots, N_k - 1, k = 1, \cdots, n\Big\}\,,$$

the ***upper sum*** and the ***lower sum*** of f with respect to the partition $\mathcal{P}$, denoted by $U(f,\mathcal{P})$ and $L(f,\mathcal{P})$ respectively, are numbers defined by

$$U(f,\mathcal{P}) = \sum_{\Delta\in\mathcal{P}}\sup_{(x,y)\in\Delta}\overline{f}^A(x,y)\nu(\Delta)\,,$$
$$L(f,\mathcal{P}) = \sum_{\Delta\in\mathcal{P}}\inf_{(x,y)\in\Delta}\overline{f}^A(x,y)\nu(\Delta)\,,$$

where $\nu(\Delta)$ is the ***volume*** of the rectangle Δ given by

$$\nu(\Delta) = (x_{i_1+1}^{(1)} - x_{i_1}^{(1)})(x_{i_2+1}^{(2)} - x_{i_2}^{(2)})\cdots(x_{i_n+1}^{(n)} - x_{i_n}^{(n)})$$

if $\Delta = [x_{i_1}^{(1)} - x_{i_1+1}^{(1)}]\times[x_{i_2}^{(2)} - x_{i_2+1}^{(2)}]\times\cdots\times[x_{i_n}^{(n)} - x_{i_n+1}^{(n)}]$, and $\overline{f}^A$ is the extension of f by zero outside A given by

$$\overline{f}^A(x) = \begin{cases} f(x) & x\in A\,,\\ 0 & x\notin A\,. \end{cases} \tag{A.12}$$

The two numbers

$$\overline{\int_A} f(x)\,dx \equiv \inf\big\{U(f,\mathcal{P})\,\big|\,\mathcal{P}\text{ is a partition of }A\big\}\,,$$

and

$$\underline{\int_A} f(x)\,dx \equiv \sup\big\{L(f,\mathcal{P})\,\big|\,\mathcal{P} \text{ is a partition of } A\big\}$$

are called the ***upper integral*** and ***lower integral*** of f over A, respective. The function f is said to be ***Riemann*** (***Darboux***) ***integrable*** (over A) if $\overline{\int_A} f(x)\,dx = \underline{\int_A} f(x)\,dx$, and in this case, we express the upper and lower integral as $\int_A f(x)\,dx$, called the ***integral*** of f over A.

Definition A.57. A partition $\mathcal{P}'$ of a bounded set $A \subseteq \mathbb{R}^n$ is said to be a ***refinement*** of another partition $\mathcal{P}$ of A if for any $\Delta' \in \mathcal{P}'$, there is $\Delta \in \mathcal{P}$ such that $\Delta' \subseteq \Delta$. A partition $\mathcal{P}$ of a bounded set $A \subseteq \mathbb{R}^n$ is said to be the ***common refinement*** of another partitions $\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_k$ of A if

1. $\mathcal{P}$ is a refinement of $\mathcal{P}_j$ for all $1 \leqslant j \leqslant k$.
2. If $\mathcal{P}'$ is a refinement of $\mathcal{P}_j$ for all $1 \leqslant j \leqslant k$, then $\mathcal{P}'$ is also a refinement of $\mathcal{P}$.

In other words, $\mathcal{P}$ is a common refinement of $\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_k$ if it is the coarsest refinement.

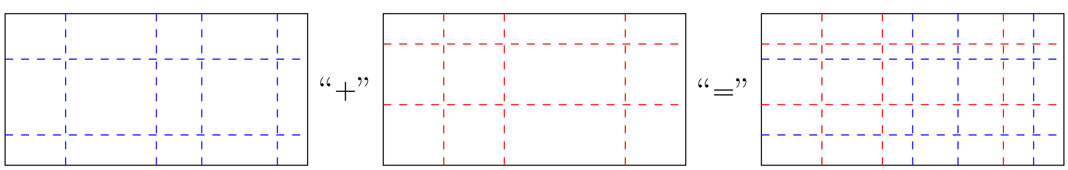


Figure A.1: The common refinement of two partitions

Qualitatively speaking, $\mathcal{P}$ is a common refinement of $\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_k$ if for each $j = 1, \cdots n$, the j-th component c_j of the vertex $(c_1, \cdots, c_n)$ of each rectangle $\Delta \in \mathcal{P}$ belongs to $\mathcal{P}_i^{(j)}$ for some $i = 1, \cdots, k$.

Proposition A.58. *Let $A \subseteq \mathbb{R}^n$ be a bounded subset, and $f : A \to \mathbb{R}$ be a bounded function. If $\mathcal{P}$ and $\mathcal{P}'$ are partitions of A and $\mathcal{P}'$ is a refinement of $\mathcal{P}$, then*

$$L(f,\mathcal{P}) \leqslant L(f,\mathcal{P}') \leqslant U(f,\mathcal{P}') \leqslant U(f,\mathcal{P})\,.$$

The following proposition provides a theoretical criteria for Riemann integrability of functions.

Proposition A.59 (Riemann's condition)**.** *Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable over A if and only if*

$$\forall\, \varepsilon > 0, \exists \text{ a partition } \mathcal{P} \text{ of } A \ni U(f,\mathcal{P}) - L(f,\mathcal{P}) < \varepsilon\,.$$

The following theorem provides an equivalent condition of Riemann integrability using Riemann sum approximation. The Riemann sum approximation is often useful in writing the limit of Riemann sums as Riemann integrals.

Theorem A.60 (Darboux)**.** *Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f : A \to \mathbb{R}$ be a bounded function with extension $\overline{f}^A$ given by* (A.12)*. Then f is Riemann integrable if and only if $\exists\, \mathrm{I} \in \mathbb{R}$ such that $\forall\, \varepsilon > 0$, $\exists\, \delta > 0 \ni$ if $\mathcal{P} = \{\Delta_1, \cdots, \Delta_N\}$ is a partition of A satisfying $\|\mathcal{P}\| < \delta$ and a set of sample points $\xi_1 \in \Delta_1$, $\xi_2 \in \Delta_2$, $\cdots$, $\xi_N \in \Delta_N$, we have*

$$\Big|\sum_{k=1}^{N} \overline{f}^A(\xi_{k+1})\nu(\Delta_k) - \mathrm{I}\Big| < \varepsilon\,. \tag{A.13}$$

The sum $\sum\limits_{k=1}^{N} \overline{f}^A(\xi_{k+1})\nu(\Delta_k)$ is called a ***Reimann sum*** *of f over A.*

Theorem A.61. *Let $A \subseteq \mathbb{R}^n$ be a bounded set, and $f_k : A \to \mathbb{R}$ be a sequence of Riemann integrable functions over A such that $\{f_k\}_{k=1}^{\infty}$ converges uniformly to f on A. Then f is Riemann integrable over A, and*

$$\lim_{k\to\infty} \int_A f_k(x)\, dx = \int_A f(x)\, dx\,. \tag{A.14}$$

From now on, we will simply use $\bar{f}$ to denote the zero extension of f when the domain outside which the zero extension is made is clear.

A.2.2 The Lebesgue Theorem

In this section, we discuss another equivalent condition of Riemann integrability, named the Lebesgue theorem. The Lebesque theorem provides a more practical way to check the Riemann integrability in the development of theory. To understand the Lebesgue theorem, we need to introduce a new concept: sets of measure zero.

Volume and Sets of Measure Zero

Definition A.62. Let $A \subseteq \mathbb{R}^n$ be a bounded set, and 1_A (or χ_A) be the characteristic function of A defined by

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

A is said to ***have volume*** if 1_A is Riemann integrable (over A), and the ***volume*** of A, denoted by $\nu(A)$, is the number $\displaystyle\int_A 1_A(x)\,dx$. A is said to have ***volume zero*** or ***content zero*** if $\nu(A) = 0$.

Remark A.63. Not all bounded set has volume.

Proposition A.64. *Let $A \subseteq \mathbb{R}^n$ be bounded. Then A has volume zero if and only if for every $\varepsilon > 0$, there exists finite (open) rectangles $S_1, \cdots, S_N$ (whose sides are parallel to the coordinate axes) such that*

$$A \subseteq \bigcup_{k=1}^{N} S_k \qquad \textit{and} \qquad \sum_{k=1}^{N} \nu(S_k) < \varepsilon\,.$$

Proof. "$\Rightarrow$" Since A has volume zero, $\displaystyle\overline{\int_A} 1_A(x)\,dx = 0$; thus for any given $\varepsilon > 0$, there exists a partition $\mathcal{P}$ of A such that

$$U(1_A, \mathcal{P}) < \overline{\int_A} 1_A(x)\,dx + \frac{\varepsilon}{2} = \frac{\varepsilon}{2}\,.$$

Since $\displaystyle\sup_{x\in\Delta} 1_A(x) = \begin{cases} 1 & \text{if } \Delta \cap A \neq \varnothing\,, \\ 0 & \text{otherwise}\,, \end{cases}$ we must have $\displaystyle\sum_{\substack{\Delta\in\mathcal{P} \\ \Delta\cap A\neq\varnothing}} \nu(\Delta) < \frac{\varepsilon}{2}$. Now if $\Delta \in \mathcal{P}$ and $\Delta \cap A \neq \varnothing$, we can find an open rectangle $\square$ such that $\Delta \subseteq \square$ and $\nu(\square) < 2\nu(\Delta)$. Let $S_1, \cdots, S_N$ be those open rectangles $\square$. Then $A \subseteq \displaystyle\bigcup_{k=1}^{N} S_k$ and $\displaystyle\sum_{k=1}^{N} \nu(S_k) < \varepsilon$.

"$\Leftarrow$" W.L.O.G. we can assume that the ratio of the maximum length and minimum length of sides of S_k is less than 2 for all $k = 1, \cdots, N$ (otherwise we can divide S_k into smaller rectangles so that each smaller rectangle satisfies this requirement). Then each S_k can be covered by a closed rectangle $\square_k$ whose sides are parallel to the coordinate axes with the property that $\nu(\square_k) \leqslant 2^{n-1}\sqrt{n}^{\,n}\nu(S_k)$. Let $\mathcal{P}$

be a partition of A such that for each $\Delta \in \mathcal{P}$ with $\Delta \cap A \neq \varnothing$, $\Delta \subseteq S_k$ for some $k = 1, \cdots, N$. Then

$$U(1_A, \mathcal{P}) = \sum_{\substack{\Delta \in \mathcal{P} \\ \Delta \cap A \neq \varnothing}} \nu(\Delta) \leqslant \sum_{k=1}^{N} \nu(\square_k) \leqslant 2^{n-1}\sqrt{n}^{\,n} \sum_{k=1}^{N} \nu(S_k) < 2^{n-1}\sqrt{n}^{\,n}\varepsilon\,;$$

thus the upper integral $\overline{\int_A} 1_A(x)\,dx = 0$. Since the lower integral cannot be negative, we must have $\overline{\int_A} 1_A(x)\,dx = \underline{\int_A} 1_A(x)\,dx = 0$ which implies that A has volume zero. □

Definition A.65. A set $A \subseteq \mathbb{R}^n$ (not necessarily bounded) is said to ***have measure zero*** or be ***a set of measure zero*** if for every $\varepsilon > 0$, there exist countable many rectangles $S_1, S_2, \cdots$ such that $\{S_k\}_{k=1}^{\infty}$ is a cover of A $\big($that is, $A \subseteq \bigcup\limits_{k=1}^{\infty} S_k\big)$ and $\sum\limits_{k=1}^{\infty} \nu(S_k) < \varepsilon$.

Proposition A.66. *Let $A \subseteq \mathbb{R}^n$ be a set of measure zero. If $B \subseteq A$, then B also has measure zero.*

Modifying the second part (or the "$\Leftarrow$" part) of the proof of Proposition A.64, we can also conclude the following

Proposition A.67. *A set $A \subseteq \mathbb{R}^n$ has measure zero if and only if for every $\varepsilon > 0$, there exist countable many open rectangles $S_1, S_2, \cdots$ whose sides are parallel to the coordinate axes such that $A \subseteq \bigcup\limits_{k=1}^{\infty} S_k$ and $\sum\limits_{k=1}^{\infty} \nu(S_k) < \varepsilon$.*

Remark A.68. If a set A has volume zero, then it has measure zero.

Proposition A.69. *Let $K \subseteq \mathbb{R}^n$ be a compact set of measure zero. Then K has volume zero.*

Proof. Let $\varepsilon > 0$ be given. Then there are countable open rectangles $S_1, S_2, \cdots$ such that

$$K \subseteq \bigcup_{k=1}^{\infty} S_k \qquad \text{and} \qquad \sum_{k=1}^{\infty} \nu(S_k) < \varepsilon\,.$$

Since $\{S_k\}_{k=1}^{\infty}$ is an open cover of K, by the compactness of K there exists $N > 0$ such that $K \subseteq \bigcup\limits_{k=1}^{N} S_k$, while $\sum\limits_{k=1}^{N} \nu(S_k) \leqslant \sum\limits_{k=1}^{\infty} \nu(S_k) < \varepsilon$. As a consequence, K has volume zero. □

Since the boundary of a rectangle has measure zero, we also have the following

Corollary A.70. *Let $S \subseteq \mathbb{R}^n$ be a bounded rectangle with positive volume. Then R is not a set of measure zero.*

Theorem A.71. *If $A_1, A_2, \cdots$ are sets of measure zero in $\mathbb{R}^n$, then $\bigcup\limits_{k=1}^{\infty} A_k$ has measure zero.*

Proof. Let $\varepsilon > 0$ be given. Since $A_k's$ are sets of measure zero, there exist countable rectangles $\{S_j^{(k)}\}_{j=1}^{\infty}$, such that

$$A_k \subseteq \bigcup_{j=1}^{\infty} S_j^{(k)} \quad \text{and} \quad \sum_{j=1}^{\infty} \nu(S_j^{(k)}) < \frac{\varepsilon}{2^{k+1}} \qquad \forall\, k \in \mathbb{N}\,.$$

Consider the collection consisting of all $S_j^{(k)}$'s. Since there are countable many rectangles in this collection, we can label them as $S_1, S_2, \cdots$, and we have

$$\bigcup_{k=1}^{\infty} A_k \subseteq \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{\infty} S_j^{(k)} = \bigcup_{\ell=1}^{\infty} S_\ell$$

and

$$\sum_{k=1}^{\infty} \nu(S_\ell) = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \nu(S_j^{(k)}) \leqslant \sum_{k=1}^{\infty} \frac{\varepsilon}{2^{k+1}} = \frac{\varepsilon}{2} < \varepsilon\,.$$

Therefore, $\bigcup\limits_{k=1}^{\infty} A_k$ has measure zero. □

Corollary A.72. *The set of rational numbers in $\mathbb{R}$ has measure zero.*

Theorem A.73. *Let $A \subseteq \mathbb{R}^n$ be bounded and $B \subseteq \mathbb{R}^m$ be a set of measure zero. Then $A \times B$ has measure zero in $\mathbb{R}^{n+m}$.*

Proof. Let $\varepsilon > 0$ be given. Since A is bounded, there exist a bounded rectangle R such that $A \subseteq R$. Since B has measure zero, there exist countable rectangles $\{S_k\}_{k=1}^{\infty} \subseteq \mathbb{R}^m$ such that

$$B \subseteq \bigcup_{k=1}^{\infty} S_k \quad \text{and} \quad \sum_{k=1}^{\infty} \nu_m(S_k) < \frac{\varepsilon}{\nu(R)}\,.$$

Then $A \times B \subseteq \bigcup\limits_{k=1}^{\infty} (R \times S_k)$, and

$$\sum_{k=1}^{\infty} \nu_{n+m}(R \times S_k) = \sum_{k=1}^{\infty} \nu_n(R)\nu_m(S_k) = \nu_n(R) \sum_{k=1}^{\infty} \nu_m(S_k)\Big) < \varepsilon\,.$$

Since $R \times S_k$ is a rectangle for all $k \in \mathbb{N}$, we conclude that $A \times B$ has measure zero. □

The Lebesgue Theorem

The Lebesgue theorem states that a function f is Riemann integrable over A if and only if the collection of discontinuities of $\overline{f}^A$, the extension of f defined by (A.12), has measure zero. To prove the theorem, we first give a quantitative measure which measures how discontinuous a discontinuity of a function can be.

Definition A.74. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. For any $x \in \mathbb{R}^n$, the ***oscillation*** of f at x is the quantity

$$\operatorname{osc}(f, x) \equiv \inf_{\delta > 0} \sup_{x_1, x_2 \in D(x,\delta)} \left| f(x_1) - f(x_2) \right|.$$

Let $h(\delta; x)$ denote the quantity of which is taken the infimum; that is,

$$h(\delta; x) \equiv \sup_{x_1, x_2 \in D(x,\delta)} \left| f(x_1) - f(x_2) \right|.$$

Then for fixed $x \in \mathbb{R}^n$, $h(\cdot; x)$ is a decreasing function. Therefore, $\operatorname{osc}(f, x) = \lim\limits_{\delta \to 0} h(\delta; x)$. We note that $h(\delta; x)$ can also be expressed as $\sup\limits_{y \in D(x,\delta)} f(y) - \inf\limits_{y \in D(x,\delta)} f(y)$.

The following lemma provides a way to examine whether a point is a discontinuity of a function or not.

Lemma A.75. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function, and $x_0 \in \mathbb{R}^n$. Then f is continuous at x_0 if and only if* $\operatorname{osc}(f, x_0) = 0$.

Proof. "$\Rightarrow$" Let $\varepsilon > 0$ be given. Since f is continuous at x_0,

$$\exists\, \delta > 0 \ni \left| f(x) - f(x_0) \right| < \frac{\varepsilon}{3} \qquad \text{whenever } x \in D(x_0, \delta).$$

In particular, for any $x_1, x_2 \in D(x_0, \delta)$,

$$\left| f(x_1) - f(x_2) \right| \leqslant \left| f(x_1) - f(x_0) \right| + \left| f(x_0) - f(x_2) \right| < \frac{2\varepsilon}{3};$$

thus $\sup\limits_{x_1, x_2 \in D(x_0,\delta)} \left| f(x_1) - f(x_2) \right| \leqslant \dfrac{2\varepsilon}{3}$ which further suggests that

$$0 \leqslant \operatorname{osc}(f, x_0) \leqslant \frac{2\varepsilon}{3} < \varepsilon.$$

Since ε is given arbitrarily, $\operatorname{osc}(f, x_0) = 0$.

"$\Leftarrow$" Let $\varepsilon > 0$ be given. By the definition of infimum, there exists $\delta > 0$ such that

$$\sup_{x_1,x_2 \in D(x_0,\delta)} \big|f(x_1) - f(x_2)\big| < \varepsilon\,.$$

In particular,

$$\big|f(x) - f(x_0)\big| \leqslant \sup_{x_1,x_2 \in D(x_0,\delta)} \big|f(x_1) - f(x_2)\big| < \varepsilon$$

for all $x \in D(x_0, \delta)$. □

Lemma A.76. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. Then for all $\varepsilon > 0$, the set $D_\varepsilon = \big\{x \in \mathbb{R}^n \,\big|\, \operatorname{osc}(f, x) \geqslant \varepsilon\big\}$ is closed.*

Proof. Suppose that $\{y_k\}_{k=1}^\infty \subseteq D_\varepsilon$ and $y_k \to y$. Then for any $\delta > 0$, there exists $N > 0$ such that $y_k \in D(y, \delta)$ for all $k \geqslant N$. Since $D(y, \delta)$ is open, for each $k \geqslant N$ there exists $\delta_k > 0$ such that $D(y_k, \delta_k) \subseteq D(y, \delta)$; thus we find that

$$\sup_{x_1,x_2 \in D(y_k,\delta_k)} \big|f(x_1) - f(x_2)\big| \leqslant \sup_{x_1,x_2 \in D(y,\delta)} \big|f(x_1) - f(x_2)\big| \qquad \forall\, k \geqslant N\,.$$

The inequality above implies that $\operatorname{osc}(f, y) \geqslant \varepsilon$; thus $y \in D_\varepsilon$ and D_ε is closed. □

Theorem A.77 (Lebesgue)**.** *Let $A \subseteq \mathbb{R}^n$ be bounded, $f : A \to \mathbb{R}$ be a bounded function, and $\bar{f}$ be the extension of f by zero outside A; that is,*

$$\bar{f}(x) = \begin{cases} f(x) & \textit{if } x \in A\,, \\ 0 & \textit{otherwise}\,. \end{cases}$$

Then f is Riemann integrable if and only if the collection of discontinuity of $\bar{f}$ is a set of measure zero.

Proof. Let $D = \big\{x \in \mathbb{R}^n \,\big|\, \operatorname{osc}(\bar{f}, x) > 0\big\}$ and $D_\varepsilon = \big\{x \in \mathbb{R}^n \,\big|\, \operatorname{osc}(\bar{f}, x) \geqslant \varepsilon\big\}$. We remark here that $D = \bigcup\limits_{k=1}^\infty D_{\frac{1}{k}}$.

"$\Rightarrow$" We show that $D_{\frac{1}{k}}$ has measure zero for all $k \in \mathbb{N}$ (if so, then Theorem A.71 implies that D has measure zero).

Let $k \in \mathbb{N}$ be fixed, and $\varepsilon > 0$ be given. By Riemann's condition there exists a partition $\mathcal{P}$ of A such that

$$\sum_{\Delta \in \mathcal{P}} \Big[\sup_{x \in \Delta} \bar{f}(x) - \inf_{x \in \Delta} \bar{f}(x)\Big]\nu(\Delta) < \frac{\varepsilon}{k}\,.$$

Define

$$
\begin{aligned}
D^{(1)}_{\frac{1}{k}} &\equiv \big\{x \in D_{\frac{1}{k}} \,\big|\, x \in \partial\Delta \text{ for some } \Delta \in \mathcal{P}\big\}\,,\\
D^{(2)}_{\frac{1}{k}} &\equiv \big\{x \in D_{\frac{1}{k}} \,\big|\, x \in \operatorname{int}(\Delta) \text{ for some } \Delta \in \mathcal{P}\big\}\,.
\end{aligned}
$$

Then $D_{\frac{1}{k}} = D^{(1)}_{\frac{1}{k}} \cup D^{(2)}_{\frac{1}{k}}$. We note that $D^{(1)}_{\frac{1}{k}}$ has measure zero since it is contained in $\bigcup_{\Delta\in\mathcal{P}} \partial\Delta$ while each $\partial\Delta$ has measure zero. Now we show that $D^{(2)}_{\frac{1}{k}}$ also has measure zero. Let $C = \big\{\Delta \in \mathcal{P} \,\big|\, \operatorname{int}(\Delta) \cap D_{\frac{1}{k}} \neq \varnothing\big\}$. Then $D^{(2)}_{\frac{1}{k}} \subseteq \bigcup\limits_{\Delta\in C} \Delta$. Moreover, we also note that if $\Delta \in C$, $\sup\limits_{x\in\Delta} \bar{f}(x) - \inf\limits_{x\in\Delta} \bar{f}(x) \geqslant \dfrac{1}{k}$. In fact, if $\Delta \in C$, there exists $y \in \operatorname{int}(\Delta) \cap D_{\frac{1}{k}}$; thus choosing $\delta > 0$ such that $D(y,\delta) \subseteq \operatorname{int}(\Delta)$,

$$
\begin{aligned}
\sup_{x\in\Delta} \bar{f}(x) - \inf_{x\in\Delta} \bar{f}(x) &= \sup_{x_1,x_2\in\Delta} \big|\bar{f}(x_1) - \bar{f}(x_2)\big| \geqslant \sup_{x_1,x_2\in D(y,\delta)} \big|\bar{f}(x_1) - \bar{f}(x_2)\big|\\
&\geqslant \inf_{\delta>0} \sup_{x_1,x_2\in D(y,\delta)} \big|\bar{f}(x_1) - \bar{f}(x_2)\big| = \operatorname{osc}(\bar{f}, y) \geqslant \frac{1}{k}\,.
\end{aligned}
$$

As a consequence,

$$
\frac{1}{k} \sum_{\Delta\in C} \nu(\Delta) \leqslant \sum_{\Delta\in\mathcal{P}} \Big[\sup_{x\in\Delta} \bar{f}(x) - \inf_{x\in\Delta} \bar{f}(x)\Big]\nu(\Delta) = U(f,\mathcal{P}) - L(f,\mathcal{P}) < \frac{\varepsilon}{k}
$$

which implies that $\sum\limits_{\Delta\in C} \nu(\Delta) < \varepsilon$. In other words, we establish that $D^{(2)}_{\frac{1}{k}}$ has measure zero. Therefore, $D_{\frac{1}{k}}$ has measure zero for all $k \in \mathbb{N}$; thus D has measure zero.

"$\Leftarrow$" Let R be a closed rectangle with sides parallel to the coordinate axes and $\bar{A} \subseteq \operatorname{int}(R)$, and $\varepsilon > 0$ be given. Define $\varepsilon' = \dfrac{\varepsilon}{2\|f\|_\infty + \nu(R)}$, where $\|f\|_\infty = \sup\limits_{x\in A} |f(x)|$.

1. Since $D_{\varepsilon'}$ is a subset of D, Proposition A.66 implies that $D_{\varepsilon'}$ has measure zero; thus Proposition A.67 provides open rectangles S_1, S_2, $\cdots$ whose sides are parallel to the coordinate axes such that $D_{\varepsilon'} \subseteq \bigcup\limits_{k=1}^{\infty} S_k$, and $\sum\limits_{k=1}^{\infty} \nu(S_k) < \varepsilon'$. In addition, we can assume that $S_k \subseteq R$ for all $k \in \mathbb{N}$ since $D_{\varepsilon'} \subseteq R$.
2. Since $D_{\varepsilon'} \subseteq R$ is bounded, Lemma A.76 suggests that $D_{\varepsilon'}$ is compact; thus $D_{\varepsilon'} \subseteq \bigcup\limits_{k=1}^{N} S_k$ for some $N \in \mathbb{N}$.

Let $\square_k = \overline{S_k}$, and $\mathcal{P}$ be a partition of R satisfying

(a) For each $\Delta \in \mathcal{P}$ with $\Delta \cap D_{\varepsilon'} \neq \varnothing$, $\Delta \subseteq \square_k$ for some $k = 1, \cdots, N$.

(b) For each $k = 1, \cdots, N$, $\square_k$ is the union of rectangles in $\mathcal{P}$.

(c) Some collection of $\Delta \in \mathcal{P}$ forms a partition $\widetilde{\mathcal{P}}$ of A.

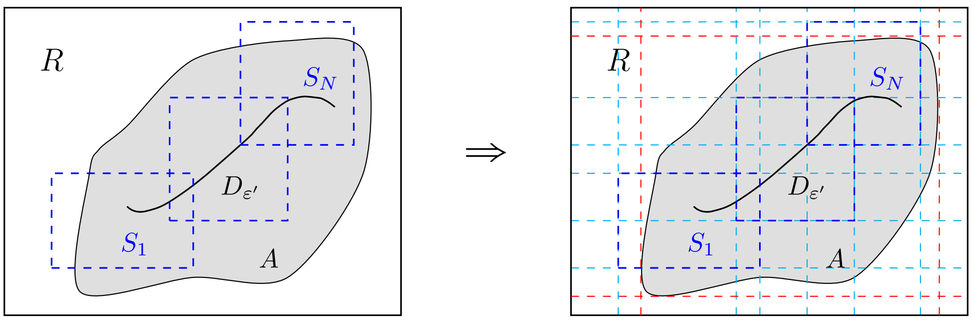


Figure A.2: Constructing partitions $\mathcal{P}$ and $\widetilde{\mathcal{P}}$ from finite rectangles $S_1, \cdots, S_N$

Rectangles in $\mathcal{P}$ fall into two families:

$$C_1 = \left\{\Delta \in \mathcal{P} \,\middle|\, \Delta \subseteq \square_k \text{ for some } k = 1, \cdots, N\right\},$$
$$C_2 = \left\{\Delta \in \mathcal{P} \,\middle|\, \Delta \nsubseteq \square_k \text{ for all } k = 1, \cdots, N\right\}.$$

By the definition of the oscillation function,

$$\forall\, x \notin D_{\varepsilon'}, \exists\, \delta_x > 0 \ni \sup_{x_1, x_2 \in D(x, \delta_x)} \left|\bar{f}(x_1) - \bar{f}(x_2)\right| < \varepsilon'.$$

Since $K = \bigcup_{\Delta \in C_2} \Delta$ is compact, there exists $r > 0$ such that for each $a \in K$, $D(a, r) \subseteq D(y, \delta_y)$ for some $y \in K$. Let $\mathcal{P}'$ be a refinement of $\mathcal{P}$ such that $\|\mathcal{P}'\| < r$. Then if $\Delta' \in \mathcal{P}'$ such that $\Delta' \subseteq \Delta$ for some $\Delta \in C_2$, for some $y \in K$ we have $\Delta' \subseteq D(y, \delta_y)$; thus

$$\sup_{x \in \Delta'} \bar{f}(x) - \inf_{x \in \Delta'} \bar{f}(x) \leqslant \sup_{x \in D(y, \delta_y)} \bar{f}(y) - \inf_{x \in D(y, \delta_y)} \bar{f}(y) = \sup_{x_1, x_2 \in D(y, \delta_y)} \left|\bar{f}(x_1) - \bar{f}(x_2)\right| < \varepsilon'.$$

As a consequence, if $\widetilde{\mathcal{P}}' = \left\{\Delta' \in \mathcal{P}' \,\middle|\, \Delta' \subseteq \Delta \text{ for some } \Delta \in \widetilde{\mathcal{P}}\right\}$, then $\widetilde{\mathcal{P}}'$ is a partition of A and

$$\begin{aligned}
U(f,\widetilde{\mathcal{P}}') - L(f,\widetilde{\mathcal{P}}') &= \Big(\sum_{\substack{\Delta'\in\mathcal{P}'\\ \Delta'\subseteq\Delta\in C_1}} + \sum_{\substack{\Delta'\in\mathcal{P}'\\ \Delta'\subseteq\Delta\in C_2}}\Big)\big(\sup_{x\in\Delta'}\bar{f}(x) - \inf_{x\in\Delta'}\bar{f}(x)\big)\nu(\Delta') \\
&\leqslant 2\|f\|_\infty \sum_{\substack{\Delta'\in\mathcal{P}'\\ \Delta'\subseteq\Delta\in C_1}} \nu(\Delta') + \varepsilon' \sum_{\substack{\Delta'\in\mathcal{P}'\\ \Delta'\subseteq\Delta\in C_2}} \nu(\Delta') \\
&\leqslant 2\|f\|_\infty \sum_{\Delta\in\mathcal{P}\cap C_1} \nu(\Delta) + \varepsilon'\nu(R) \\
&\leqslant 2\|f\|_\infty \sum_{k=1}^{N} \nu(S_k) + +\varepsilon'\nu(R) < \big(2\|f\|_\infty + \nu(R)\big)\varepsilon' = \varepsilon\,;
\end{aligned}$$

thus f is Riemann integrable over A by Riemann's condition. □

EXAMPLE A.78. Let $A = \mathbb{Q}\cap[0,1]$, and $f : A \to \mathbb{R}$ be the constant function $f \equiv 1$. Then

$$\bar{f}(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q}\cap[0,1]\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

The collection of points of discontinuity of $\bar{f}$ is $[0,1]$ which, by Corollary A.70, cannot be a set of measure zero; thus f is not Riemann integrable.

Another way to see that f is not Riemann integrable is $U(f,\mathcal{P}) = 1$ and $L(f,\mathcal{P}) = 0$ for all partitions $\mathcal{P}$ of A.

COROLLARY A.79. *A bounded set $A \subseteq \mathbb{R}^n$ has volume if and only if the boundary of A has measure zero.*

Proof. 1. If $x_0 \notin \partial A$, then there exists $\delta > 0$ such that either $D(x_0,\delta) \subseteq A$ or $D(x_0,\delta) \subseteq A^\complement$; thus $\overline{1_A}$ is continuous at $x_0 \notin \partial A$ since $\overline{1_A}(x)$ is constant for all $x \in D(x_0,\delta)$.

2. On the other hand, if $x_0 \in \partial A$, then there exists $x_k \in A$, $y_k \in A^\complement$ such that $x_k \to x_0$ and $y_k \to x_0$ as $k \to \infty$. This implies that $\overline{1_A}$ cannot be continuous at x_0 since $\overline{1_A}(x_k) = 1$ while $\overline{1_A}(y_k) = 0$ for all $k \in \mathbb{N}$.

As a consequence, the collection of discontinuity of $\overline{1_A}$ is exactly ∂A, and the corollary follows from Lebesgue's theorem. □

COROLLARY A.80. *Let $A \subseteq \mathbb{R}^n$ be bounded and have volume. A bounded function $f : A \to \mathbb{R}$ with a finite or countable number of points of discontinuity is Riemann integrable.*

Proof. We note that $\{x \in \mathbb{R}^n \,|\, \operatorname{osc}(\bar{f}, x) > 0\} \subseteq \partial A \cup \{x \in A \,|\, f \text{ is discontinuous at } x\}$.
□

Remark A.81. In addition to the set inclusion listed in the proof of Corollary A.80, we also have

$$\{x \in A \,|\, f \text{ is discontinuous at } x\} \subseteq \{x \in \mathbb{R}^n \,|\, \operatorname{osc}(\bar{f}, x) > 0\}.$$

Therefore, if $A \subseteq \mathbb{R}^n$ is bounded and has volume, then a bounded function $f : A \to \mathbb{R}$ is Riemann integrable if and only if the collection of points of discontinuity of f has measure zero.

Corollary A.82. *A bounded function is integrable over a compact set of measure zero.*

Proof. If $f : K \to \mathbb{R}$ is bounded, and K is a compact set of measure zero, then the collection of discontinuities of $\bar{f}$ is a subset of K. □

Corollary A.83. *Suppose that $A, B \subseteq \mathbb{R}^n$ are bounded sets with volume, and $f : A \to \mathbb{R}$ is Riemann integrable over A. Then f is Riemann integrable over $A \cap B$.*

Proof. By the inclusion

$$\{x \in \operatorname{int}(A \cap B) \,|\, \operatorname{osc}(\overline{f}^{A\cap B}, x) > 0\} \subseteq \{x \in \mathbb{R}^n \,|\, \operatorname{osc}(\overline{f}^{A}, x) > 0\},$$

we find that

$$\begin{aligned}\{x \in \mathbb{R}^n \,|\, \operatorname{osc}(\overline{f}^{A\cap B}, x) > 0\} &\subseteq \partial(A \cap B) \cup \{x \in \operatorname{int}(A \cap B) \,|\, \operatorname{osc}(\overline{f}^{A\cap B}, x) > 0\} \\ &\subseteq \partial A \cup \partial B \cup \{x \in \mathbb{R}^n \,|\, \operatorname{osc}(\overline{f}^{A}, x) > 0\}.\end{aligned}$$

Since ∂A and ∂B both have measure zero, the integrability of f over $A \cap B$ then follows from the integrability of f over A and the Lebesgue Theorem. □

Remark A.84. Suppose that $A \subseteq \mathbb{R}^n$ is a bounded set of measure zero. Even if $f : A \to \mathbb{R}$ is continuous, f may not be Riemann integrable. For example, the function f given in Example A.78 is not Riemann integrable even though f is continuous on A.

Remark A.85. When $f : A \to \mathbb{R}$ is Riemann integrable over A, it is not necessary that A has volume. For example, the zero function is Riemann integrable over $A = \mathbb{Q} \cap [0, 1]$ even though A does not has volume.

COROLLARY A.86 (Lebesgue's Differentiation Theorem (for Riemann integrable function))**.** *Let $A \subseteq \mathbb{R}^{\mathrm{n}}$ be a bounded, open set such that ∂A has measure zero. Suppose that $f : A \to \mathbb{R}$ is a bounded and Riemann integrable function. Then*

$$\lim_{r\to 0} \frac{1}{\nu(B(x_0,r)\cap A)} \int_{B(x_0,r)\cap A} f(x)\,dx = f(x_0) \text{ for almost every } x \in A\,. \tag{A.15}$$

Proof. Let $\varepsilon > 0$ be given, and suppose that f is continuous at x_0. Then there exists $\delta > 0$ such that

$$\big|f(x) - f(x_0)\big| < \frac{\varepsilon}{2} \qquad \forall\, x \in B(x_0,\delta)\cap A\,.$$

Since ∂A has measure zero, by the fact that $\partial(B(x_0,r)\cap A \subseteq \partial B(x_0,r) \cup \partial A$ we find that $\partial(B(x_0,r)\cap A)$ also has measure zero for all $r > 0$. In other words, $B(x_0,r)\cap A$ has volume. Then if $0 < r < \delta$,

$$\begin{aligned}
&\Big|\frac{1}{\nu(B(x_0,r)\cap A)} \int_{B(x_0,r)\cap A} f(x)\,dx - f(x_0)\Big| \\
&\quad= \Big|\frac{1}{\nu(B(x_0,r)\cap A)} \int_{B(x_0,r)\cap A} \big(f(x) - f(x_0)\big)\,dx\Big| \\
&\quad\leqslant \frac{1}{\nu(B(x_0,r)\cap A)} \int_{B(x_0,r)\cap A} \big|f(x) - f(x_0)\big|\,dx \\
&\quad\leqslant \frac{\varepsilon}{2}\frac{1}{\nu(B(x_0,r)\cap A)} \int_{B(x_0,r)\cap A} 1\,dx = \frac{\varepsilon}{2} < \varepsilon\,.
\end{aligned}$$

This implies that (A.15) holds for all x_0 at which f is continuous. The theorem then follows from the Lebesgue theorem. □

A.2.3 Properties of the Integrals

PROPOSITION A.87. *Let $A \subseteq \mathbb{R}^n$ be bounded, and $f, g : A \to \mathbb{R}$ be bounded. Then*

(a) *If $B \subseteq A$, then $\underline{\int_A} (f1_B)(x)\,dx = \underline{\int_B} f(x)\,dx$ and $\overline{\int_A} (f1_B)(x)\,dx = \overline{\int_B} f(x)\,dx$.*

(b) $\underline{\int_A} f(x)\,dx + \underline{\int_A} g(x)\,dx \leqslant \underline{\int_A} (f+g)(x)\,dx \leqslant \overline{\int_A} (f+g)(x)\,dx \leqslant \overline{\int_A} f(x)\,dx + \overline{\int_A} g(x)\,dx$.

(c) *If $c \geqslant 0$, then $\underline{\int_A} (cf)(x)\,dx = c\underline{\int_A} f(x)\,dx$ and $\overline{\int_A} (cf)(x)\,dx = c\overline{\int_A} f(x)\,dx$. If $c < 0$, then $\underline{\int_A} (cf)(x)\,dx = c\overline{\int_A} f(x)\,dx$ and $\overline{\int_A} (cf)(x)\,dx = c\underline{\int_A} f(x)\,dx$.*

(d) *If* $f \leqslant g$ *on* A*, then* $\underline{\int_A} f(x)\,dx \leqslant \underline{\int_A} g(x)\,dx$ *and* $\overline{\int_A} f(x)\,dx \leqslant \overline{\int_A} g(x)\,dx$.

(e) *If* A *has volume zero, then* f *is Riemann integrable over* A*, and* $\int_A f(x)\,dx = 0$.

Remark A.88. Let $A \subseteq \mathbb{R}^n$ be bounded and $f, g : A \to \mathbb{R}$ be bounded. Then (b) of Proposition A.87 also implies that

$$\underline{\int_A} (f-g)(x)\,dx \leqslant \underline{\int_A} f(x)\,dx - \underline{\int_A} g(x)\,dx$$

and

$$\overline{\int_A} f(x)\,dx - \overline{\int_A} g(x)\,dx \leqslant \overline{\int_A} (f-g)(x)\,dx\,.$$

Corollary A.89. *Let* $A, B \subseteq \mathbb{R}^n$ *be bounded such that* $A \cap B$ *has volume zero, and* $f : A \cup B \to \mathbb{R}$ *be bounded. Then*

$$\underline{\int_A} f(x)\,dx + \underline{\int_B} f(x)\,dx \leqslant \underline{\int_{A\cup B}} f(x)\,dx \leqslant \overline{\int_{A\cup B}} f(x)\,dx \leqslant \overline{\int_A} f(x)\,dx + \overline{\int_B} f(x)\,dx\,.$$

The following theorem is a direct consequence of Proposition A.87.

Theorem A.90. *Let* $A \subseteq \mathbb{R}^n$ *be bounded,* $c \in \mathbb{R}$*, and* $f, g : A \to \mathbb{R}$ *be Riemann integrable. Then*

1. $f \pm g$ *is Riemann integrable, and* $\int_A (f \pm g)(x)\,dx = \int_A f(x)\,dx \pm \int_A g(x)\,dx$.
2. cf *is Riemann integrable, and* $\int_A (cf)(x)\,dx = c\int_A f(x)\,dx$.
3. $|f|$ *is Riemann integrable, and* $\left|\int_A f(x)\,dx\right| \leqslant \int_A |f(x)|dx$.
4. *If* $f \leqslant g$*, then* $\int_A f(x)\,dx \leqslant \int_A g(x)\,dx$.
5. *If* A *has volume and* $|f| \leqslant M$*, then* $\left|\int_A f(x)\,dx\right| \leqslant M\nu(A)$.

Theorem A.91. *Let* $A \subseteq \mathbb{R}^n$ *be bounded, and* $f : A \to \mathbb{R}$ *be a bounded integrable function.*

1. *If* A *has measure zero, then* $\int_A f(x)\,dx = 0$.

2. *If $f(x) \geqslant 0$ for all $x \in A$, and $\displaystyle\int_A f(x)\,dx = 0$, then the set $\{x \in A \,|\, f(x) \neq 0\}$ has measure zero.*

Remark A.92. Combining Corollary A.82 and Theorem A.91, we conclude that the integral of a bounded function over a compact set of measure zero is zero.

Remark A.93. Let $A = \mathbb{Q} \cap [0,1]$ and $f : A \to \mathbb{R}$ be the constant function $f \equiv 1$. We have shown in Example A.78 that f is not Riemann integrable. We note that A has no volume since $\partial A = [0,1]$ which is not a set of measure zero. However, A has measure zero since it consists of countable number of points.

1. Since f is continuous on A, the condition that A has volume in Corollary A.80 cannot be removed.

2. Since A has measure zero, the condition that f is Riemann integrable in Theorem A.91 cannot be removed.

Theorem A.94 (Mean Value Theorem for Integrals)**.** *Let A be a subset of $\mathbb{R}^n$ such that A has volume and is compact and connected. Suppose that $f : A \to \mathbb{R}$ is continuous, then there exists $x_0 \in A$ such that*

$$\int_A f(x)\,dx = f(x_0)\nu(A)\,.$$

*The quantity $\displaystyle\frac{1}{\nu(A)}\int_A f(x)\,dx$ is called the **average** of f over A.*

Definition A.95. Let $A \subseteq \mathbb{R}^n$ be a set and $f : A \to \mathbb{R}$ be a function. For $B \subseteq A$, the ***restriction of** f **to** B* is the function $f\big|_B : A \to \mathbb{R}$ given by $f|_B = f 1_B$. In other words,

$$f\big|_B(x) = \begin{cases} f(x) & \text{if } x \in B\,, \\ 0 & \text{if } x \in A\backslash B\,. \end{cases}$$

The following lemma is a direct consequence of Proposition A.87 (a).

Lemma A.96. *Let $A \subseteq \mathbb{R}^n$ be bounded, and $f : A \to \mathbb{R}$ be a bounded function. Suppose that $B \subseteq A$, and $f\big|_B$ is Riemann integrable over A. Then f is Riemann integrable over B, and*

$$\int_A f\big|_B(x)\,dx = \int_B f(x)\,dx\,.$$

Theorem A.97. *Let A, B be bounded subsets of $\mathbb{R}^n$ be such that $A \cap B$ has measure zero, and $f : A \cup B \to \mathbb{R}$ be such that $f\big|_{A\cap B}$, $f\big|_A$ and $f\big|_B$ are all Riemann integrable over $A \cup B$. Then f is integrable over $A \cup B$, and*

$$\int_{A\cup B} f(x)\,dx = \int_A f(x)\,dx + \int_B f(x)\,dx\,.$$

A.2.4 The Fubini Theorem

If $f : [a, b] \to \mathbb{R}$ is continuous, the fundamental theorem of Calculus can be applied to computed the integral of f over $[a, b]$. In the following two sections, we focus on how the integral of f over $A \subseteq \mathbb{R}^n$, where $n \geqslant 2$, can be computed if the integral exists.

Definition A.98. Let $S = A \times B$ be the product of two bounded sets $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$, and $f : S \to \mathbb{R}$ be bounded. For each fixed $x \in A$, the lower integral of the function $f(x, \cdot) : B \to \mathbb{R}$ is denoted by $\underline{\int}_B f(x, y)\,dy$, and the upper integral of $f(x, \cdot) : B \to \mathbb{R}$ is denoted by $\overline{\int}_B f(x, y)\,dy$. If for each $x \in A$ the upper integral and the lower integral of $f(x, \cdot) : B \to \mathbb{R}$ are the same, we simply write $\int_B f(x, y)\,dy$ for the integrals of $f(x, \cdot)$ over $[c, d]$. The integrals $\overline{\int}_A f(x, y)\,dx$, $\underline{\int}_A f(x, y)\,dx$ and $\int_A f(x, y)\,dx$ are defined in a similar way.

Now we state and prove the general Fubini Theorem.

Theorem A.99 (Fubini's Theorem)**.** *Let $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$ be rectangles, and $f : A \times B \to \mathbb{R}$ be bounded. For $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, write $z = (x, y)$. Then*

$$\underline{\int}_{A\times B} f(z)\,dz \leqslant \underline{\int}_A \Big(\underline{\int}_B f(x,y)dy\Big)dx \leqslant \overline{\int}_A \Big(\overline{\int}_B f(x,y)dy\Big)dx \leqslant \overline{\int}_{A\times B} f(z)\,dz \tag{A.16}$$

and

$$\underline{\int}_{A\times B} f(z)\,dz \leqslant \underline{\int}_B \Big(\underline{\int}_A f(x,y)dx\Big)dy \leqslant \overline{\int}_B \Big(\overline{\int}_A f(x,y)dx\Big)dy \leqslant \overline{\int}_{A\times B} f(z)\,dz\,. \tag{A.17}$$

In particular, if $f: A \times B \to \mathbb{R}$ is Riemann integrable, then

$$\begin{aligned}\int_{A\times B} f(z)\,dz &= \int_A \Big(\underline{\int_B} f(x,y)dy\Big)dx = \int_A \Big(\overline{\int_B} f(x,y)dy\Big)dx \\ &= \int_B \Big(\underline{\int_A} f(x,y)dx\Big)dy = \int_B \Big(\overline{\int_A} f(x,y)dx\Big)dy\,.\end{aligned}$$

Proof. It suffices to prove (A.16). Let $\varepsilon > 0$ be given. Choose a partition $\mathcal{P}$ of $A \times B$ such that $L(f,\mathcal{P}) > \underline{\int_{A\times B}} f(z)\,dz - \varepsilon$. Since $\mathcal{P}$ is a partition of $A \times B$, there exist partition $\mathcal{P}_x$ of A and partition $\mathcal{P}_y$ of B such that $\mathcal{P} = \{\Delta = R \times S \,|\, R \in \mathcal{P}_x, S \in \mathcal{P}_y\}$. By Proposition A.87 and Corollary A.89, we find that

$$\begin{aligned}\underline{\int_A} \Big(\underline{\int_B} f(x,y)\,dy\Big)dx &\geqslant \sum_{R\in\mathcal{P}_x} \underline{\int_R} \Big(\sum_{S\in\mathcal{P}_y} \underline{\int_S} f(x,y)\,dy\Big)dx \\ &\geqslant \sum_{R\in\mathcal{P}_x}\sum_{S\in\mathcal{P}_y} \underline{\int_R}\Big(\underline{\int_S} f(x,y)\,dy\Big)dx \\ &\geqslant \sum_{R\in\mathcal{P}_x, S\in\mathcal{P}_y} \inf_{(x,y)\in R\times S} f(x,y)\nu_m(S)\nu_n(R) \\ &= \sum_{\Delta\in\mathcal{P}} \inf_{(x,y)\in\Delta} f(x,y)\nu_{n+m}(\Delta) = L(f,\mathcal{P}) > \underline{\int_{A\times B}} f(z)dz - \varepsilon\,.\end{aligned}$$

Since $\varepsilon > 0$ is given arbitrarily, we conclude that

$$\underline{\int_{A\times B}} f(z)\,dz \leqslant \underline{\int_B}\Big(\underline{\int_A} f(x,y)dx\Big)dy\,.$$

Similarly, $\overline{\int_A}\Big(\overline{\int_B} f(x,y)dy\Big)dx \leqslant \overline{\int_{A\times B}} f(z)\,dz$; thus (A.16) is concluded. □

Corollary A.100. *Let $S \subseteq \mathbb{R}^n$ be a bounded set with volume, $\varphi_1, \varphi_2 : S \to \mathbb{R}$ be continuous maps such that $\varphi_1(x) \leqslant \varphi_2(x)$ for all $x \in S$, $A = \{(x,y) \in \mathbb{R}^n \times \mathbb{R} \,|\, x \in S, \varphi_1(x) \leqslant y \leqslant \varphi_2(x)\}$, and $f: A \to \mathbb{R}$ be continuous. Then f is Riemann integrable over A, and*

$$\int_A f(x,y)\,d(x,y) = \int_S \Big(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,dy\Big)dx\,. \tag{A.18}$$

Proof. Since ∂A has measure zero, and f is continuous on A, Corollary A.80 implies that f is Riemann integrable over A. Let R be the smallest closed rectangle with sides parallel to the coordinate axes and $S \subseteq R$, and $m = \min\limits_{x\in S} \varphi_1(x)$ and $M = \max\limits_{x\in S} \varphi_2(x)$. Then $A \subseteq R \times [m, M]$; thus Theorem A.97 and the Fubini Theorem imply that

$$\int_A f(x,y)\, d(x,y) = \int_{R\times[m,M]} \overline{f}^A(x,y)\, d(x,y) = \int_R \Big(\underline{\int_m^M} \overline{f}^A(x,y)\, dy \Big) dx\,.$$

Let $g(x) = \underline{\int_m^M} \overline{f}^A(x,y)\, dy$. Then $g(x) = 0$ if $x \notin S$; thus with the help of Lemma A.96 the identity above further implies that

$$\int_A f(x,y)\, d(x,y) = \int_R g(x) 1_S(x)\, dx = \int_S g(x)\, dx = \int_S \Big(\underline{\int_m^M} \overline{f}^A(x,y)\, dy \Big) dx\,. \tag{A.19}$$

On the other hand, for each fixed $x \in S$, let $A_x = \big\{ y \in \mathbb{R} \,\big|\, \varphi_1(x) \leqslant y \leqslant \varphi_2(x) \big\}$. Then $\overline{f}^A(x,y) = f(x,y) 1_{A_x}(y)$ for all $(x,y) \in R \times [m, M]$ or equivalently, $\overline{f}^A(x,\cdot) = f(x,\cdot)|_{A_x}$ for all $x \in S$; thus Proposition A.87 (a) implies that

$$\underline{\int_m^M} \overline{f}^A(x,y)\, dy = \int_{A_x} f(x,y)\, dy = \int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\, dy \qquad \forall\, x \in S\,. \tag{A.20}$$

Combining (A.19) and (A.20), we conclude (A.18). $\square$

Example A.101. In this example we compute the volume of the n-dimensional unit ball ω_n. By the Fubini theorem,

$$\omega_n = \int_{-1}^{1} \int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} \cdots \int_{-\sqrt{1-x_1^2-\cdots-x_{n-1}^2}}^{\sqrt{1-x_1^2-\cdots-x_{n-1}^2}} dx_n \cdots dx_1\,.$$

Note that the integral $\int_{-\sqrt{1-x_1^2}}^{\sqrt{1-x_1^2}} \cdots \int_{-\sqrt{1-x_1^2-\cdots-x_{n-1}^2}}^{\sqrt{1-x_1^2-\cdots-x_{n-1}^2}} dx_n \cdots dx_2$ is in fact $\omega_{n-1}(1-x_1^2)^{\frac{n-1}{2}}$; thus

$$\omega_n = \int_{-1}^{1} \omega_{n-1}(1-x^2)^{\frac{n-1}{2}}\, dx = 2\,\omega_{n-1} \int_0^{\frac{\pi}{2}} \cos^n\theta d\theta\,. \tag{A.21}$$

Integrating by parts,

$$\begin{aligned}\int_0^{\frac{\pi}{2}} \cos^n\theta\, d\theta &= \int_0^{\frac{\pi}{2}} \cos^{n-1}\theta\, d(\sin\theta) = \cos^{n-1}\theta \sin\theta \Big|_{\theta=0}^{\theta=\frac{\pi}{2}} + (n-1)\int_0^{\frac{\pi}{2}} \cos^{n-2}\theta \sin^2\theta\, d\theta \\ &= (n-1)\int_0^{\frac{\pi}{2}} \cos^{n-2}\theta (1-\cos^2\theta)\, d\theta\end{aligned}$$

which implies that

$$\int_0^{\frac{\pi}{2}} \cos^n\theta\,d\theta = \frac{n-1}{n}\int_0^{\frac{\pi}{2}} \cos^{n-2}\theta\,d\theta\,.$$

As a consequence,

$$\int_0^{\frac{\pi}{2}} \cos^n\theta\,d\theta = \begin{cases} \dfrac{(n-1)(n-3)\cdots 2}{n(n-2)\cdots 3}\displaystyle\int_0^{\frac{\pi}{2}} \cos\theta\,d\theta & \text{if } n \text{ is odd}\,, \\ \dfrac{(n-1)(n-3)\cdots 1}{n(n-2)\cdots 2}\displaystyle\int_0^{\frac{\pi}{2}} d\theta & \text{if } n \text{ is even}\,; \end{cases}$$

thus the recursive formula (A.21) implies that $\omega_n = \dfrac{2\,\omega_{n-2}}{n}\pi$. Further computations shows that

$$\omega_n = \begin{cases} \dfrac{(2\pi)^{\frac{n-1}{2}}}{n(n-2)\cdots 3}\omega_1 & \text{if } n \text{ is odd}\,, \\ \dfrac{(2\pi)^{\frac{n-2}{2}}}{n(n-2)\cdots 4}\omega_2 & \text{if } n \text{ is even}\,. \end{cases}$$

Let Γ be the Gamma function defined by $\Gamma(t) = \displaystyle\int_0^\infty x^{t-1}e^{-x}\,dx$ for $t > 0$. Then $\Gamma(x+1) = x\Gamma(x)$ for all $x > 0$, $\Gamma(1) = 1$ and $\Gamma\big(\frac{1}{2}\big) = \sqrt{\pi}$. By the fact that $\omega_1 = 2$ and $\omega_2 = \pi$, we can express ω_n as

$$\omega_n = \frac{\pi^{\frac{n}{2}}}{\Gamma\big(\frac{n+2}{2}\big)}\,.$$

A.2.5 Change of Variables Formula

The Fubini theorem can be used to find the integral of a (Riemann integrable) function over a rectangular domain if the iterated integrals can be evaluated; however, like the integral of a function of one variable, in many cases we need to make use of several changes of variables in order to transform the integral to another integral that can be easily evaluated. In this section, we state the change of variables formula for the integral of functions of several variables.

THEOREM A.102 (Change of Variables Formula)**.** *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be an open bounded set, and $g : \mathcal{U} \to \mathbb{R}^n$ be an one-to-one $\mathscr{C}^1$ mapping with $\mathscr{C}^1$ inverse; that is, $g^{-1} : g(\mathcal{U}) \to \mathcal{U}$ is also continuously differentiable. Assume that the Jacobian of g, $\mathbf{J}_g = \det([Dg])$, does not vanish in $\mathcal{U}$, and $E \subset\subset \mathcal{U}$ has volume. Then $g(E)$ has volume. Moreover, if*

$f : g(E) \to \mathbb{R}$ is bounded and integrable, then $(f \circ g)\mathbf{J}_g$ is integrable over E, and

$$\int_{g(E)} f(y)\,dy = \int_E (f \circ g)(x)\big|\mathbf{J}_g(x)\big|\,dx = \int_E (f \circ g)(x)\Big|\frac{\partial(g_1, \cdots, g_n)}{\partial(x_1, \cdots, x_n)}\Big|\,dx\,.$$

Remark A.103. The condition that g has to be defined on a larger open set $\mathcal{U}$ can be relaxed using the Monotone Convergence Theorem. See Theorem A.176 (another change of variables formula for more general situations) for the precise statement.

A.3 Uniform Convergence and the Space of Continuous Functions

A.3.1 Pointwise and Uniform Convergence

Definition A.104. Let (M, d) and (N, ρ) be two metric spaces, $A \subseteq M$ be a set, and $f_k, f : A \to N$ be functions for $k = 1, 2, \cdots$. The sequence of function $\{f_k\}_{k=1}^{\infty}$ is said to ***converge pointwise*** to f if

$$\lim_{k\to\infty} \rho\big(f_k(a), f(a)\big) = 0 \qquad \forall\, a \in A.$$

We often write $f_k \to f$ p.w. if f_k converges pointwise to f.

Let $B \subseteq A$ be a subset. The sequence of functions $\{f_k\}_{k=1}^{\infty}$ is said to ***converge uniformly*** to f on B (or $\{f_k\}_{k=1}^{\infty}$ converges to f uniformly on B) if

$$\lim_{k\to\infty} \sup_{x\in B} \rho\big(f_k(x), f(x)\big) = 0\,.$$

In other words, $\{f_k\}_{k=1}^{\infty}$ converges uniformly to f on B if for every $\varepsilon > 0$, $\exists\, N > 0$ such that

$$\rho\big(f_k(x), f(x)\big) < \varepsilon \qquad \forall\, k \geqslant N \text{ and } x \in B\,.$$

Proposition A.105. *Let (M, d) and (N, ρ) be two metric spaces, $A \subseteq M$ be a set, and $f_k, f : A \to N$ be functions for $k = 1, 2, \cdots$. If $\{f_k\}_{k=1}^{\infty}$ converges uniformly to f on A, then $\{f_k\}_{k=1}^{\infty}$ converges pointwise to f.*

Proposition A.106 (Cauchy criterion for uniform convergence)**.** *Let (M, d) and (N, ρ) be two metric spaces, $A \subseteq M$ be a set, and $f_k : A \to N$ be a sequence of functions. Suppose that (N, ρ) is complete. Then $\{f_k\}_{k=1}^{\infty}$ converges uniformly on $B \subseteq A$ if and only if for every $\varepsilon > 0$, $\exists\, N > 0$ such that*

$$\rho\big(f_k(x), f_\ell(x)\big) < \varepsilon \qquad \forall\, k, \ell \geqslant N \text{ and } x \in B\,.$$

Theorem A.107. *Let (M,d) and (N,ρ) be two metric spaces, $A \subseteq M$ be a set, and $f_k : A \to N$ be a sequence of continuous functions converging to $f : A \to N$ uniformly on A. Then f is continuous on A; that is,*

$$\lim_{x\to a} f(x) = \lim_{x\to a}\lim_{k\to\infty} f_k(x) = \lim_{k\to\infty}\lim_{x\to a} f_k(x) = f(a)\,.$$

Remark A.108. The uniform limit of sequence of continuous function might not be uniformly continuous. For example, let $A = (0,1)$ and $f_k(x) = \dfrac{1}{x}$ for all $k \in \mathbb{N}$. Then $\{f_k\}_{k=1}^{\infty}$ converges uniformly to $f(x) = \dfrac{1}{x}$, but the limit function is not uniformly continuous on A.

Theorem A.109. *Let $I \subseteq \mathbb{R}$ be a finite interval, $f_k : I \to \mathbb{R}$ be a sequence of differentiable functions, and $g : I \to \mathbb{R}$ be a function. Suppose that $\big\{f_k(a)\big\}_{k=1}^{\infty}$ converges for some $a \in I$, and $\{f_k'\}_{k=1}^{\infty}$ converges uniformly to g on I. Then*

1. *$\{f_k\}_{k=1}^{\infty}$ converges uniformly to some function f on I.*

2. *The limit function f is differentiable on I, and $f'(x) = g(x)$ for all $x \in I$; that is,*
$$\lim_{k\to\infty} f_k'(x) = \lim_{k\to\infty}\frac{d}{dx}f_k(x) = \frac{d}{dx}\lim_{k\to\infty} f_k(x) = f'(x)\,.$$

Theorem A.110. *Let $f_k : [a,b] \to \mathbb{R}$ be a sequence of Riemann integrable functions which converges uniformly to f on $[a,b]$. Then f is Riemann integrable, and*

$$\lim_{k\to\infty}\int_a^b f_k(x)dx = \int_a^b \lim_{k\to\infty} f_k(x)dx = \int_a^b f(x)dx\,. \tag{A.22}$$

Theorem A.111 (Dini's Theorem)**.** *Let K be a compact set, and $f_k : K \to \mathbb{R}$ be continuous for all $k \in \mathbb{N}$ such that $\{f_k\}_{k=1}^{\infty}$ converges pointwise to a continuous function $f : K \to \mathbb{R}$. Suppose that $f_k \leqslant f_{k+1}$ for all $k \in \mathbb{N}$. Then $\{f_k\}_{k=1}^{\infty}$ converges uniformly to f on K.*

A.3.2 The Space of Continuous Functions

Definition A.112. Let (M,d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a normed vector space, and $A \subseteq M$ be a subset. We define $\mathscr{C}(A;\mathcal{V})$ as the collection of all continuous functions on A with value in $\mathcal{V}$; that is,

$$\mathscr{C}(A;\mathcal{V}) = \big\{f : A \to \mathcal{V}\,\big|\, f \text{ is continuous on } A\big\}\,.$$

Let $\mathscr{C}_b(A;\mathcal{V})$ be the subspace of $\mathscr{C}(A;\mathcal{V})$ which consists of all bounded continuous functions on A; that is,

$$\mathscr{C}_b(A;\mathcal{V}) = \left\{ f \in \mathscr{C}(A;\mathcal{V}) \,\middle|\, f \text{ is bounded} \right\}.$$

Every $f \in \mathscr{C}_b(A;\mathcal{V})$ is associated with a non-negative real number $\|f\|_\infty$ given by

$$\|f\|_\infty = \sup\left\{ \|f(x)\| \,\middle|\, x \in A \right\} = \sup_{x\in A} \|f(x)\|$$

The number $\|f\|_\infty$ is called the ***sup-norm*** of f.

Proposition A.113. *Let (M,d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a normed vector space, $A \subseteq M$ be a subset.*

1. *$\mathscr{C}(A;\mathcal{V})$ and $\mathscr{C}_b(A;\mathcal{V})$ are vector spaces.*
2. *$\big(\mathscr{C}_b(A;\mathcal{V}), \|\cdot\|_\infty\big)$ is a normed vector space.*
3. *If $K \subseteq M$ is compact, then $\mathscr{C}(K;\mathcal{V}) = \mathscr{C}_b(K;\mathcal{V})$.*

Remark A.114. In general $\|\cdot\|_\infty$ is not a "norm" on $\mathscr{C}(A;\mathcal{V})$. For example, the function $f(x) = \dfrac{1}{x}$ belongs to $\mathscr{C}((0,1);\mathbb{R})$ and $\|f\|_\infty = \infty$. Note that to be a norm $\|f\|_\infty$ has to take values in $\mathbb{R}$, and $\infty \notin \mathbb{R}$.

Proposition A.115. *Let (M,d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a normed vector space, $A \subseteq M$ be a subset, and $f_k, f \in \mathscr{C}_b(A;\mathcal{V})$ for all $k \in \mathbb{N}$. Then $\{f_k\}_{k=1}^\infty$ converges uniformly to f on A if and only if $\{f_k\}_{k=1}^\infty$ converges to f in $\big(\mathscr{C}_b(A;\mathcal{V}), \|\cdot\|_\infty\big)$.*

Theorem A.116. *Let (M,d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a normed vector space, and $A \subseteq M$ be a subset. If $(\mathcal{V}, \|\cdot\|)$ is complete, so is $\big(\mathscr{C}_b(A;\mathcal{V}), \|\cdot\|_\infty\big)$.*

Definition A.117. A ***Banach space*** is a complete normed vector space.

A.3.3 The Arzelà-Ascoli Theorem

In Dini's theorem, we prove that in the space of continuous functions, under the additional assumption that the pointwise convergence is monotone, pointwise convergence is uniform convergence. On the other hand, in general pointwise convergent sequences of functions do not converge monotonically so monotone convergence is not a good criteria for uniform convergence. In this section, we investigate the difference between pointwise convergence and uniform convergence. To be more precise, we are looking for a condition such that

> A pointwise convergence sequence of continuous functions satisfies ***this condition*** if and only if this sequence converges uniformly.

This condition differentiates the pointwise convergence and the uniform convergence of sequences of continuous functions, and will play an important role for judging whether a subset of the space of continuous functions are compact or not.

Equi-continuous family of functions

The first part of this section is devoted to the investigation of the difference between the pointwise convergence and the uniform convergence of sequence of continuous functions.

Definition A.118. Let (M,d) be a metric space, $(\mathcal{V},\|\cdot\|)$ be a normed vector space, and $A\subseteq M$ be a subset. A subset $B\subseteq \mathscr{C}_b(A;\mathcal{V})$ is said to be ***equi-continuous*** if

$$\forall\,\varepsilon>0, \exists\,\delta>0 \ni \|f(x_1)-f(x_2)\|<\varepsilon \text{ whenever } d(x_1,x_2)<\delta,\ x_1,x_2\in A, \text{ and } f\in B\,.$$

Remark A.119. 1. If $B\subseteq \mathscr{C}_b(A;\mathcal{V})$ is equi-continuous, and C is a subset of B, then C is also equi-continuous.

2. In an equi-continuous set of functions B, every $f\in B$ is uniformly continuous.

Lemma A.120. *Let (M,d) be a metric space, $(\mathcal{V},\|\cdot\|)$ be a normed vector space, and $K\subseteq M$ be a compact subset. If $B\subseteq \mathscr{C}(K;\mathcal{V})$ is pre-compact, then B is equi-continuous.*

Proof. Suppose the contrary that B is not equi-continuous. Then $\exists\,\varepsilon>0$ such that

$$\forall\,k\in\mathbb{N}, \exists\,x_k,y_k\in K \text{ and } f_k\in B \ni d(x_k,y_k)<\frac{1}{k} \text{ but } \|f_k(x_k)-f_k(y_k)\|\geqslant\varepsilon\,.$$

Since B is pre-compact in $\big(\mathscr{C}(K;\mathcal{V}),\|\cdot\|_\infty\big)$ and K is compact in (M,d), there exists a subsequence $\big\{f_{k_j}\big\}_{j=1}^\infty$ and $\{x_{k_j}\}_{j=1}^\infty$ such that $\big\{f_{k_j}\big\}_{j=1}^\infty$ converges uniformly to some function $f \in \big(\mathscr{C}(K;\mathcal{V}), \|\cdot\|_\infty\big)$ and $\{x_{k_j}\}_{j=1}^\infty$ converges to some $a \in K$. We must also have $\{y_{k_j}\}_{j=1}^\infty$ converges to a since $d(x_{k_j}, y_{k_j}) < \dfrac{1}{k_j}$.

Since f is continuous at a,

$$\exists\, \delta > 0 \ni \|f(x) - f(a)\| < \frac{\varepsilon}{5} \qquad \text{if } x \in D(a,\delta) \cap K.$$

Moreover, since $\{f_{k_j}\}_{j=1}^\infty$ converges to f uniformly on K and $x_{k_j}, y_{k_j} \to a$ as $j \to \infty$, $\exists\, N > 0$ such that

$$\|f_{k_j}(x) - f(x)\| < \frac{\varepsilon}{5} \qquad \text{if } j \geqslant N \text{ and } x \in K$$

and

$$d(x_{k_j}, a) < \delta \quad \text{and} \quad d(y_{k_j}, a) < \delta \qquad \text{if } j \geqslant N\,.$$

As a consequence, for all $j \geqslant N$,

$$\begin{aligned}\varepsilon \leqslant \|f_{k_j}(x_{k_j}) - f_{k_j}(y_{k_j})\| \leqslant \|f_{k_j}(x_{k_j}) - f(x_{k_j})\| + \|f(x_{k_j}) - f(a)\| \\ + \|f(y_{k_j}) - f(a)\| + \|f(y_{k_j}) - f_{k_j}(y_{k_j})\| < \frac{4\varepsilon}{5}\end{aligned}$$

which is a contradiction. □

Alternative proof of Lemma A.120. Suppose the contrary that B is not equi-continuous. Then $\exists\, \varepsilon > 0$ such that

$$\forall\, k \in \mathbb{N}, \exists\, x_k, y_k \in K \text{ and } f_k \in B \ni d(x_k, y_k) < \frac{1}{k} \text{ but } \|f_k(x_k) - f_k(y_k)\| \geqslant \varepsilon\,.$$

Since B is pre-compact in $\big(\mathscr{C}(K;\mathcal{V}), \|\cdot\|_\infty\big)$, there exists a subsequence $\{f_{k_j}\}_{j=1}^\infty$ converges to some function f in $\big(\mathscr{C}(K;\mathcal{V}), \|\cdot\|_\infty\big)$. By Proposition A.115, $\{f_{k_j}\}_{j=1}^\infty$ converges uniformly to f on K; thus there exists $N_1 > 0$ such that

$$\big\|f_{k_j}(x) - f(x)\big\| < \frac{\varepsilon}{4} \quad \forall\, j \geqslant N_1 \text{ and } x \in K\,.$$

Since $f \in \mathscr{C}(K;\mathcal{V})$ and K is compact, f is uniformly continuous on K; thus

$$\exists\, \delta > 0 \ni \|f(x) - f(y)\| < \frac{\varepsilon}{4} \quad \text{if } d(x,y) < \delta \text{ and } x, y \in K\,.$$

For one of such a δ, there exists $N_2 > 0$ such that

$$d(x_k, y_k) < \delta \quad \forall\, k \geqslant N_2\,.$$

Therefore, $d(x_{k_j}, y_{k_j}) < \delta$ if $j \geqslant N_2$ (this is because $k_j \geqslant j$ for all $j \in \mathbb{N}$); thus for all $j \geqslant \max\{N_1, N_2\}$,

$$\begin{aligned}\varepsilon &\leqslant \|f_{k_j}(x_{k_j}) - f_{k_j}(y_{k_j})\| \\ &\leqslant \|f_{k_j}(x_{k_j}) - f(x_{k_j})\| + \|f(x_{k_j}) - f(y_{k_j})\| + \|f(y_{k_j}) - f_{k_j}(y_{k_j})\| < \frac{3\varepsilon}{4}\end{aligned}$$

which is a contradiction. □

Corollary A.121. *Let (M, d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a normed vector space, and $K \subseteq M$ be a compact subset. If $\{f_k\}_{k=1}^\infty$ converges uniformly on K, then $\{f_k\}_{k=1}^\infty$ is equi-continuous.*

Corollary A.121 shows that if $\{f_k\}_{k=1}^\infty$ converges uniformly on a compact set K, then $\{f_k\}_{k=1}^\infty$ must be equi-continuous. The inverse statement, on the other hand, cannot be true. For example, taking $\{f_k\}_{k=1}^\infty$ to be a sequence of constant functions $f_k(x) = k$. Then $\{f_k\}_{k=1}^\infty$ obviously does not converge, not even any subsequence. Therefore, we would like to study under what additional conditions, equi-continuity of a sequence of functions (defined on a compact set K) indeed converges uniformly. The following lemma is an answer to the question.

Lemma A.122. *Let (M, d) be a metric space, $(\mathcal{V}, \|\cdot\|)$ be a Banach space, $K \subseteq M$ be a compact set, and $\{f_k\}_{k=1}^\infty \subseteq \mathscr{C}(K; \mathcal{V})$ be a sequence of equi-continuous functions. If $\{f_k\}_{k=1}^\infty$ converges pointwise on a dense subset E of K (that is, $E \subseteq K \subseteq \mathrm{cl}(E)$), then $\{f_k\}_{k=1}^\infty$ converges uniformly on K.*

Proof. Let $\varepsilon > 0$ be given. By the equi-continuity of $\{f_k\}_{k=1}^\infty$,

$$\exists\, \delta > 0 \ni \|f_k(x) - f_k(y)\| < \frac{\varepsilon}{3} \quad \text{if } d(x, y) < \delta,\ x, y \in K \text{ and } k \in \mathbb{N}\,.$$

Since K is compact, K is totally bounded; thus

$$\exists \{y_1, \cdots, y_m\} \subseteq K \ni K \subseteq \bigcup_{j=1}^m D\big(y_j, \frac{\delta}{2}\big)\,.$$

By the denseness of E in K, for each $j = 1, \cdots, m$, $\exists\, z_j \in E$ such that $d(z_j, y_j) < \dfrac{\delta}{2}$. Moreover, $D\big(y_j, \dfrac{\delta}{2}\big) \subseteq D(z_j, \delta)$; thus $K \subseteq \bigcup\limits_{j=1}^m D(z_j, \delta)$. Since $\{f_k\}_{k=1}^\infty$ converges pointwise on E, $\{f_k(z_j)\}_{k=1}^\infty$ converges as $k \to \infty$ for all $j = 1, \cdots, m$. Therefore,

$$\exists\, N_j > 0 \ni \|f_k(z_j) - f_\ell(z_j)\| < \frac{\varepsilon}{3} \qquad \forall\, k, \ell \geqslant N_j\,.$$

Let $N = \max\{N_1, \cdots, N_m\}$, then

$$\|f_k(z_j) - f_\ell(z_j)\| < \frac{\varepsilon}{3} \qquad \forall\, k, \ell \geqslant N \text{ and } j = 1, \cdots, m\,.$$

Now we are in the position of concluding the lemma. If $x \in K$, there exists $z_j \in E$ such that $d(x, z_j) < \delta$; thus if we further assume that $k, \ell \geqslant N$,

$$\|f_k(x) - f_\ell(x)\| \leqslant \|f_k(x) - f_k(z_j)\| + \|f_k(z_j) - f_\ell(z_j)\| + \|f_\ell(z_j) - f_\ell(x)\| < \varepsilon\,.$$

By Proposition A.106, $\{f_k\}_{k=1}^\infty$ converges uniformly on K. ▫

Remark A.123. Corollary A.121 and Lemma A.122 suggest that "a sequence $\{f_k\}_{k=1}^\infty \subseteq \mathscr{C}(K; \mathcal{V})$ converges uniformly on K if and only if $\{f_k\}_{k=1}^\infty$ is equi-continuous and pointwise convergent (on a dense subset of K)".

Compact sets in $\mathscr{C}(K; \mathcal{V})$

The next subject in this section is to obtain a (useful) criterion of determining the compactness (or pre-compactness) of a subset $B \subseteq \mathscr{C}(K; \mathcal{V})$ which guarantees the existence of a convergent subsequence $\{f_{k_j}\}_{j=1}^\infty$ of a given sequence $\{f_k\}_{k=1}^\infty \subseteq B$ in $\big(\mathscr{C}(K; \mathcal{V}), \|\cdot\|_\infty\big)$.

Lemma A.124 (Cantor's Diagonal Process)**.** *Let E be a countable set, $(\mathcal{V}, \|\cdot\|)$ be a Banach space, and $f_k : E \to \mathcal{V}$ be a sequence of functions. Suppose that for each $x \in E$, $\{f_k(x)\}_{k=1}^\infty$ is pre-compact in $\mathcal{V}$. Then there exists a subsequence of $\{f_k\}_{k=1}^\infty$ that converges pointwise on E.*

Proof. Since E is countable, $E = \{x_\ell\}_{\ell=1}^\infty$.

1. Since $\{f_k(x_1)\}_{k=1}^\infty$ is pre-compact in $(\mathcal{V}, \|\cdot\|)$, there exists a subsequence $\{f_{k_j}\}_{j=1}^\infty$ such that $\{f_{k_j}(x_1)\}_{j=1}^\infty$ converges in $(\mathcal{V}, \|\cdot\|)$.

2. Since $\{f_k(x_2)\}_{k=1}^\infty$ is pre-compact in $(\mathcal{V}, \|\cdot\|)$, the sequence $\{f_{k_j}(x_2)\}_{j=1}^\infty \subseteq \{f_k(x_2)\}_{k=1}^\infty$ has a convergent subsequence $\{f_{k_{j_\ell}}(x_2)\}_{\ell=1}^\infty$.

Continuing this process, we obtain a sequence of sequences S_1, S_2, $\cdots$ such that

1. S_k consists of a subsequence of $\{f_k\}_{k=1}^\infty$ which converges at x_k, and

2. $S_k \supseteq S_{k+1}$ for all $k \in \mathbb{N}$.

Let g_k be the k-th element of S_k. Then the sequence $\{g_k\}_{k=1}^{\infty}$ is a subsequence of $\{f_k\}_{k=1}^{\infty}$ and $\{g_k\}_{k=1}^{\infty}$ converges at each point of E. ▫

The condition that "$\big\{f_k(x)\big\}_{k=1}^{\infty}$ is pre-compact in $\mathcal{V}$ for each $x \in E$" in Lemma A.124 motivates the following

Definition A.125. Let (M,d) be a metric space, $(\mathcal{V},\|\cdot\|)$ be a normed vector space, and $A \subseteq M$ be a subset. A subset $B \subseteq \mathscr{C}_b(A;\mathcal{V})$ is said to be ***pointwise***
compact / ***pre-compact*** / ***bounded*** if the set $B_x \equiv \big\{f(x) \,\big|\, f \in B\big\}$ is compact / pre-compact / bounded in $(\mathcal{V},\|\cdot\|)$ for all $x \in A$.

Now, we consider compact sets in $\mathscr{C}(K;\mathcal{V})$. Let $B \subseteq \mathscr{C}(K;\mathcal{V})$ be a compact set. Given a sequence $\{f_k\}_{k=1}^{\infty} \subseteq B$, we would like to know if it is possible to find a subsequence $\big\{f_{k_j}\big\}_{j=1}^{\infty}$ which converges in sup-norm. If there is a dense subset E of K such that $\{f_k\}_{k=1}^{\infty}$ $\{f_k\}_{k=1}^{\infty}$ is pointwise pre-compact, by the Diagonal Process (Lemma A.124) we can find a pointwise convergent subsequence $\big\{f_{k_j}\big\}_{j=1}^{\infty}$. With the help of Lemma A.122, we immediately know that by imposing the equi-continuity condition, pointwise convergence implies uniform convergence. Therefore, naturally we require that B satisfies pointwise pre-compactness and equi-continuous to guarantee that B is a compact set of $\mathscr{C}(K;\mathcal{V})$.

The existence of a dense subset of a compact set K is guaranteed by the following

Lemma A.126. *A compact set K in a metric space (M,d) is separable; that is, there exists a countable subset E of K such that* $\mathrm{cl}(E) = K$.

Theorem A.127. *Let (M,d) be a metric space, $(\mathcal{V},\|\cdot\|)$ be a Banach space, $K \subseteq M$ be a compact set, and $B \subseteq \mathscr{C}(K;\mathcal{V})$ be equi-continuous and pointwise pre-compact. Then B is pre-compact in $\big(\mathscr{C}(K;\mathcal{V}),\|\cdot\|_\infty\big)$.*

Proof. We show that every sequence $\{f_k\}_{k=1}^{\infty}$ in B has a convergent subsequence. Since K is compact, there is a countable dense subset E of K (Lemma A.126), and the diagonal process (Lemma A.124) suggests that there exists $\big\{f_{k_j}\big\}_{j=1}^{\infty}$ that converges pointwise on E. Since E is dense in K, by Lemma A.122 $\big\{f_{k_j}\big\}_{j=1}^{\infty}$ converges uniformly on K; thus $\big\{f_{k_j}\big\}_{j=1}^{\infty}$ converges in $\big(\mathscr{C}(K;\mathcal{V}),\|\cdot\|_\infty\big)$ by Proposition A.115. ▫

Remark A.128. Lemma A.120 and Theorem A.127 suggest that "a set $B \subseteq \mathscr{C}(K;\mathcal{V})$ is pre-compact if and only if B is equi-continuous and pointwise pre-compact". (That B is pre-compact implies that B is pointwise pre-compact is left as an exercise).

Corollary A.129. *Let (M,d) be a metric space, and $K \subseteq M$ be a compact set. Assume that $B \subseteq \mathscr{C}(K;\mathbb{R})$ is equi-continuous and pointwise bounded on K. Then every sequence in B has a uniformly convergent subsequence.*

Proof. By the Bolzano-Weierstrass theorem the boundedness of $\{f_k(x)\}_{k=1}^{\infty}$ suggests that $\{f_k(x)\}_{k=1}^{\infty}$ is pre-compact for all $x \in E$. Therefore, we can apply Theorem A.127 under the setting $(\mathcal{V}, \|\cdot\|) = (\mathbb{R}, |\cdot|)$ to conclude the corollary. $\square$

The following theorem is the fundamental result on compact sets in $\mathscr{C}(K;\mathcal{V})$.

Theorem A.130 (The Arzelà-Ascoli Theorem)**.** *Let (M,d) be a metric space, $(\mathcal{V},\|\cdot\|)$ be a Banach space, $K \subseteq M$ be a compact set, and $B \subseteq \mathscr{C}(K;\mathcal{V})$. Then B is compact in $\big(\mathscr{C}(K;\mathcal{V}), \|\cdot\|_\infty\big)$ if and only if B is closed, equi-continuous, and pointwise compact.*

Proof. "$\Leftarrow$" This direction is conclude by Theorem A.127 and the fact that B is closed.

"$\Rightarrow$" By Lemma A.120, it suffices to shows that B is pointwise compact. Let $x \in K$ and $\{f_k(x)\}_{k=1}^{\infty}$ be a sequence in B_x. Since B is compact, there exists a subsequence $\{f_{k_j}\}_{j=1}^{\infty}$ that converges uniformly to some function $f \in B$. In particular, $\{f_{k_j}(x)\}_{j=1}^{\infty}$ converges to $f(x) \in B_x$. In other words, we find a subsequence $\{f_{k_j}(x)\}_{j=1}^{\infty}$ of $\{f_k(x)\}_{k=1}^{\infty}$ that converges to a point in B_x. This implies that B_x is sequentially compact; thus B_x is compact. $\square$

A.3.4 The Contraction Mapping Principle and its Applications

Definition A.131. Let (M,d) be a metric space, and $\Phi : M \to M$ be a mapping. Φ is said to be a ***contraction mapping*** if there exists a constant $k \in [0,1)$ such that

$$d\big(\Phi(x),\Phi(y)\big) \leqslant k d(x,y) \qquad \forall\, x,y \in M\,.$$

Definition A.132. Let (M,d) be a metric space, and $\Phi : M \to M$ be a mapping. A point $x_0 \in M$ is called a ***fixed-point*** for Φ if $\Phi(x_0) = x_0$.

THEOREM A.133 (Contraction Mapping Principle)**.** *Let (M, d) be a complete metric space, and $\Phi : M \to M$ be a contraction mapping. Then Φ has a unique fixed-point.*

Proof. Let $x_0 \in M$, and define $x_{n+1} = \Phi(x_n)$ for all $n \in \mathbb{N} \cup \{0\}$. Then

$$d(x_{n+1}, x_n) = d\big(\Phi(x_n), \Phi(x_{n-1})\big) \leqslant kd(x_n, x_{n-1}) \leqslant k^n d(x_1, x_0)\,;$$

thus if $n > m$,

$$\begin{aligned} d(x_n, x_m) &\leqslant d(x_m, x_{m+1}) + d(x_{m+1}, x_{m+2}) + \cdots + d(x_{n-1}, x_n) \\ &\leqslant (k^m + k^{m+1} + \cdots + k^{n-1}) d(x_1, x_0) \\ &\leqslant k^m(1 + k + k^2 + \cdots) d(x_1, x_0) = \frac{k^m}{1-k} d(x_1, x_0)\,. \end{aligned} \tag{A.23}$$

Since $k \in [0, 1)$, $\lim\limits_{m\to\infty} \dfrac{k^m}{1-k} d(x_1, x_0) = 0$; thus

$$\forall\, \varepsilon > 0, \exists\, N > 0 \ni d(x_n, x_m) < \varepsilon \quad \forall\, n, m \geqslant N\,.$$

In other words, $\{x_n\}_{n=1}^\infty$ is a Cauchy sequence. Since (M, d) is complete, $x_n \to x$ as $n \to \infty$ for some $x \in M$. Finally, since $\Phi(x_n) = x_{n+1}$ for all $n \in \mathbb{N}$, by the continuity of Φ we obtain that

$$\Phi(x) = \lim_{n\to\infty} \Phi(x_n) = \lim_{n\to\infty} x_{n+1} = x$$

which guarantees the existence of a fixed-point.

Suppose that for some $x, y \in M$, $\Phi(x) = x$ and $\Phi(y) = y$. Then

$$d(x, y) = d\big(\Phi(x), \Phi(y)\big) \leqslant kd(x, y)$$

which suggests that $d(x, y) = 0$ or $x = y$. Therefore, the fixed-point of Φ is unique. ▫

REMARK A.134. The proof of the contraction mapping principle also suggests an iterative way, $x_{k+1} = \Phi(x_k)$, of finding the fixed-point of a contraction mapping Φ. Using (A.23), the convergence rate of $\{x_m\}_{m=1}^\infty$ to the fixed-point x is measured by

$$d(x_m, x) = \lim_{n\to\infty} d(x_m, x_n) \leqslant \frac{k^m}{1-k} d(x_1, x_0)\,.$$

Therefore, the smaller the contraction constant k, the faster the convergence.

REMARK A.135. Theorem A.133 sometimes is also called the ***Banach fixed-point theorem***.

The existence and uniqueness of the solution to ODEs

In this sub-section we are concerned with if there is a solution to the initial value problem of ordinary differential equation:

$$x'(t) = f\big(x(t), t\big) \qquad \forall\, t \in [t_0, t_0 + \Delta t], \tag{A.24a}$$

$$x(t_0) = x_0\,, \tag{A.24b}$$

where $x : [t_0, t_0 + \Delta t] \to \mathbb{R}^n$ and $f : \mathbb{R}^n \times [t_0, t_0 + \Delta t] \to \mathbb{R}^n$ are vector-valued functions, and $x_0 \in \mathbb{R}^n$ is a vector. Another question we would like to answer is "if (A.24) indeed has a solution, is the solution unique?"

Theorem A.136 (Fundamental Theorem of ODE)**.** *Suppose that for some $r > 0$, $f : D(x_0, r) \times [t_0, T] \to \mathbb{R}^n$ is continuous and is Lipschitz in the spatial variable; that is,*

$$\exists K > 0 \ni \big\|f(x,t) - f(y,t)\big\|_2 \leqslant K\|x - y\|_2 \qquad \forall\, x, y \in D(x_0, r) \text{ and } t \in [t_0, T]\,.$$

Then there exists $0 < \Delta t \leqslant T - t_0$ such that there exists a unique solution to (A.24).

Proof. For any $x \in \mathscr{C}([t_0, T]; \mathbb{R}^n)$, define

$$\Phi(x)(t) = x_0 + \int_{t_0}^{t} f\big(x(s), s\big)\,ds\,.$$

We note that if $x(t)$ is a solution to (A.24), then x is a fixed point of Φ (for $t \in [t_0, t_0 + \Delta t]$). Therefore, the problem of finding a solution to (A.24) transforms to a problem of finding a fixed-point of Φ.

To guarantee the existence of a unique fixed-point, we appeal to the contraction mapping principle. To be able to apply the contraction mapping principle, we need to specify the metric space (M, d). Let

$$\Delta t = \min\Big\{T - t_0, \frac{r}{Kr + 2\|f(x_0, \cdot)\|_\infty}, \frac{1}{2K}\Big\}\,, \tag{A.25}$$

and define

$$M = \Big\{x \in \mathscr{C}\big([t_0, t_0 + \Delta t]; \mathbb{R}^n\big) \,\Big|\, \|x - x_0\|_\infty \leqslant \frac{r}{2}\Big\}$$

with the metric induced by the sup-norm $\|\cdot\|_\infty$ of $\mathscr{C}\big([t_0, t_0 + \Delta t]; \mathbb{R}^n\big)$. Then

1. We first show that $\Phi : M \to M$. To see this, we observe that

$$
\begin{aligned}
&\big\|\Phi(x) - x_0\big\|_\infty \\
&\quad = \Big\| \int_{t_0}^t f\big(x(s), s\big) ds \Big\|_\infty = \Big\| \int_{t_0}^t \big[f\big(x(s), s\big) - f(x_0, s)\big] ds + \int_{t_0}^t f(x_0, s) ds \Big\|_\infty \\
&\quad \leqslant \int_{t_0}^{t_0+\Delta t} \big\|f\big(x(s), s\big) - f(x_0, s)\big\|_2 ds + \int_{t_0}^{t_0+\Delta t} \big\|f(x_0, s)\big\|_2 ds \\
&\quad \leqslant K \int_{t_0}^{t_0+\Delta t} \|x(s) - x_0\|_2 ds + \Delta t \big\|f(x_0, \cdot)\big\|_\infty \\
&\quad \leqslant \Delta t \Big[K\|x - x_0\|_\infty + \big\|f(x_0, \cdot)\big\|_\infty\Big];
\end{aligned}
$$

thus if $x \in M$, (A.25) implies that $\|\Phi(x) - x_0\|_\infty \leqslant \dfrac{r}{2}$.

2. Next we show that Φ is a contraction mapping. To see this, we compute $\big\|\Phi(x) - \Phi(y)\big\|_\infty$ for $x, y \in M$ and find that

$$
\begin{aligned}
\big\|\Phi(x) - \Phi(y)\big\|_\infty &\leqslant \Big\| \int_{t_0}^t \big[f\big(x(s), s\big) - f\big(y(s), s\big)\big] ds \Big\|_\infty \\
&\leqslant \int_{t_0}^{t_0+\Delta t} K\|x(s) - y(s)\|_2 ds \leqslant K\Delta t \|x - y\|_\infty \leqslant \frac{1}{2}\|x - y\|_\infty\,;
\end{aligned}
$$

thus $\Phi : M \to M$ is a contraction mapping.

3. Finally we show that (M, d) is complete. It suffices to show that M is a closed subset of $\mathscr{C}([t_0, t_0 + \Delta t]; \mathbb{R}^n)$. Let $\{x_k\}_{k=1}^\infty$ be a uniformly convergent sequence with limit x. Since $\|x_k(t) - x_0\|_2 \leqslant \dfrac{r}{2}$ for all $t \in [t_0, t_0 + \Delta t]$, passing k to the limit we find that $\|x(t) - x_0\|_2 \leqslant \dfrac{r}{2}$ for all $t \in [t_0, t_0 + \Delta t]$ which implies that $\|x - x_0\|_\infty \leqslant \dfrac{r}{2}$; thus $x \in M$.

Therefore, by the contraction mapping principle, there exists a unique fixed point $x \in M$ which suggests that there exists a unique solution to (A.24). □

Remark A.137. In the iterative process above of solving ODE, the iterative relation

$$
x_{k+1}(t) = x_0 + \int_{t_0}^t f\big(x_k(s), s\big) ds
$$

is called the ***Picard iteration***.

A.4 The Inverse Function Theorem

The inverse function theorem is the primary tool to determine if a function has an inverse. In general, as long as a function is not injective, we cannot define its inverse. For example, the trigonometric functions $y = \sin x$, $y = \cos x$ and $y = \tan x$ are periodic, so "global" inverse functions do not exist. However, we also know that there are inverse functions of those functions such as $y = \sin^{-1} x = \arcsin x$, $y = \cos^{-1} x = \arctan x$ and $y = \tan^{-1} x = \arctan x$. The existence of such inverse functions is due to the fact that we restrict the domain of the original trigonometric functions so that they becomes one-to-one on those domains (so that their inverse exist).

To know what condition might guarantee the existence of an inverse function in a sufficiently small region, we first look at the case of functions of a single variable. For the inverse function theorem for functions of a single variable to hold, we require that the derivative does not vanish. We conjecture that the non-vanishing derivative condition corresponds to the condition of the invertibility of the bounded linear map $(Df)(x)$.

Suppose that $f \in \mathscr{C}'$, then Theorem A.10 shows that if $(Df)(x_0)$ is invertible, then in a neighborhood of x_0 (Df) is also invertible. Therefore, the following inverse function theorem uses only the condition that (Df) is invertible at one point.

Theorem A.138 (Inverse Function Theorem). *Let $\mathcal{D} \subseteq \mathbb{R}^n$ be open, $x_0 \in \mathcal{D}$, $f : \mathcal{D} \to \mathbb{R}^n$ be of class $\mathscr{C}^1$, and $(Df)(x_0)$ be invertible. Then there exist an open neighborhood $\mathcal{U}$ of x_0 and an open neighborhood $\mathcal{V}$ of $f(x_0)$ such that*

1. *$f : \mathcal{U} \to \mathcal{V}$ is one-to-one and onto;*
2. *The inverse function $f^{-1} : \mathcal{V} \to \mathcal{U}$ is of class $\mathscr{C}^1$;*
3. *If $x = f^{-1}(y)$, then $(Df^{-1})(y) = \big((Df)(x)\big)^{-1}$;*
4. *If f is of class $\mathscr{C}^r$ for some $r > 1$, so is f^{-1}.*

Remark A.139. Since $f^{-1} : \mathcal{V} \to \mathcal{U}$ is continuous, for any open subset $\mathcal{W}$ of $\mathcal{U}$ $f(\mathcal{W}) = (f^{-1})^{-1}(\mathcal{W})$ is open relative to $\mathcal{V}$, or $f(\mathcal{W}) = \mathcal{O} \cap \mathcal{V}$ for some open set $\mathcal{O} \subseteq \mathbb{R}^n$. In other words, if $\mathcal{U}$ is an open neighborhood of x_0 given by the inverse function theorem, then $f(\mathcal{W})$ is also open for all open subsets $\mathcal{W}$ of $\mathcal{U}$. We call this property as f is a ***local open mapping*** at x_0.

Remark A.140. Since $(Df)(x_0) \in \mathscr{B}(\mathbb{R}^n, \mathbb{R}^n)$, the condition that $(Df)(x_0)$ is invertible can be replaced by that the determinant of the Jacobian matrix of f at x_0 is not zero; that is,

$$\det\big(\big[(Df)(x_0)\big]\big) \neq 0 \,.$$

The determinant of the Jacobian matrix of f at x_0 is called the ***Jacobian*** of f at x_0. The Jacobian of f at x is sometimes denoted by $\mathbf{J}_f(x)$ or $\dfrac{\partial(f_1, \cdots, f_n)}{\partial(x_1, \cdots, x_n)}$.

Corollary A.141. *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be open, $f : \mathcal{U} \to \mathbb{R}^n$ be of class $\mathscr{C}^1$, and $(Df)(x)$ be invertible for all $x \in \mathcal{U}$. Then $f(\mathcal{W})$ is open for every open set $\mathcal{W} \subseteq \mathcal{U}$.*

Having established the local version of inverse function theorem for functions of several variables (Theorem A.138), next we focus on the existence of a global inverse function. Based on the inverse function theorem for functions of a single variable, we might conjecture again that the invertibility of $(Df)(x)$ for all x guarantees the existence of a global inverse function. The example below provides a counter-example; in particlar, the condition that (Df) is invertible everywhere does not guarantee the injectivity of functions.

Example A.142. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$f(x, y) = (e^x \cos y, e^x \sin y) \,.$$

Then

$$\big[(Df)(x, y)\big] = \begin{bmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{bmatrix}.$$

It is easy to see that the Jacobian of f at any point is not zero (thus $(Df)(x)$ is invertible for all $x \in \mathbb{R}^2$). On the other hand, f is not globally one-to-one since $f(x, y) = f(x, y + 2\pi)$, and hence cannot have a global inverse.

Theorem A.143 (Global Inverse Function)**.** *Let $\mathcal{D} \subseteq \mathbb{R}^n$ be open, $f : \mathcal{D} \to \mathbb{R}^n$ be of class $\mathscr{C}^1$, and $(Df)(x)$ be invertible for all $x \in K$. Suppose that K is a connected compact subset of $\mathcal{D}$, and $f : \partial K \to \mathbb{R}^n$ is one-to-one. Then $f : K \to \mathbb{R}^n$ is one-to-one.*

Proof. Define $E = \big\{x \in K \,\big|\, \exists\, y \in K, y \neq x \ni f(x) = f(y)\big\}$. Our goal is to show that $E = \varnothing$.

Claim 1: E is closed.

Proof of claim 1: Suppose the contrary that E is not closed. Then there exists $\{x_k\}_{k=1}^{\infty} \subseteq E$, $x_k \to x$ as $k \to \infty$ but $x \in K \backslash E$. Since $x_k \in E$, by the definition of E there exists $y_k \in E$ such that $y_k \neq x_k$ and $f(x_k) = f(y_k)$. By the compactness of K, there exists a convergent subsequence $\big\{y_{k_j}\big\}_{j=1}^{\infty}$ of $\{y_k\}_{k=1}^{\infty}$ with limit $y \in K$. Since $x \notin E$ and $f(x_{k_j}) = f(y_{k_j}) \to f(y)$ as $j \to \infty$, we must have $x = y$; thus $y_{k_j} \to x$ as $j \to \infty$.

Since $(Df)(x)$ is invertible, by the inverse function theorem there exists $\delta > 0$ such that $f : D(x, \delta) \to \mathbb{R}^n$ is one-to-one. By the convergence of sequences $\big\{x_{k_j}\big\}_{j=1}^{\infty}$ and $\big\{y_{k_j}\big\}_{j=1}^{\infty}$, there exists $N > 0$ such that

$$x_{k_j}, y_{k_j} \in D(x, \delta) \qquad \forall\, j \geqslant N\,.$$

This implies that $f : D(x, \delta) \to \mathbb{R}^n$ cannot be one-to-one $\big($since $x_{k_j} \neq y_{k_j}$ but $f\big(x_{k_j}\big) = f\big(y_{k_j}\big)\big)$, a contradiction. Therefore, E is closed.

Claim 2: E is open relative to K; that is, for every $x \in E$, there exists an open set $\mathcal{U}$ such that $x \in \mathcal{U}$ and $\mathcal{U} \cap K \subseteq E$.

Proof of claim 2: Let $x_1 \in E$. Then there is $x_2 \in E$, $x_2 \neq x_1$, such that $f(x_1) = f(x_2)$. Since $(Df)(x_1)$ and $(Df)(x_2)$ are invertible, by the inverse function theorem there exist open neighborhoods $\mathcal{U}_1$ of x_1 and $\mathcal{U}_2$ of x_2, as well as open neighborhoods $\mathcal{V}_1$, $\mathcal{V}_2$ of $f(x_1)$, such that $f : \mathcal{U}_1 \to \mathcal{V}_1$ and $f : \mathcal{U}_2 \to \mathcal{V}_2$ are both one-to-one and onto. Since $x_1 \neq x_2$, W.L.O.G. we can assume that $\mathcal{U}_1 \cap \mathcal{U}_2 = \varnothing$. Since $\mathcal{V}_1 \cap \mathcal{V}_2$ is open, the continuity of f implies that $f^{-1}(\mathcal{V}_1 \cap \mathcal{V}_2) = \mathcal{O} \cap \mathcal{D}$ for some open set $\mathcal{O}$; thus

$$f : \mathcal{U}_1 \cap \mathcal{O} \cap K \to \mathcal{V}_1 \cap \mathcal{V}_2 \cap f(K) \text{ is one-to-one and onto}\,,$$
$$f : \mathcal{U}_2 \cap \mathcal{O} \cap K \to \mathcal{V}_1 \cap \mathcal{V}_2 \cap f(K) \text{ is one-to-one and onto}\,.$$

Let $\mathcal{U} = \mathcal{U}_1 \cap \mathcal{O}$. Then every $x \in \mathcal{U} \cap K$ corresponds to a unique $\widetilde{x} \in \mathcal{U}_2 \cap \mathcal{O} \cap K$ such that $f(x) = f(\widetilde{x})$. Since $\mathcal{U}_1 \cap \mathcal{U}_2 = \varnothing$, we must have $x \neq \widetilde{x}$. Therefore, $x \in E$, or equivalently, $\mathcal{U} \cap K \subseteq E$.

Now we show that $E = \varnothing$. Since K is connected, E is open relative to K and E is closed, $E = K$ or $E = \varnothing$. Suppose the case that $E = K$. Let $x \in \partial K \subseteq E$. Then there exists $y \in E$ such that $y \neq x$ and $f(x) = f(y)$. Since $f : \partial K \to \mathbb{R}^n$ is one-to-one, $y \notin \partial K$. Therefore, we have shown that if $E = K$, then $f(\partial K) \subseteq f\big(\mathrm{int}(K)\big)$.

Since K is compact and f is continuous, $f(K)$ is compact; thus there is $b \in \mathbb{R}^n$ such that $b \notin f(K)$. Consider the function $\varphi : K \to \mathbb{R}$ defined by

$$\varphi(x) = \frac{1}{2}\|f(x) - b\|_{\mathbb{R}^n}^2 = \frac{1}{2}\sum_{j=1}^{n} |f_i(x) - b_i|^2 .$$

Then φ is a continuous function on K; thus φ attains its maximum at $x_0 \in K$. Since $f(\partial K) \subseteq f\big(\mathrm{int}(K)\big)$, we can assume that $x_0 \in \mathrm{int}(K)$; thus $(D\varphi)(x_0) = 0$. As a consequence,

$$\big[(Df)(x_0)\big]^{\mathrm{T}}\big[f(x_0) - b\big] = 0 .$$

By the choice of b, $f(x_0) - b \neq 0$; thus we must have that $(Df)(x_0)$ is not invertible, a contradiction. ▫

The following corollary can be obtained by the extension argument and applying the global inverse function theorem.

COROLLARY A.144. *Let $\Omega \subseteq \mathbb{R}^n$ be a connected $H^{\mathrm{k}+1}$-domain for some $\mathrm{k} > \frac{n}{2}$, and $f : \Omega \to \mathbb{R}^n$ be an $H^{\mathrm{k}+1}$-map such that $\det(\nabla f) > 0$ in $\bar{\Omega}$. If $f : \partial\Omega \to \mathbb{R}^n$ is injective, then $f : \Omega \to \mathbb{R}^n$ is one-to-one.*

A.5 The Monotone and Bounded Convergence Theorems

Now we turn our attention to the validity of (A.22) if $\{f_k\}_{k=1}^{\infty}$ converges pointwise to f. When the uniform convergence is removed from the assumptions, the limit function f may not be Riemann integrable. Moreover, even if the limit function f is Riemann integrable, there are counter-examples for the validity of (A.22). In this section, we provide two theorems in which the conditions of the uniform convergence of the sequence of functions is replaced by some other condition to guarantee the convergence of the integrals to the integral of the limit function.

We begin with the following lemma, which focuses on the approximation of a non-negative bounded function by continuous functions in a certain sense.

LEMMA A.145. *Let $f : [a, b] \to \mathbb{R}$ be a non-negative bounded function. Then for each $\varepsilon > 0$, there exist continuous non-negative functions $g, h : [a, b] \to \mathbb{R}$ such that*

$0 \leqslant g, h \leqslant f$ *and*

$$\underline{\int_a^b} f(x)\,dx < \int_a^b g(x)\,dx + \varepsilon \qquad \text{and} \qquad \overline{\int_a^b} f(x)\,dx > \int_a^b h(x)\,dx - \varepsilon\,.$$

Proof. We only prove the case for the lower integral.

Let $\varepsilon > 0$ be given, and $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b\}$ be a partition of $[a, b]$ such that $L(f, \mathcal{P}) > \underline{\int_a^b} f(x)dx - \frac{\varepsilon}{2}$. Let $s(x)$ be the (non-negative) step function given by

$$s(x) = \sum_{k=1}^{n-1} \inf_{x\in[x_{k-1},x_k]} f(x) \mathbf{1}_{[x_{k-1},x_k)}(x) + \inf_{x\in[x_{n-1},b]} \mathbf{1}_{[x_{n-1},b]}(x)$$

which is a linear combination of characteristic functions. Then

$$\underline{\int_a^b} f(x)\,dx < \int_a^b s(x)\,dx + \frac{\varepsilon}{2} \tag{A.26}$$

since the integral of s over $[a, b]$ is exactly the lower sum $L(f, \mathcal{P})$. On the other hand, for such a simple function s we can always find non-negative continuous function $g : [a, b] \to \mathbb{R}$ (for example, g can be a trapezoidal function) such that $g \leqslant s$ and

$$\int_a^b s(x)\,dx < \int_a^b g(x)\,dx + \frac{\varepsilon}{2}\,. \tag{A.27}$$

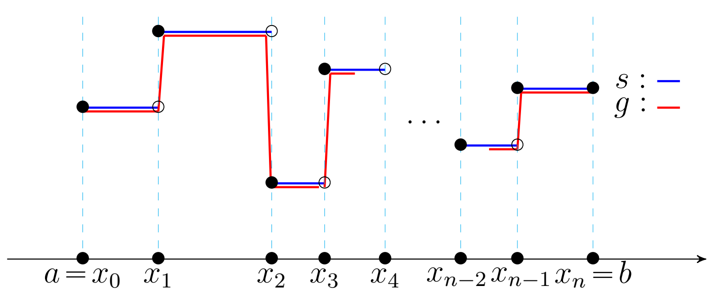


Figure A.3: One way of constructing g given simple function s

The combination of (A.26) and (A.27) then concludes the lemma. □

The Dini Theorem (Theorem A.111) suggests that the monotone pointwise convergence of sequence of continuous functions to continuous functions on a compact set is in fact uniform. Therefore, the monotone pointwise convergence of a sequence of functions seems to possess better convergence behavior. In fact, in terms of the convergence of the integrals this is almost the case, and we have the following

Theorem A.146. *Let $\{f_k\}_{k=1}^\infty$ be a decreasing sequence of bounded functions on $[a,b]$; that is, $f_k \geqslant f_{k+1}$ for all $k \in \mathbb{N}$. If $\lim\limits_{k\to\infty} f_k(x) = 0$ for all $x \in [a,b]$, then*

$$\lim_{k\to\infty} \underline{\int_a^b} f_k(x)\,dx = 0 \ \Big(= \underline{\int_a^b} \lim_{k\to\infty} f_k(x)\,dx \Big) .$$

Proof. Let $\varepsilon > 0$ be given. By Lemma A.145, for each $k \in \mathbb{N}$ there exists a continuous function $g_k : [a,b] \to \mathbb{R}$ such that $0 \leqslant g_k \leqslant f_k$ and

$$\underline{\int_a^b} f_k(x)\,dx < \int_a^b g_k(x)\,dx + \frac{\varepsilon}{2^{k+1}} . \tag{A.28}$$

Define $h_k = \min\{g_1, \cdots, g_k\}$. Then h_k is continuous on $[a,b]$, $h_k \geqslant h_{k+1}$, and $0 \leqslant h_k \leqslant g_k \leqslant f_k$ for all $k \in \mathbb{N}$. Since $\{f_k\}_{k=1}^\infty$ converges pointwise to the zero function, $\lim\limits_{k\to\infty} h_k(x) = 0$ for all $x \in [a,b]$; thus the Dini theorem implies that $\{h_k\}_{k=1}^\infty$ converges uniformly to the zero function on $[a,b]$. Therefore, by Theorem A.110 there exists $N > 0$ such that

$$\int_a^b h_k(x)\,dx < \frac{\varepsilon}{4} \qquad \forall\, k \geqslant N . \tag{A.29}$$

On the other hand, for $1 \leqslant \ell \leqslant k$, $\max\{g_\ell, \cdots, g_k\} \leqslant \max\{f_\ell, \cdots, f_k\} = f_\ell$; thus

$$\int_a^b \big(\max\{g_\ell, \cdots, g_k\} - g_\ell\big)\,dx \leqslant \underline{\int_a^b} f_\ell(x)\,dx - \int_a^b g_\ell(x)\,dx < \frac{\varepsilon}{2^{\ell+1}} .$$

Moreover, for each $1 \leqslant j \leqslant k$,

$$0 \leqslant g_k = g_j + (g_k - g_j) \leqslant g_j + \big(\max\{g_j, \cdots, g_k\} - g_j\big) \leqslant g_j + \sum_{\ell=1}^{k-1} \big(\max\{g_\ell, \cdots, g_k\} - g_\ell\big) ,$$

so for $x \in [a,b]$ and $k \in \mathbb{N}$, choosing $1 \leqslant j \leqslant k$ such that $h_k(x) = g_j(x)$, the inequality above implies that

$$0 \leqslant g_k(x) \leqslant h_k(x) + \sum_{\ell=1}^{k-1} \big(\max\{g_\ell, \cdots, g_k\}(x) - g_\ell(x)\big) \qquad \forall\, x \in [a,b] .$$

As a consequence,

$$0 \leqslant \int_a^b g_k(x)\,dx \leqslant \int_a^b h_k(x)\,dx + \sum_{\ell=1}^{k-1} \frac{\varepsilon}{2^{\ell+1}} \leqslant \int_a^b h_k(x)\,dx + \frac{\varepsilon}{2} ;$$

thus (A.28) and (A.29) imply that

$$0 \leqslant \underline{\int_a^b} f_k(x) dx < \varepsilon \qquad \forall\, k \geqslant N .$$

□

Corollary A.147 (Monotone Convergence Theorem). *Let $f_k : [a,b] \to \mathbb{R}$ be a sequence of Riemann integrable functions such that $\{f_k\}_{k=1}^{\infty}$ converges pointwise to a Riemann integrable function f on $[a,b]$. Suppose that $f_k \leqslant f_{k+1}$ for all $k \in \mathbb{N}$. Then*

$$\int_a^b f(x)\,dx = \lim_{k\to\infty} \int_a^b f_k(x)\,dx\,.$$

Proof. Let $g_k = f - f_k$. Then $\{g_k\}_{k=1}^{\infty}$ is a decreasing sequence of bounded functions on $[a,b]$ and $\lim\limits_{k\to\infty} g_k(x) = 0$ for all $x \in [a,b]$. Therefore, the integrability of f_k and f, as well as Theorem A.146, imply that

$$\begin{aligned} 0 &= \lim_{k\to\infty} \underline{\int_a^b} g_k(x)\,dx = \lim_{k\to\infty} \underline{\int_a^b} (f - f_k)(x)\,dx = \lim_{k\to\infty} \int_a^b (f - f_k)(x)dx \\ &= \int_a^b f(x)\,dx - \lim_{k\to\infty} \int_a^b f_k(x)\,dx\,. \end{aligned}$$ □

Corollary A.148 (Arzelà's Bounded Convergence Theorem). *Let $f_k : [a,b] \to \mathbb{R}$ be a sequence of Riemann integrable functions such that $\{f_k\}_{k=1}^{\infty}$ converges pointwise to a Riemann integrable function f on $[a,b]$. Suppose that there exists a constant $M > 0$ such that $|f_k(x)| \leqslant M$ for all $x \in [a,b]$ and $k \in \mathbb{N}$. Then*

$$\int_a^b f(x)\,dx = \lim_{k\to\infty} \int_a^b f_k(x)\,dx\,.$$

Proof. Let $\varepsilon > 0$ be given. Define $g_n(x) = \sup\limits_{k\geqslant n} |f_k(x) - f(x)|$. Then $\{g_n\}_{n=1}^{\infty}$ is a decreasing sequence of bounded functions on $[a,b]$ and $\lim\limits_{n\to\infty} g_n(x) = 0$ for all $x \in [a,b]$. Therefore, Theorem A.146 implies that there exists $N > 0$ such that

$$\underline{\int_a^b} g_n(x)\,dx < \varepsilon \qquad \forall\, n \geqslant N\,.$$

Moreover, observing that $0 \leqslant |f_k(x) - f(x)| \leqslant g_k(x)$ for all $k \in \mathbb{N}$, by the integrability of f_k and f we conclude that

$$\int_a^b |f_k(x) - f(x)|\,dx = \underline{\int_a^b} |f_k(x) - f(x)|\,dx \leqslant \underline{\int_a^b} g_k(x)\,dx < \varepsilon \qquad \forall\, k \geqslant N\,.$$ □

Combining the Fubini theorem with Theorem A.146, we can conclude the Monotone Convergence Theorem and the Bounded Convergence Theorem for multiple integrals.

Theorem A.149 (Monotone Convergence Theorem)**.** *Let A be a rectangle in $\mathbb{R}^n$, and $f_k : A \to \mathbb{R}$ be Riemann integrable for all $k \in \mathbb{N}$ such that $\{f_k\}_{k=1}^\infty$ converges pointwise to a Riemann integrable function f on A. Suppose $\{f_k\}_{k=1}^\infty$ is a monotone sequence of functions; that is, $f_k \leqslant f_{k+1}$ or $f_k \geqslant f_{k+1}$ for all $k \in \mathbb{N}$. Then*

$$\int_A f(x)\,dx = \lim_{k\to\infty} \int_A f_k(x)\,dx\,.$$

Proof. W.L.O.G. we assume that $f_k \geqslant f_{k+1}$ for all $k \in \mathbb{N}$. We first prove the case $n = 2$ and write $A = [a,b] \times [c,d]$. Define $g_k(x) = \underline{\int_c^d} \big(f_k(x,y) - f(x,y)\big)dy$. Then $g_k \geqslant g_{k+1}$ for all $k \in \mathbb{N}$. Moreover, Theorem A.146 implies that $\{g_k\}_{k=1}^\infty$ converges pointwise to 0, and the Fubini theorem (Theorem A.99) implies that g_k is Riemann integrable over $[a,b]$ for all $k \in \mathbb{N}$. Therefore, by the monotone convergence theorem for functions of one variable (Corollary A.147) we find that

$$\begin{aligned} 0 &= \lim_{k\to\infty} \int_a^b g_k(x)\,dx = \lim_{k\to\infty} \int_a^b \Big(\underline{\int_c^d} \big(f_k(x,y) - f(x,y)\big)dy \Big)dx \\ &= \lim_{k\to\infty} \int_A \big(f_k(x,y) - f(x,y)\big)\,d\mathbb{A}\,. \end{aligned}$$

Now suppose that the conclusion holds for the case $n = N$. Then for $n = N+1$, write $A = R \times [c,d]$ for some rectangle R in $\mathbb{R}^N$, and define g_k by

$$g_k(x_1, \cdots, x_N) = \underline{\int_c^d} \big(f_k(x_1, \cdots, x_{N+1}) - f(x_1, \cdots, x_{N+1})\big)dx_{N+1}\,.$$

Then Theorem A.146 again implies that $\{g_k\}_{k=1}^\infty$ converges monotonically to 0 on R, and the Fubini theorem (Theorem A.99) implies that g_k is Riemann integrable over R for all $k \in \mathbb{N}$. Write $x' = (x_1, \cdots, x_N)$. Then the validity of the monotone convergence theorem for N-tuple integrals implies that

$$\begin{aligned} 0 &= \lim_{k\to\infty} \int_R g_k(x')\,dx' = \lim_{k\to\infty} \int_R \Big(\underline{\int_c^d} \big(f_k(x', x_{N+1}) - f(x', x_{N+1})\big)dx_{N+1} \Big)\,dx' \\ &= \lim_{k\to\infty} \int_A \big(f_k(x) - f(x)\big)\,dx\,. \end{aligned}$$

□

Theorem A.150 (Bounded Convergence Theorem)**.** *Let A be a rectangle in $\mathbb{R}^n$, and $f_k : A \to \mathbb{R}$ be Riemann integrable for all $k \in \mathbb{N}$ such that $\{f_k\}_{k=1}^\infty$ converges pointwise to a Riemann integrable function f on A. Suppose that there exists a constant $M > 0$*

such that $\big|f_k(x)\big| \leqslant M$ *for all* $x \in A$ *and* $k \in \mathbb{N}$*. Then*

$$\int_A f(x)\,dx = \lim_{k\to\infty}\int_A f_k(x)\,dx\,.$$

Proof. Write $A = [a_1, b_1] \times \cdots \times [a_n, b_n]$. For $1 \leqslant j \leqslant n-1$, define

$$g_k^{(j)}(x_1, \cdots, x_j) = \underline{\int_{a_{j+1}}^{b_{j+1}}} \cdots \underline{\int_{a_n}^{b_n}} \sup_{\ell \geqslant k}\big|f_\ell(x_1, \cdots, x_n) - f(x_1, \cdots, x_n)\big|\,dx_n \cdots dx_{j+1}\,.$$

Then $g_k^{(j)} : [a_1, b_1] \times \cdots \times [a_j, b_j] \to \mathbb{R}$ is a bounded decreasing sequence of functions for all $j \in \{1, \cdots, n-1\}$. Moreover, since $\{f_k\}_{k=1}^\infty$ converges pointwise to f, we must have

$$\lim_{k\to\infty}\sup_{\ell\geqslant k}\big|f_\ell(x_1, \cdots, x_n) - f(x_1, \cdots, x_n)\big| = \limsup_{k\to\infty}\big|f_\ell(x_1, \cdots, x_n) - f(x_1, \cdots, x_n)\big| = 0\,.$$

Therefore, Theorem A.146 implies that that $\{g_k^{(n-1)}\}_{k=1}^\infty$ converges pointwise to 0 on $[a_1, b_1] \times \cdots \times [a_{n-1}, b_{n-1}]$. By the fact that

$$g_k^{(j)}(x_1, \cdots, x_j) = \underline{\int_{a_{j+1}}^{b_{j+1}}} g_k^{(j+1)}(x_1, \cdots, x_{j+1})\,dx_{j+1}\,,$$

we also find that $\{g_k^{(n-2)}\}_{k=1}^\infty$ converges pointwise to 0. By induction, we conclude that $\{g_k^{(j)}\}_{k=1}^\infty$ converges pointwise to 0 for all $j \in \{1, \cdots, n-1\}$; thus Theorem A.146 and the Fubini theorem (Theorem A.99) imply that

$$\begin{aligned}
0 &= \lim_{k\to\infty}\underline{\int_{a_1}^{b_1}} g_k^{(1)}(x_1)\,dx_1 \\
&= \lim_{k\to\infty}\underline{\int_{a_1}^{b_1}} \cdots \underline{\int_{a_n}^{b_n}} \sup_{\ell\geqslant k}\big|f_\ell(x_1, \cdots, x_n) - f(x_1, \cdots, x_n)\big|\,dx_n \cdots dx_1 \\
&\geqslant \limsup_{k\to\infty}\underline{\int_{a_1}^{b_1}} \cdots \underline{\int_{a_n}^{b_n}} \big|f_k(x_1, \cdots, x_n) - f(x_1, \cdots, x_n)\big|\,dx_n \cdots dx_1 \\
&\geqslant \limsup_{k\to\infty}\underline{\int_A} \big|f_k(x) - f(x)\big|dx = \limsup_{k\to\infty}\int_A \big|f_k(x) - f(x)\big|dx\,.
\end{aligned}$$

□

REMARK A.151. 1. If A is a bounded set with volume, we can choose a rectangle $S \supseteq A$ and consider $g_k = \overline{f}_k^A$ as well as $g = \overline{f}^A$. Then $g_k, g : S \to \mathbb{R}$ satisfy the assumptions in Theorem A.149 and A.150; thus Lemma A.96 implies that

$$\lim_{k\to\infty}\int_A f_k(x)\,dx = \lim_{k\to\infty}\int_R g_k(x)\,dx = \int_R g(x)\,dx = \int_A f(x)\,dx\,.$$

In other words, the Monotone Convergence Theorem and the Bounded Convergence Theorem holds for more general domain A, or to be more precise, for bounded set A with volume.

2. The Monotone Convergence Theorem (MCT) can be viewed as a corollary of the Bounded Convergence Theorem (BCT) since under the assumptions of MCT, we can apply BCT $\big($choose $M = \max\big\{\sup\limits_{x\in A} f(x), \sup\limits_{x\in A} f_1(x)\big\}\big)$ directly to conclude the MCT. Here we prove MCT without the help of BCT to demonstrate the power of the Fubini Theorem.

   On the other hand, unlike the case in the theory of Lebesgue integrals we cannot prove BCT using MCT since the infimum or supremum of a sequence of Riemann integrable functions might not be Riemann integrable anymore.

A.6 Improper Integrals

The Riemann integral deals with the "integrals" of bounded functions over bounded sets; however, often times we need to integrate unbounded functions over unbounded sets, such as finding the area under an unbounded function above x-axis. The improper integral is an answer to this particular situation.

A.6.1 Definition and basic properties

Let $A \subseteq \mathbb{R}^n$ be a set whose boundary ∂A has measure zero, and $f : A \to \mathbb{R}$ be non-negative such that the collection of points of discontinuity of f has measure zero; that is,

$$\text{the set } \big\{x \in A \,\big|\, f \text{ is discontinuous at } x\big\} \text{ has measure zero}.$$

If A is bounded and f is bounded, then the Lebesgure Theorem (Theorem A.77 or Remark A.81) implies that f is Riemann integrable over A. Now suppose that A is unbounded but f is still bounded. Define $A_k = A \cap D(0,k)$ and $f_k = f 1_{A_k}$; that is, f_k is the restriction of f to A_k or

$$f_k(x) = \begin{cases} f(x) & \text{if } x \in A_k\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{A.30}$$

By the fact that $\partial(A \cap D(0,k)) \subseteq \partial A \cup \partial D(0,k)$, ∂A_k has measure zero for all $k > 0$; thus Corollary A.79 implies that A_k has volume. By the fact that

$$\{x \in A_k \,|\, f_k \text{ is discontinuous at } x\} \subseteq \partial A_k \cup \{x \in A \,|\, f \text{ is discontinuous at } x\},$$

we find that f_k is Riemann integrable over A_k, and Theorem A.97 implies that

$$\int_{A_k} f(x)\,dx = \int_{A_k} f_k(x)\,dx = \int_A f_k(x)\,dx\,.$$

Since $0 \leqslant f_k \leqslant f_{k+1}$ for all $k \in \mathbb{N}$ and $\{f_k\}_{k=1}^\infty$ converges pointwise to f, in view of the Monotone Convergence Theorem (Theorem A.149) (which has to be proved for the case of improper integrals), we intend to define the integral of (bounded function) f over (unbounded set) A by

$$\int_A f(x)\,dx = \lim_{k\to\infty} \int_A f_k(x)\,dx$$

if the limit exists. We remark that if f is not sign-definite, then $\{f_k\}_{k=1}^\infty$ defined by (A.30) no longer converges monotonically to f.

Now suppose that A is bounded but f is unbounded. Define f_k by

$$f_k(x) = (f \wedge k)(x) = \min\{f(x), k\} = \begin{cases} f(x) & \text{if } f(x) \leqslant k\,, \\ k & \text{otherwise}\,. \end{cases} \tag{A.31}$$

Then clearly $\{x \in A \,|\, f_k \text{ is discontinuous at } x\} \subseteq \{x \in A \,|\, f \text{ is discontinuous at } x\}$; thus f_k is Riemann integrable over A. Since $0 \leqslant f_k \leqslant f_{k+1}$ for all $k \in \mathbb{N}$ and $\{f_k\}_{k=1}^\infty$ converges pointwise to f, in view of the Monotone Convergence Theorem (Theorem A.149) again we intend to define the integral of (unbounded function) f over (bounded set) A by

$$\int_A f(x)\,dx = \lim_{k\to\infty} \int_A f_k(x)\,dx$$

if the limit exists.

Remark A.152. Instead of (A.31), one might want to define f_k by

$$f_k(x) = \begin{cases} f(x) & \text{if } f(x) \leqslant k\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

The sequence $\{f_k\}_{k=1}^\infty$ still monotonically converges to f; however, it is not easy to see if the collection of points of discontinuity of f_k has measure zero since the set

$\{x \in A \,|\, f(x) = k\}$ could be large. In other words, by defining f_k in this way we do not know the integrability of f_k over A; thus it is meaningless to define the improper integral as the limit of $\displaystyle\int_A f_k(x)\,dx$.

Finally we consider the case that A is an unbounded set and f is unbounded on A. Several ways of defining the improper integrals are possible. One approach is to define the improper integral of f over A by

$$\int_A f(x)\,dx = \lim_{k\to\infty} \int_{A\cap D(0,k)} f(x)\,dx\,,$$

where $\displaystyle\int_{A\cap D(0,k)} f(x)\,dx$ is itself the improper integral of (unbounded) f over (bounded) $A \cap D(0,k)$. Another approach is to define the improper integral of f over A by

$$\int_A f(x)\,dx = \lim_{k\to\infty} \int_A (f \wedge k)(x)\,dx\,,$$

where $\displaystyle\int_A (f \wedge k)(x)\,dx$ is itself the improper integral of (bounded) $(f \wedge k)$ over (unbounded) A. We shall prove that these two approaches lead to the same limit.

Suppose that $\alpha = \displaystyle\lim_{k\to\infty} \int_{A\cap D(0,k)} f(x)\,dx$ and $\beta = \displaystyle\lim_{k\to\infty} \int_A (f \wedge k)(x)\,dx$ for some $\alpha, \beta \in \mathbb{R}$. Let $\varepsilon > 0$ be given. Then there exists $L > 0$ such that

$$0 \leqslant \alpha - \int_{A\cap D(0,\ell)} f(x)\,dx < \frac{\varepsilon}{2} \qquad \forall\, \ell \geqslant L\,.$$

By the definition of improper integral of unbounded function over bounded set, for each $\ell \in \mathbb{N}$ there exists $N(\ell) > 0$ such that

$$0 \leqslant \int_{A\cap D(0,\ell)} f(x)\,dx - \int_{A\cap D(0,\ell)} (f \wedge k)(x)\,dx < \frac{\varepsilon}{2} \qquad \forall\, k \geqslant N(\ell)\,.$$

W.L.O.G. we can assume that $N(\ell) \geqslant \ell$; thus if $k \geqslant N(L)$,

$$\begin{aligned}
0 &\leqslant \alpha - \int_{A\cap D(0,k)} (f \wedge k)(x)\,dx \\
&= \alpha - \int_{A\cap D(0,L)} f(x)\,dx + \int_{A\cap D(0,L)} f(x)\,dx - \int_{A\cap D(0,L)} (f \wedge k)(x)\,dx \\
&\quad + \int_{A\cap D(0,L)} (f \wedge k)(x)\,dx - \int_{A\cap D(0,k)} (f \wedge k)(x)\,dx \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} + \int_{A\cap D(0,L)} (f \wedge k)(x)\,dx - \int_{A\cap D(0,k)} (f \wedge k)(x)\,dx \leqslant \varepsilon\,,
\end{aligned}$$

where the last inequality is concluded by the non-negativity of $f \wedge k$ and the inclusion $D(0,L) \subseteq D(0,k)$. In other words, we proved that

$$\alpha = \lim_{k\to\infty} \int_{A\cap D(0,k)} f(x)\,dx = \lim_{k\to\infty} \int_{A\cap D(0,k)} (f\wedge k)(x)\,dx\,. \tag{A.32}$$

Similar argument can be used to show that $\beta = \lim\limits_{k\to\infty} \int_{A\cap D(0,k)} (f\wedge k)(x)\,dx$; thus we established the identity $\alpha = \beta$. We note that the integral $\int_{A\cap D(0,k)} (f\wedge k)(x)\,dx$ is monotone increasing in k by the non-negativity of f, so the limit $\lim\limits_{k\to\infty} \int_{A\cap D(0,k)} (f\wedge k)(x)\,dx$ either is finite or diverges to infinite.

The discussion above motivates the following

Definition A.153. Let $A \subseteq \mathbb{R}^n$ be a set whose boundary ∂A has measure zero, and $f : A \to \mathbb{R}$ be a non-negative function such that the collection of points of discontinuity of f has measure zero. f is said to be ***integrable*** over A if the limit

$$\int_A f(x)\,dx \equiv \lim_{k\to\infty} \int_{A\cap D(0,k)} (f\wedge k)(x)\,dx \tag{A.33}$$

is finite, and in such a case $\int_A f(x)\,dx$ is called the integral of f over A.

Remark A.154. 1. For non-negative function $f : A :\to \mathbb{R}$ (with f and A satisfying assumptions in Definition A.153), if the limit $\int_{A\cap D(0,k)} f(x)\,dx$ is infinite, we still call $\int_A f(x)\,dx$ the integral of f over A. However, in this case f is not integrable over A.

2. Let $A \subseteq \mathbb{R}^n$ and $f : A \to \mathbb{R}$ be given in Definition A.153. If $F \subseteq A$ is a measure zero set whose boundary ∂F also has measure zero, then

$$\int_F f(x)\,dx = \lim_{k\to\infty} \int_{F\cap D(0,k)} (f\wedge k)(x)\,dx = 0\,,$$

where Theorem A.91 is used to evaluate the integral.

3. By the Monotone Convergence Theorem (Theorem A.149), (A.33) always holds if $f : A \to \mathbb{R}$ is Riemann integrable; thus Riemann integrable functions are integrable.

Note that by defining the improper integrals in this way, several properties listed in Theorem A.90 also hold. For example, if $0 \leqslant f \leqslant g$ and $f, g : A \to \mathbb{R}$ are integrable, then $\int_A f(x)\,dx \leqslant \int_A g(x)\,dx$. Moreover, a result similar to Theorem A.97 also holds. To be more precise , we state the result as follows.

Theorem A.155. *Let $A, B \subseteq \mathbb{R}^n$ be sets whose boundaries ∂A and ∂B have measure zero, and $f : A \cup B \to \mathbb{R}$ be a non-negative function such that the collection of points of discontinuity of f has measure zero. If $A \cap B$ has measure zero, then*

$$\int_{A\cup B} f(x)\,dx = \int_A f(x)\,dx + \int_B f(x)\,dx\,.$$

Proof. To simplify the notation, for each $k \in \mathbb{N}$ we let $f_k = f \wedge k$, and $A_k = A \cap D(0,k)$ as well as $B_k = B \cap D(0,k)$. Since

$$\partial(A_k \cup B_k) = \partial\big((A \cup B) \cap D(0,k)\big) \subseteq \partial(A \cup B) \cup \partial D(0,k) \subseteq \partial A \cup \partial B \cup \partial D(0,k)\,,$$
$$\partial(A_k \cap B_k) = \partial\big((A \cap B) \cap D(0,k)\big) \subseteq \partial(A \cap B) \cup \partial D(0,k) \subseteq \partial A \cup \partial B \cup \partial D(0,k)\,,$$

we find that under the assumptions of this theorem, $A_k \cup B_k$ and $A_k \cap B_k$ have volume for each $k \in \mathbb{N}$. Therefore, Remark A.81 implies that for each $k \in \mathbb{N}$, $f_k 1_{A_k}$, $f_k 1_{B_k}$ and $f_k 1_{A_k \cap B_k}$ are all Riemann integrable over $A_k \cup B_k$. Since $A_k \cap B_k$ has measure zero, Theorem A.97 implies that

$$\begin{aligned}\int_{(A\cup B)\cap D(0,k)} (f\wedge k)\,dx &= \int_{A_k\cup B_k} f_k(x)\,dx = \int_{A_k} f_k(x)\,dx + \int_{B_k} f_k(x)\,dx \\ &= \int_{A\cap D(0,k)} (f\wedge k)\,dx + \int_{B\cap D(0,k)} (f\wedge k)\,dx\,,\end{aligned}$$

and the theorem is concluded by passing k to the limit. □

Example A.156. Let $f : [1,\infty) \to \mathbb{R}$ be given by $f(x) = x^p$ for some $p \in \mathbb{R}$. If $p > 0$, then f is unbounded, and in this case

$$(f \wedge k)(x) = \begin{cases} x^p & \text{if } 1 \leqslant x \geqslant k^{\frac{1}{p}}\,, \\ k & \text{if } x > k^{\frac{1}{p}}\,; \end{cases}$$

thus

$$\int_{[1,\infty)\cap(-k,k)} (f\wedge k)(x)\,dx = \int_1^{k^{\frac{1}{p}}} x^p dx + \int_{k^{\frac{1}{p}}}^k k dx = \frac{1}{p+1}(k^{1+\frac{1}{p}} - 1) + k(k - k^{\frac{1}{p}})$$

whose limit (as $k \to \infty$) does not exist.

When $p \leqslant 0$, f is bounded by 1 on $[1,\infty)$. Therefore,

$$\int_{[1,\infty)\cap(-k,k)} (f \wedge k)(x)\,dx = \int_1^k f(x)dx = \begin{cases} \dfrac{1}{p+1}(k^{p+1}-1) & \text{if } p \neq -1\,, \\ \log k & \text{if } p = -1\,. \end{cases}$$

It is easy to see that the limit (as $k \to \infty$) exists only when $p < -1$. Therefore, f is integrable over $(0,1)$ if $p < -1$, and in this case

$$\int_{[1,\infty)} f(x)\,dx = \lim_{k\to\infty} \frac{1}{p+1}(k^{p+1}-1) = -\frac{1}{p+1}\,.$$

Example A.157. Let $f : (0,1) \to \mathbb{R}$ be given by $f(x) = x^p$ for some $p \in \mathbb{R}$. If $p \geqslant 0$, f is continuous on $(0,1)$, so f is Riemann integrable over $(0,1)$. If $p < 0$, f is unbounded on $(0,1)$, so the Riemann integral of f no longer makes sense. Nevertheless, we can find the improper integral of f using (A.33): for each $k \in \mathbb{N}$,

$$(f \wedge k)(x) = \begin{cases} x^p & \text{if } x \geqslant k^{\frac{1}{p}}\,, \\ k & \text{if } 0 < x < k^{\frac{1}{p}}\,; \end{cases}$$

thus

$$\int_0^1 (f \wedge k)(x)\,dx = \int_0^{k^{\frac{1}{p}}} k dx + \int_{k^{\frac{1}{p}}}^1 x^p dx = \begin{cases} \dfrac{1}{p+1}(pk^{1+\frac{1}{p}}+1) & \text{if } p \neq -1\,, \\ 1+\log k & \text{if } p = -1\,. \end{cases}$$

It is easy to see that the limit (as $k \to \infty$) exists only when $p > -1$. Therefore, f is integrable over $(0,1)$ if $p > -1$, and in this case

$$\int_{(0,1]} f(x)dx = \lim_{k\to\infty} \frac{1}{p+1}(pk^{1+\frac{1}{p}}+1) = \frac{1}{p+1}\,.$$

Those who are familiar with the improper integrals introduced in Calculus might be confused with the way we compute the improper integrals in Example A.156 and A.157. In fact, there are other ways of evaluating the improper integrals for functions of one variable, and the following theorem is useful for this particular purpose.

Theorem A.158. 1. *Let $f : [a,\infty) \to \mathbb{R}$ be bounded, non-negative, and continuous except perhaps on a set of measure zero. Then*

$$\int_{[a,\infty)} f(x)\,dx = \lim_{R\to\infty} \int_a^R f(x)\,dx\,. \tag{A.34}$$

2. *Let $f : (a,b] \to \mathbb{R}$ be non-negative, bounded on $[a+\varepsilon, b]$ for all $\varepsilon > 0$, and continuous except perhaps on a set of measure zero. Then*
$$\int_{(a,b]} f(x)\,dx = \lim_{\varepsilon\to 0^+} \int_{a+\varepsilon}^b f(x)\,dx\,. \tag{A.35}$$

Proof. 1. For each $R > 0$ sufficiently large, define $M_R = \sup\limits_{x\in[a,R]} f(x)$. Then for large $R > 0$,
$$\int_a^R f(x)\,dx = \int_a^R (f\wedge M_R)(x)\,dx \leqslant \int_{[a,\infty)} f(x)\,dx\,.$$
Passing R to the limit, we find that
$$\lim_{R\to\infty} \int_a^R f(x)\,dx \leqslant \int_{[a,\infty)} f(x)\,dx\,. \tag{A.36}$$
On the other hand, note that for $k \geqslant |a|$,
$$\int_{(a,\infty]\cap(-k,k)} (f\wedge k)(x)\,dx = \int_a^k (f\wedge k)(x)\,dx \leqslant \int_a^k f(x)\,dx \leqslant \lim_{R\to\infty}\int_a^R f(x)\,dx\,.$$
Passing k to the limit, we find that
$$\int_{[a,\infty)} f(x)\,dx = \lim_{k\to\infty}\int_{(a,\infty]\cap(-k,k)} (f\wedge k)(x)\,dx \leqslant \lim_{R\to\infty}\int_a^R f(x)\,dx\,. \tag{A.37}$$
Combining (A.36) and (A.37), we conclude (A.34).

2. For each $\varepsilon > 0$ sufficiently small, define $M_\varepsilon \equiv \sup\limits_{x\in[a+\varepsilon,b]} f(x)$. Then for small $\varepsilon > 0$,
$$\int_{a+\varepsilon}^b f(x)\,dx = \int_{a+\varepsilon}^b (f\wedge M_\varepsilon)(x)\,dx \leqslant \int_{(a,b]} f(x)\,dx\,.$$
Passing ε to the limit, we find that
$$\lim_{\varepsilon\to 0^+}\int_{a+\varepsilon}^b f(x)\,dx \leqslant \int_{(a,b]} f(x)\,dx\,. \tag{A.38}$$
On the other hand, note that
$$\int_a^b (f\wedge k)(x)\,dx = \lim_{\varepsilon\to 0^+}\int_{a+\varepsilon}^b (f\wedge k)(x)\,dx \leqslant \lim_{\varepsilon\to 0^+}\int_{a+\varepsilon}^b f(x)\,dx\,.$$
Passing k to the limit, we find that
$$\int_a^b f(x)\,dx = \lim_{\varepsilon\to 0^+}\int_{(a,b]\cap(-k,k)} (f\wedge k)(x)\,dx \leqslant \lim_{\varepsilon\to 0^+}\int_{a+\varepsilon}^b f(x)\,dx\,. \tag{A.39}$$
Combining (A.38) and (A.39), we concluded (A.35). □

Remark A.159. In view of (A.34) and (A.35), we also have the following notation for improper integrals for functions of one variable:

$$\int_a^\infty f(x)\,dx \equiv \int_{[a,\infty)} f(x)\,dx \qquad \text{and} \qquad \int_a^b f(x)\,dx = \int_{(a,b]} f(x)\,dx\,.$$

Example A.160. Let $f(x) = x^p$ as in Example A.156 and A.157. Since

$$\int_1^R x^p\,dx = \begin{cases} \dfrac{1}{p+1}(R^{p+1}-1) & \text{if } p \neq -1\,, \\ \log R & \text{if } p = -1\,, \end{cases}$$

and

$$\int_\varepsilon^1 x^p\,dx = \begin{cases} \dfrac{1}{p+1}(1-\varepsilon^{1+p}) & \text{if } p \neq -1\,, \\ -\log\varepsilon & \text{if } p = -1\,, \end{cases}$$

by Theorem A.158 we find that f is integrable over $[1,\infty)$ if and only if $p < -1$ and f is integrable over $(0,1]$ if and only if $p > -1$. These are the conclusions that we have obtained in Example A.156 and A.157.

Example A.161 (The Gamma function)**.** For each $t > 0$, define $\Gamma(t) = \displaystyle\int_0^\infty x^{t-1}e^{-x}\,dx$.

1. For $1 \leqslant t < \infty$, the integrand is bounded and non-negative. In fact, $x^{t-1}e^{-x} \leqslant M_t e^{-\frac{x}{2}}$ for some constant $M_t > 0$ (we can choose $M_t = \sup\limits_{x\in[0,\infty)} x^{t-1}e^{-\frac{x}{2}}$). Since
$$\int_0^R x^{t-1}e^{-x}\,dx \leqslant \int_0^R M_t e^{-\frac{x}{2}}dx \leqslant -2M_t e^{-\frac{x}{2}}\Big|_{x=0}^{x=R} \leqslant 2M_t < \infty\,;$$
we find that $\Gamma(t)$ is well-defined for $1 \leqslant t < \infty$.

2. For $0 < t < 1$, the integrand is unbounded near 0; thus by Theorem A.155 we rewrite
$$\int_0^\infty x^{t-1}e^{-x}\,dx = \int_0^1 x^{t-1}e^{-x}\,dx + \int_1^\infty x^{t-1}e^{-\frac{x}{2}}e^{-\frac{x}{2}}dx\,.$$
Since $x^{t-1}e^{-x} \leqslant x^{t-1}$ on $(0,1]$ and $x^{t-1}e^{-x} \leqslant e^{-x}$ on $[1,\infty)$, for $\varepsilon > 0$ we have
$$\int_\varepsilon^1 x^{t-1}e^{-x}dx \leqslant \int_\varepsilon^1 x^{t-1}dx = \frac{1}{t}x^t\Big|_{x=\varepsilon}^1 = \frac{1}{t}(1-\varepsilon^t) \leqslant \frac{1}{t}$$
and for $R > 1$,
$$\int_1^R x^{t-1}e^{-x}dx \leqslant \int_1^R e^{-x}dx = -e^{-x}\Big|_{x=1}^{x=R} = e^{-1} - e^{-R} \leqslant e^{-1}\,.$$
Therefore, $\Gamma(t)$ is also well-defined for $0 < t < 1$.

The following theorem provides different ways of computing the improper (multiple) integrals.

Theorem A.162. *Let $A \subseteq \mathbb{R}^n$ be a set whose boundary ∂A has measure zero, and $f : A \to \mathbb{R}$ be non-negative such that the collection of points of discontinuity of f has measure zero. Then f is integrable over A if and only if for each sequence $\{B_k\}_{k=1}^\infty \subseteq \mathbb{R}^n$ of bounded sets with volume satisfying*

1. *$B_k \subseteq B_{k+1}$ for all $k \in \mathbb{N}$;*
2. *for all $R > 0$ we have $D(0,R) \subseteq B_k$ for sufficient large $k \in \mathbb{N}$;*

the limit $\displaystyle\lim_{k\to\infty} \int_{A\cap B_k} (f \wedge k)(x)\, dx$ exists.

Proof. "$\Leftarrow$" Simply choose $B_k = D(0,k)$ to conclude the integrability of f over A.

"$\Rightarrow$" For each $\ell \in \mathbb{N}$, there exists $N(\ell) \geqslant \ell$ such that $D(0,\ell) \subseteq B_k$ for all $k \geqslant N(\ell)$. Then

$$\int_{A\cap D(0,\ell)} (f\wedge\ell)(x)\, dx \leqslant \int_{A\cap B_k} (f\wedge\ell)(x)\, dx \leqslant \int_{A\cap B_k} (f\wedge k)(x)\, dx \qquad \forall\, k \geqslant N(\ell)\,.$$

Since $\displaystyle\int_{A\cap B_k} (f\wedge k)(x)\, dx = \int_A \big((f\wedge k)1_{B_k}\big)(x)\, dx \leqslant \int_A f(x)\, dx$, by the sandwich lemma we conclude that

$$\int_A f(x)\, dx = \lim_{\ell\to\infty} \int_{A\cap D(0,\ell)} (f\wedge\ell)(x)\, dx = \lim_{k\to\infty} \int_{A\cap B_k} (f\wedge k)(x)\, dx\,. \qquad \square$$

In other words, as long as $\{B_k\}_{k=1}^\infty$ "expands to the whole space", one can evaluate the improper integral using

$$\int_A f(x)\, dx = \lim_{k\to\infty} \int_{A\cap B_k} (f\wedge k)(x)\, dx\,.$$

One particular sequence of sets $\{B_k\}_{k=1}^\infty$ is given by $B_k = [-k,k] \times \cdots \times [-k,k]$.

Example A.163. Consider the improper integral $\displaystyle\int_{-\infty}^{\infty} e^{-x^2}\, dx$. Instead of evaluating this improper integral directly, we consider the improper integral $\displaystyle\int_{\mathbb{R}^2} e^{-(x^2+y^2)}\, d\mathbb{A}$. Note that Theorem A.162 suggests that

$$\int_{\mathbb{R}^2} e^{-(x^2+y^2)}\, d\mathbb{A} = \lim_{k\to\infty} \int_{[-k,k]\times[-k,k]} e^{-(x^2+y^2)}\, d\mathbb{A} = \lim_{k\to\infty} \int_{D(0,k)} e^{-(x^2+y^2)}\, d\mathbb{A}\,.$$

By the Fubini theorem,

$$\lim_{k\to\infty}\int_{[-k,k]\times[-k,k]} e^{-(x^2+y^2)}\,d\mathbb{A} = \lim_{k\to\infty}\int_{-k}^{k}\Big(\int_{-k}^{k} e^{-(x^2+y^2)}dy\Big)dx = \Big(\int_{-\infty}^{\infty} e^{-x^2}dx\Big)^2,$$

while the change of variables formula (with $(x,y)=(r\cos\theta, r\sin\theta)$) implies that

$$\begin{aligned}\lim_{k\to\infty}\int_{D(0,k)} e^{-(x^2+y^2)}\,d\mathbb{A} &= \lim_{k\to\infty}\int_{[0,k]\times[0,2\pi]} e^{-r^2}r\,d(r,\theta) = \lim_{k\to\infty}\int_0^{2\pi}\Big(\int_0^k e^{-r^2}rdr\Big)d\theta \\ &= \lim_{k\to\infty}\int_0^{2\pi}\Big(\frac{e^{-r^2}}{-2}\Big|_{r=0}^{r=k}\Big)d\theta = \lim_{k\to\infty}\pi(1-e^{-k^2}) = \pi\,.\end{aligned}$$

Since $\displaystyle\int_{-\infty}^{\infty} e^{-x^2}\,dx \geqslant 0$, we must have

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}\,.$$

Now we define the improper integrals for general functions. We imitate the idea of truncation (that is, $f\wedge k$ if $f\geqslant 0$) and define

$$f_k(x) = \begin{cases} f(x) & \text{if } |f(x)|\leqslant k\,,\\ k & \text{if } f(x)>k\,,\\ -k & \text{if } f(x)<-k\,.\end{cases}$$

Note that $f_k = (-k)\vee(f\wedge k) = (f^+\wedge k) - (f^-\wedge k)$, where $\vee$ outputs the maximum of values from both sides of $\vee$, and $f^+ = \max\{f,0\} = f\vee 0$ and $f^- = \max\{-f,0\} = (-f)\vee 0$ are the positive part and the negative part of f, respectively. Moreover, if the collection of points of discontinuity of f has measure zero, so does the collection of points of discontinuity of f_k. In other words, f_k is integrable over $A\cap D(0,k)$ for all $k\in\mathbb{N}$.

As in the previous discussion, we intend to define the improper integral of f over A as the limit of $\displaystyle\int_{A\cap D(0,k)} f_k(x)\,dx$ (as $k\to\infty$) provided that the limit exists.

Definition A.164. Let $A\subseteq\mathbb{R}^n$ be a set whose boundary ∂A has measure zero, and $f: A\to\mathbb{R}$ be a function such that the collection of points of discontinuity of f has measure zero. f is said to be integrable if the limit

$$\begin{aligned}\int_A f(x)\,dx &\equiv \lim_{k\to\infty}\int_{A\cap D(0,k)}\big((-k)\vee(f\wedge k)\big)(x)dx \\ &= \lim_{k\to\infty}\int_{A\cap D(0,k)}\big((f^+\wedge k)(x) - (f^-\wedge k)(x)\big)dx\end{aligned}$$

exists, and in such case $\displaystyle\int_A f(x)\,dx$ is called the integral of f over A.

Note that if f is not sign-definite, the integral $\displaystyle\int_{A\cap D(0,k)}\big((-k)\vee(f\wedge k)(x)\big)dx$ is in general not monotone in k; thus in general we do not even know if the limit superior and limit inferior of the integral agree. The following test, similar to the one introduced in the series, provided a useful criterion for the integrability of functions.

Theorem A.165 (Comparison Test)**.** *Let $A \subseteq \mathbb{R}^n$ be a set whose boundary ∂A has measure zero, $f, g : A \to \mathbb{R}$ be functions such that the collection of points of discontinuity of f and g have measure zero. If $|f| \leqslant g$ on A and g is integrable over A, then f is integrable over A.*

Proof. To simplify the notation, for each $k \in \mathbb{N}$ we let $f_k = (-k)\vee f \wedge k$ and define $I_k = \displaystyle\int_{A\cap D(0,k)} f_k(x)\,dx$. Since

$$\big|\big((-k)\vee\varphi\wedge k\big)(x)-\varphi(x)\big| \leqslant \psi(x)-(\psi\wedge k)(x) \qquad \text{if } |\varphi|\leqslant\psi \text{ and } k\in\mathbb{N}\,, \tag{A.40}$$

where the validity of (A.40) is left as an exercise, we have for $k \geqslant \ell$,

$$\begin{aligned}\big|f_k(x)-f_\ell(x)\big| &\leqslant \big|f_k(x)-f(x)\big|+\big|f_\ell(x)-f(x)\big| \leqslant g(x)-(g\wedge k)(x)+g(x)-(g\wedge \ell)(x)\\ &\leqslant 2\big(g(x)-(g\wedge\ell)(x)\big)\,.\end{aligned}$$

As a consequence, for $k \geqslant \ell$,

$$\begin{aligned}|I_k-I_\ell| &= \Big|\int_{A\cap D(0,k)} f_k(x)\,dx-\int_{A\cap D(0,\ell)} f_\ell(x)\,dx\Big|\\ &= \Big|\int_{A\cap D(0,k)\setminus D(0,\ell)} f_k(x)\,dx-\int_{A\cap D(0,\ell)}\big(f_k(x)-f_\ell(x)\big)dx\Big|\\ &\leqslant \int_{A\cap D(0,k)\setminus D(0,\ell)} g(x)\,dx+2\int_{A\cap D(0,\ell)}\big(g(x)-(g\wedge\ell)(x)\big)dx\\ &= \int_{A\cap D(0,k)} g(x)\,dx-\int_{A\cap D(0,\ell)} g(x)\,dx+2\int_{A\cap D(0,\ell)}\big(g(x)-(g\wedge\ell)(x)\big)dx\,.\end{aligned}$$

Since g is integrable, identity (A.32) then implies that $\displaystyle\lim_{\substack{k\geqslant\ell\\ \ell\to\infty}}|I_k-I_\ell| = 0$; thus $\{I_k\}_{k=1}^\infty$ is Cauchy. The absolute convergence of $\displaystyle\int_A f(x)\,dx$ is trivial since $f^\pm \leqslant g$. $\square$

Example A.166. Let $f:[0,\infty)\to\mathbb{R}$ be given by $f(x)=\dfrac{\sin x}{x^2+1}$. Then $\big|f(x)\big|\leqslant\dfrac{1}{x^2+1}$ and the function $y=\dfrac{1}{x^2+1}$ is integrable over $[0,\infty)$ since

$$\lim_{R\to\infty}\int_0^R\frac{1}{x^2+1}dx=\lim_{R\to\infty}\tan^{-1}x\Big|_{x=0}^{x=R}=\lim_{R\to\infty}\tan^{-1}R=\frac{\pi}{2}\,.$$

Even though the limit of the integral $\displaystyle\int_{A\cap D(0,k)}\big((-k)\vee(f\wedge k)(x)\big)dx$ might not exists; however, there are special cases that the limit of the integral above exists (including the possibility that it diverges to infinity). For example, if the limit $\displaystyle\lim_{k\to\infty}\int_{A\cap D(0,k)}(f^+\wedge k)(x)\,dx$ or $\displaystyle\lim_{k\to\infty}\int_{A\cap D(0,k)}(f^-\wedge k)(x)\,dx$ is finite, then

$$\begin{aligned}\int_{A\cap D(0,k)}\big((-k)\vee(f\wedge k)(x)\big)dx&=\lim_{k\to\infty}\int_{A\cap D(0,k)}\big((f^+\wedge k)(x)-(f^-\wedge k)(x)\big)dx\\&=\lim_{k\to\infty}\int_{A\cap D(0,k)}(f^+\wedge k)(x)\,dx-\lim_{k\to\infty}\int_{A\cap D(0,k)}(f^-\wedge k)(x)\,dx\,.\end{aligned}$$

This observation motivates the following

Definition A.167. Let $A\subseteq\mathbb{R}^n$ be a set whose boundary ∂A has measure zero, and $f:A\to\mathbb{R}$ be an integrable function. The improper integral $\displaystyle\int_A f(x)\,dx$ is said to be ***absolutely convergent*** if

$$\int_A f^+(x)\,dx<\infty\quad\text{and}\quad\int_A f^-(x)\,dx<\infty\,,$$

and in this case $\displaystyle\int_A f(x)\,dx=\int_A f^+(x)\,dx-\int_A f^-(x)\,dx$, and f is said to be ***absolutely integrable*** over A.

The improper integral $\displaystyle\int_A f(x)\,dx$ is said to be ***conditionally convergent*** if

$$\int_A f^+(x)\,dx=\int_A f^-(x)\,dx=\infty\,.$$

Example A.168. Let $f:[0,\infty)\to\mathbb{R}$ be given by $f(x)=\dfrac{\sin x}{x}$. Then for all $0<r<R$, using the integration by parts formula we obtain that

$$\Big|\int_r^R\frac{\sin x}{x}dx\Big|=\Big|\frac{1-\cos x}{x}\Big|_{x=r}^{x=R}+\int_r^R\frac{1-\cos x}{x^2}dx\Big|\leqslant\frac{2}{R}+\frac{2}{r}+\int_r^R\frac{2}{x^2}dx=\frac{4}{r}\,.$$

Let $I_k = \displaystyle\int_0^k \frac{\sin x}{x} dx$. Then the inequality above implies that $\{I_k\}_{k=1}^\infty$ is Cauchy in $\mathbb{R}$, so the limit $\displaystyle\int_0^\infty \frac{\sin x}{x} dx = \lim_{k\to\infty} I_k$ exists. However,

$$\int_0^\infty f^+(x)\,dx = \sum_{k=1}^\infty \int_{(2k-2)\pi}^{(2k-1)\pi} \frac{\sin x}{x} dx \geqslant \sum_{k=1}^\infty \frac{1}{(2k-1)\pi} \int_{(2k-2)\pi}^{(2k-1)\pi} \sin x dx = \frac{2}{\pi} \sum_{k=1}^\infty \frac{1}{2k-1} = \infty$$

and

$$\int_0^\infty f^-(x)\,dx = \sum_{k=1}^\infty \int_{(2k-1)\pi}^{2k\pi} \frac{-\sin x}{x} dx \geqslant \sum_{k=1}^\infty \frac{1}{2k\pi} \int_{(2k-1)\pi}^{2k\pi} (-\sin x) dx = \frac{1}{\pi} \sum_{k=1}^\infty \frac{1}{k} = \infty\,.$$

Therefore, the improper integral $\displaystyle\int_0^\infty \frac{\sin x}{x}\,dx$ is conditionally convergent.

REMARK A.169. For absolutely integrable function $f : A \to \mathbb{R}$, one can compute the integral of f over A by

$$\begin{aligned}
\int_A f(x)\,dx &= \lim_{k\to\infty} \Big(\int_{A\cap D(0,k)} \big((f^+ \wedge k)(x) - (f^- \wedge k)(x)\big)\,dx \Big) \\
&= \lim_{k\to\infty} \int_{A\cap D(0,k)} \big((f^+ \wedge k)(x)\,dx - \lim_{k\to\infty} \int_{A\cap D(0,k)} (f^- \wedge k)(x)\big)\,dx \\
&= \lim_{k\to\infty} \int_{A\cap D(0,k)} f^+(x)\,dx - \lim_{k\to\infty} \int_{A\cap D(0,k)} f^-(x)\,dx \\
&= \lim_{k\to\infty} \int_{A\cap D(0,k)} \big(f^+(x) - f^-(x)\big)\,dx = \lim_{k\to\infty} \int_{A\cap D(0,k)} f(x)\,dx\,. \qquad \text{(A.41)}
\end{aligned}$$

where (A.32) is used to conclude the third equality. In (A.41), the set $D(0,k)$ can also be replaced by increasing sequence of set $\{B_k\}_{k=1}^\infty$ as introduced in Theorem A.162.

A.6.2 The Monotone Convergence Theorem and the Dominated Convergence Theorem

In the remaining part of this section, we present some important theorems introduced in Section A.2.4 and A.2.5 under the new settings of improper integrals. Since the improper integrals are defined as the limit of Riemann integrals, we first need to prove those convergence theorems, including the Monotone Convergence Theorem and the Dominated Convergence Theorem (the improper integral version of Bounded Convergence Theorem) for improper integrals, and then apply the Fubini theorem and the change of variables formula for Riemann integrals to conclude the counter-parts for improper integrals by passing to the limit. To summarize briefly, the Fubini theorem and the change of variables formula can be proved for absolutely integrable functions, and the absolute integrability of a function can often be seen using the Tonelli Theorem which will also be presented.

Noting that since we are concerned with the improper integrals, the sequence of functions under consideration in general can be unbounded, so the assumption that the sequence of functions is uniformly bounded by a constant in the Bounded Convergence Theorem in general is pointless.

Theorem A.170 (Dominated Convergence Theorem)**.** *Let $A \subseteq \mathbb{R}^n$ be a set whose boundary ∂A has measure zero, $f : A \to \mathbb{R}$ be a function such that the collection of points of discontinuity of f has measure zero, and $f_n : A \to \mathbb{R}$ be non-negative integrable for all $n \in \mathbb{N}$ such that $\{f_n\}_{n=1}^{\infty}$ converges pointwise to f. Suppose that there exists an integrable function g such that $|f_n| \leqslant g$ for all $n \in \mathbb{N}$. Then f is integrable, and*

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

Proof. Since $\big|f_n(x)\big| \leqslant g(x)$ for all $x \in A$, $\big|f(x)\big| \leqslant g(x)$ for all $x \in A$. By the integrability of g, the comparison test (Theorem A.165) implies that f is also integrable.

Let $\varepsilon > 0$ be given. Since f, g are integrable, there exists $K > 0$ such that

$$0 \leqslant \int_A g(x)\,dx - \int_{A\cap D(0,k)} (g \wedge k)(x)\,dx < \frac{\varepsilon}{3} \qquad \forall\, k \geqslant K \tag{A.42}$$

and

$$\Big|\int_A f(x)\,dx - \int_{A\cap D(0,k)} \big((-k) \vee (f \wedge k)\big)(x)\,dx\Big| < \frac{\varepsilon}{3} \qquad \forall\, k \geqslant K\,.$$

Moreover, since $(-k) \vee (f_n \wedge k) \to (-k) \vee (f \wedge k)$ p.w. as $n \to \infty$ (due to the pointwise convergence of $\{f_n\}_{n=1}^{\infty}$ to f), and $\big|(-k) \vee (f_n \wedge k)\big| \leqslant k$ on $A \cap D(0,k)$, the Bounded Convergence Theorem (Theorem A.150) implies that for each $k \in \mathbb{N}$ there exists $N(k) > 0$ such that

$$\Big|\int_{A\cap D(0,k)} \big((-k)\vee(f_n\wedge k)\big)(x)\,dx - \int_{A\cap D(0,k)} \big((-k)\vee(f\wedge k)\big)(x)\,dx\Big| < \frac{\varepsilon}{3} \qquad \forall\, n \geqslant N(k)\,.$$

Note that by Theorem A.155, (A.42) also suggests that for $k \geqslant K$,

$$\begin{aligned}\int_{A\cap D(0,k)^\complement} g(x)\,dx &= \int_A g(x)\,dx - \int_{A\cap D(0,k)} g(x)\,dx \\ &< \frac{\varepsilon}{3} + \int_{A\cap D(0,k)} (g\wedge k)(x)\,dx - \int_{A\cap D(0,k)} g(x)\,dx \leqslant \frac{\varepsilon}{3}\,.\end{aligned}$$

We also note that by Theorem A.155 again,

$$\begin{aligned}\Big|\int_A f(x)\,dx - \int_A f_n(x)\,dx\Big| &\leqslant \Big|\int_A f(x)\,dx - \int_{A\cap D(0,K)} \big((-K)\vee(f\wedge K)\big)(x)\,dx\Big| \\ &+ \Big|\int_{A\cap D(0,K)} \big((-K)\vee(f\wedge K)\big)(x)\,dx - \int_{A\cap D(0,K)} \big((-K)\vee(f_n\wedge K)\big)(x)\,dx\Big| \\ &+ \Big|\int_{A\cap D(0,K)} \big((-K)\vee(f_n\wedge K)\big)(x)\,dx - \int_{A\cap D(0,K)} f_n(x)\,dx\Big| + \Big|\int_{A\cap D(0,K)^\complement} f_n(x)\,dx\Big|,\end{aligned}$$

and by the fact that $|f_n| \leqslant g$, (A.40) implies that

$$\big|\big((-K)\vee(f_n\wedge K)\big)(x) - f_n(x)\big| \leqslant g(x) - (g\wedge K)(x)\,.$$

Therefore, for $n \geqslant N(K)$,

$$\begin{aligned}&\Big|\int_A f(x)\,dx - \int_A f_n(x)\,dx\Big| \\ &\qquad< \frac{2\varepsilon}{3} + \int_{A\cap D(0,K)} \big(g(x) - (g\wedge K)(x)\big)\,dx + \int_{A\cap D(0,K)^\complement} g(x)\,dx \\ &\qquad\leqslant \frac{2\varepsilon}{3} + \int_A g(x)\,dx - \int_{A\cap D(0,k)} (g\wedge K)(x)\,dx < \varepsilon\,.\end{aligned}$$

□

The Monotone Convergence Theorem for improper integrals, unlike the case in the Riemann integrals, is no longer an immediate consequence of the Dominated Convergence Theorem since the "integral" of the limit function might be infinite. It requires a little bit more attention to get proved.

Theorem A.171 (Monotone Convergence Theorem for Improper Integrals)**.** *Let $A \subseteq \mathbb{R}^n$ be a set whose boundary ∂A has measure zero, $f : A \to \mathbb{R}$ be a function such that the collection of points of discontinuity of f has measure zero, and $f_n : A \to \mathbb{R}$ be non-negative integrable for all $n \in \mathbb{N}$ such that $\{f_n\}_{n=1}^\infty$ converges pointwise to f. Suppose that $\{f_n\}_{n=1}^\infty$ is a monotone sequence of functions; that is, $f_n \geqslant f_{n+1}$ or $f_n \leqslant f_{n+1}$ for all $n \in \mathbb{N}$. Then*

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

Proof. W.L.O.G. we assume that $0 \leqslant f_n \leqslant f_{n+1}$ for all $n \in \mathbb{N}$ for in another case we let $g_n = f_1 - f_n$ which constitutes an increasing non-negative integrable sequence of functions with limit $f_1 - f$. Moreover, by the Dominated Convergence Theorem (Theorem A.170), we only need to consider the case that $\displaystyle\int_A f(x)\,dx = \infty$ and show that $\displaystyle\lim_{n\to\infty}\int_A f_n(x)\,dx = \infty$. We also assume the non-trivial case that $\displaystyle\int_A f_n(x)\,dx < \infty$ for all $n \in \mathbb{N}$.

Let $M > 0$ be given. Since $\displaystyle\int_A f(x)\,dx = \infty$, there exists $K > 0$ such that

$$\int_{A\cap D(0,k)} (f\wedge k)(x)\,dx \geqslant 2M \qquad \forall\, k \geqslant K\,.$$

By the Monotone Convergence Theorem for Riemann integrals (Theorem A.149), for each $k \in \mathbb{N}$ there exists $N(k) > 0$ such that

$$-M \leqslant \int_{A\cap D(0,k)} (f_n\wedge k)(x)\,dx - \int_{A\cap D(0,k)} (f\wedge k)(x)\,dx \leqslant 0 \qquad \forall\, n \geqslant N(k)\,.$$

Therefore, for all $k \geqslant K$ and $n \geqslant N(K)$,

$$\begin{aligned}\int_A f_n(x)\,dx &= \int_A f_n(x)\,dx - \int_{A\cap D(0,K)} (f_n\wedge K)(x)\,dx + \int_{A\cap D(0,K)} (f_n\wedge K)(x)\,dx \\ &\quad - \int_{A\cap D(0,K)} (f\wedge K)(x)\,dx + \int_{A\cap D(0,K)} (f\wedge K)(x)\,dx \\ &\geqslant \int_A f_n(x)\,dx - \int_{A\cap D(0,k)} (f_n\wedge k)(x)\,dx + M\,.\end{aligned}$$

Passing k to the limit, we find that $\displaystyle\int_A f_n(x)\,dx \geqslant M$ for all $n \geqslant N(K)$. □

A.6.3 The Fubini theorem and the Tonelli theorem

Now we present the Fubini theorem. In the case of improper integrals, for the Fubini Theorem to hold it requires absolute integrability of functions.

Theorem A.172 (Fubini). *Let $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$ be sets whose boundary ∂A and ∂B have measure zero in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and $f : A\times B \to \mathbb{R}$ be a function such that $f(x,\cdot)$ is integrable over B for all $x \in A$ and $f(\cdot,y)$ is integrable over A for all $y \in B$. If f is absolutely integrable over $A\times B$, then*

$$\int_{A\times B} f(x,y)d(x,y) = \int_A \Big(\int_B f(x,y)dy\Big)dx = \int_B \Big(\int_A f(x,y)dx\Big)dy\,.$$

Proof. We only prove the first equality since the proof for the other one is similar.

First we note that if $f(x,\cdot)$ is integrable over B for all $x \in A$, so are $f^+(x,\cdot)$ and $f^-(x,\cdot)$. Since $\displaystyle\int_{B\cap[-k,k]^m}(f^+\wedge k)(x,y)dy$ and $\displaystyle\int_{B\cap[-k,k]^m}(f^-\wedge k)(x,y)dy$ are both increasing in k and bounded from above, by the Monotone Convergence Theorem (Theorem A.171),

$$\int_{B\cap[-k,k]^m} f^+(x,y)dy \nearrow \int_B f^+(x,y)dy \qquad \text{as} \quad k\to\infty$$

and

$$\int_{B\cap[-k,k]^m} f^-(x,y)dy \nearrow \int_B f^-(x,y)dy \qquad \text{as} \quad k\to\infty\,.$$

By the Monotone Convergence Theorem again, we find that

$$\lim_{k\to\infty}\int_{A\cap[-k,k]^n}\Big(\int_{B\cap[-k,k]^m}(f^+\wedge k)(x,y)dy\Big)dy = \int_A\Big(\int_B f^+(x,y)dy\Big)dx < \infty$$

and

$$\lim_{k\to\infty}\int_{A\cap[-k,k]^n}\Big(\int_{B\cap[-k,k]^m}(f^-\wedge k)(x,y)dy\Big)dy = \int_A\Big(\int_B f^-(x,y)dy\Big)dx < \infty\,.$$

By Theorem A.162 and the Fubini theorem (Theorem A.99),

$$\begin{aligned}
&\int_{A\times B} f(x,y)d(x,y)\\
&\quad= \lim_{k\to\infty}\Big[\int_{(A\times B)\cap[-k,k]^{n+m}}\big[(f^+\wedge k)(x,y)-(f^-\wedge k)(x,y)\big]d(x,y)\Big]\\
&\quad= \lim_{k\to\infty}\int_{A\cap[-k,k]^n}\Big(\int_{B\cap[-k,k]^n}\big[(f^+\wedge k)(x,y)-(f^-\wedge k)(x,y)\big]dy\Big)dx\\
&\qquad= \int_A\Big(\int_B f^+(x,y)dy\Big)dx - \int_A\Big(\int_B f^-(x,y)dy\Big)dx = \int_A\Big(\int_B f(x,y)dy\Big)dx\,.
\end{aligned}$$

□

Evaluating integrals by integrating iteratively, which is the main idea of the Fubini theorem, is an important tool to compute integrals. Therefore, how to check the absolute integrability of functions so that the Fubini theorem can be applied is an important subject. So far the only test we have for determining the absolute integrability of functions is the comparison test (Theorem A.165); however, for a given function f sometimes it is not easy to find an absolutely integrable upper bound. Nevertheless, we can compute the integral of $|f|$ by computing the iterated integrals of $|f|$ directly (without knowing if we can do this first), thanks to the Tonelli theorem.

Theorem A.173 (Tonelli)**.** *Let $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$ be sets such that boundary $\partial(A \times B)$, ∂A and ∂B have measure zero in $\mathbb{R}^{n+m}$, $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and $f : A \times B \to \mathbb{R}$ be a non-negative function satisfying*

1. *the collection of points of discontinuity of f has measure zero in $\mathbb{R}^{n+m}$;*
2. *the collection of points of discontinuity of $f(x, \cdot)$ has measure zero in $\mathbb{R}^m$ for all $x \in A$;*
3. *the collection of points of discontinuity of $\int_B f(\cdot, y)\,dy$ has measure zero in $\mathbb{R}^n$.*

Then

$$\int_{A\times B} f(x,y)d(x,y) = \int_A \Big(\int_B f(x,y)dy \Big) dx\,. \tag{A.43}$$

Remark A.174. If $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$ are rectangular sets (not necessarily bounded), then $\partial(A \times B)$, ∂A and ∂B have measure zero in $\mathbb{R}^{n+m}$, $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. In general, the requirement for $\partial(A \times B)$, ∂A and ∂B have measure zero (in corresponding spaces) is used to guaranteed that the truncation of f and $f(x, \cdot)$ and $\int_B f(\cdot, y)\,dy$ are Riemann integrable over truncated sets $(A \times B) \cap [-k, k]^{n+m}$, $B \cap [-k, k]^m$ and $A \cap [-k, k]^n$.

We also note that even if $f : A \times B \to \mathbb{R}$ is continuous, $F(x) \equiv \int_B f(x, y)\,dy$ might still be discontinuous at some points. For example, let $A = [-1, 1]$, $B = [1, \infty)$ and $f(x, y) = |x|y^{-1-|x|}$. Then

$$F(x) = \begin{cases} 0 & \text{if } x = 0\,, \\ 1 & \text{otherwise}\,, \end{cases}$$

which is discontinuous at $x = 0$. In general, we do not know if the collection of points of discontinuity of F has measure zero in $\mathbb{R}^n$ even if f is continuous on $A \times B$.

Proof of Theorem A.173. We first prove that for each $\ell \in \mathbb{N}$,

$$\int_{(A\cap[-\ell,\ell]^n)\times B} (f \wedge \ell)(x,y)d(x,y) = \int_{A\cap[-\ell,\ell]^n} \Big(\int_B (f \wedge \ell)(x,y)dy \Big) dx\,. \tag{A.44}$$

Once the identity above is proved, then by the fact that $\big\{(f \wedge \ell)1_{[-\ell,\ell]^n\times B}\big\}_{\ell=1}^{\infty}$ converges to f pointwise and the convergence is monotone, the Monotone Convergence Theorem (Theorem A.171) implies (A.43) immediately.

Let $\ell \in \mathbb{N}$ be fixed. For each $k \geqslant \ell$, define $g_k(x) = \int_{B\cap[-k,k]^m} (f \wedge \ell)(x,y)\,dy$. Then $g_k \leqslant g_{k+1}$ for all $k \to \mathbb{N}$, and the Monotone Convergence Theorem (Theorem A.171) suggests that

$$\lim_{k\to\infty} g_k(x) = \int_B (f\wedge \ell)(x,y)dy \qquad \forall\, x \in A\,.$$

In other words, $\{g_k\}_{k=1}^\infty$ converges pointwise to $\int_B (f\wedge\ell)(\cdot,y)dy$. Therefore, we apply the Monotone Convergence Theorem again to conclude that

$$\lim_{k\to\infty} \int_{A\cap[-\ell,\ell]^n} g_k(x)\,dx = \int_{A\cap[-\ell,\ell]^n} \Big(\int_B (f\wedge\ell)(x,y)\,dy\Big)dx\,.$$

On the other hand, the Fubini theorem (Theorem A.99) implies that

$$\int_{A\cap[-\ell,\ell]^n} g_k(x)\,dx = \int_{(A\cap[-\ell,\ell]^n)\times(B\cap[-k,k]^m)} (f\wedge\ell)(x,y)d(x,y)\,;$$

thus the Monotone Convergence Theorem (Theorem A.171) implies that

$$\lim_{k\to\infty} \int_{A\cap[-\ell,\ell]^n} g_k(x)\,dx = \int_{(A\cap[-\ell,\ell]^n)\times B} (f\wedge\ell)(x,y)d(x,y)$$

which concludes (A.44). □

Corollary A.175. *Let $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$ be sets such that boundary $\partial(A\times B)$, ∂A and ∂B have measure zero in $\mathbb{R}^{n+m}$, $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, $f : A \to \mathbb{R}$ and $g : B \to \mathbb{R}$ be absolutely integrable. Then the function $h : A\times B \to \mathbb{R}$ given by $h(x,y) = f(x)g(y)$ is absolutely integrable, and*

$$\int_{A\times B} h(x,y)d(x,y) = \Big(\int_A f(x)\,dx\Big)\Big(\int_B g(y)\,dy\Big)\,.$$

Proof. By Theorem A.73, the collection of points of discontinuity of $|h|$ has measure zero in $\mathbb{R}^{n+m}$. Moreover, by the integrability of g we find that the collection of points of discontinuity of $\big|h(x,\cdot)\big|$ has measure zero in $\mathbb{R}^m$ for each $x \in A$. Since $|f|$ is integrable over A and

$$\int_B \big|h(x,y)\big|\,dy = \big|f(x)\big| \int_B \big|g(y)\big|\,dy\,,$$

the collection of points of discontinuity of $\int_B |h(\cdot,y)|dy$ has measure zero in $\mathbb{R}^n$. In other words, h satisfies condition 1-3 in the Tonelli theorem (Theorem A.173); thus we have

$$\int_{A\times B}\big|h(x,y)\big|d(x,y)=\int_A\Big(\int_B\big|h(x,y)\big|dy\Big)dx=\Big(\int_A\big|f(x)\big|\,dx\Big)\Big(\int_B\big|g(y)\big|\,dy\Big)<\infty\,.$$

Therefore, h is absolutely integrable over $A\times B$. Since $h(x,\cdot)$ is integrable over B for all $x\in A$ and $h(\cdot,y)$ is integrable over A for all $y\in B$, the Fubini theorem (Theorem A.172) further suggests that

$$\int_{A\times B}h(x,y)d(x,y)=\int_A\Big(\int_B h(x,y)dy\Big)dx=\Big(\int_A f(x)\,dx\Big)\Big(\int_B g(y)\,dy\Big)\,. \qquad \square$$

A.6.4 Change of variables formula

As in the proof of the Fubini theorem (Theorem A.172) and the Tonelli theorem (Theorem A.173), we can also apply the Monotone Convergence Theorem (Theorem A.171) to conclude the change of variables formula for improper integrals.

Theorem A.176 (Change of Variables Formula)**.** *Let $\mathcal{U}\subseteq\mathbb{R}^n$ be an open set whose boundary $\partial\mathcal{U}$ has measure zero, and $g:\mathcal{U}\to\mathbb{R}^n$ be an one-to-one $\mathscr{C}^1$ mapping with $\mathscr{C}^1$ inverse; that is, $g^{-1}:g(\mathcal{U})\to\mathcal{U}$ is also continuously differentiable. Assume that the Jacobian of g, $\mathbf{J}_g=\det([Dg])$, does not vanish in $\mathcal{U}$. If $f:g(\mathcal{U})\to\mathbb{R}$ is absolutely integrable, then $(f\circ g)\mathbf{J}_g$ is absolutely integrable over $\mathcal{U}$, and*

$$\int_{g(\mathcal{U})}f(y)\,dy=\int_{\mathcal{U}}(f\circ g)(x)\big|\mathbf{J}_g(x)\big|\,dx=\int_{\mathcal{U}}(f\circ g)(x)\Big|\frac{\partial(g_1,\cdots,g_n)}{\partial(x_1,\cdots,x_n)}\Big|\,dx\,.$$

Proof. Let $\{\mathcal{U}_k\}_{k=1}^{\infty}$ be a sequence of open sets such that $\bigcup\limits_{k=1}^{\infty}\mathcal{U}_k=\mathcal{U}$, and for each $k\in\mathbb{N}$, $\partial\mathcal{U}_k$ has measure zero, $\mathcal{U}_k\subset\subset\mathcal{U}$, and $\mathcal{U}_k\subseteq\mathcal{U}_{k+1}$. We note that such sequence of sets always exists if $\mathcal{U}$ is open. Define $f_k^+=f^+\wedge k$ and $f_k^-=f^-\wedge k$. Then the change of variables formula (Theorem A.102) implies that $g(\mathcal{U}_k)$ has volume, and

$$\int_{\mathcal{U}_k}(f_k^\pm\circ g)(x)\big|\mathbf{J}_g(x)\big|\,dx=\int_{g(\mathcal{U}_k)}f_k^\pm(y)\,dy\,.$$

Passing k to the limit, by the Monotone Convergence Theorem (Theorem A.171) and the absolute integrability of f we find that

$$\int_{\mathcal{U}}(f^\pm\circ g)(x)\big|\mathbf{J}_g(x)\big|\,dx=\int_{g(U)}f^\pm(y)\,dy<\infty\,;$$

thus

$$\int_{g(U)} f(y)\,dy = \int_{g(\mathcal{U})} \big(f^+(y) - f^-(y)\big)\,dy = \int_{\mathcal{U}} (f\circ g)(x)\big|\mathbf{J}_g(x)\big|\,dx\,. \qquad \square$$

Remark A.177. In Theorem A.176, except that the integrals under consideration could be improper integrals, there is no need to have a larger open set $\mathcal{V}$ so that $\bar{\mathcal{U}} \subseteq \mathcal{V}$ which is required in the proof of Theorem A.102. We also note that the change of variables formula is valid for non-negative functions whose point of discontinuity forms a measure zero set.

A.7 The Divergence and Stokes Theorem

A.7.1 The metric tensor and the first fundamental form

Definition A.178 (Metric)**.** Let $\Sigma \subseteq \mathbb{R}^{\rm n}$ be a $(n-1)$-dimensional manifold. The metric tensor associated with the local parametrization $\{\mathcal{V}, \psi\}$ (at $p \in \Sigma$) is the matrix $g = [g_{\alpha\beta}]_{({\rm n}-1)\times(n-1)}$ given by

$$g_{\alpha\beta} = \psi,_\alpha \cdot \psi,_\beta = \sum_{i=1}^{\rm n} \frac{\partial \psi^i}{\partial y_\alpha}\frac{\partial \psi^i}{\partial y_\beta} \qquad \text{in} \quad \mathcal{V}\,.$$

Proposition A.179. *Let $\Sigma \subseteq \mathbb{R}^{\rm n}$ be a $(n-1)$-dimensional manifold, and $g = [g_{\alpha\beta}]_{({\rm n}-1)\times({\rm n}-1)}$ be the metric tensor associated with the local parametrization $\{\mathcal{V}, \psi\}$ (at $p \in \Sigma$). Then the metric tensor g is positive definite; that is,*

$$\sum_{\alpha,\beta=1}^{{\rm n}-1} g_{\alpha\beta} v^\alpha v^\beta > 0 \qquad \forall\, \boldsymbol{v} = \sum_{\gamma=1}^{{\rm n}-1} v^\gamma \frac{\partial \psi}{\partial y^\gamma} \neq \mathbf{0}\,.$$

Proof. Since $D\psi$ has full rank on $\mathcal{V}$, every tangent vector $\boldsymbol{v}$ can be expressed as the linear combination of $\Big\{\dfrac{\partial \psi}{\partial y_1}, \cdots, \dfrac{\partial \psi}{\partial y_{{\rm n}-1}}\Big\}$. Write $\boldsymbol{v} = \sum\limits_{\gamma=1}^{{\rm n}-1} v^\gamma \dfrac{\partial \psi}{\partial y^\gamma}$. Then if $\boldsymbol{v} \neq \mathbf{0}$,

$$0 < \|\boldsymbol{v}\|^2_{\mathbb{R}^{\rm n}} = \sum_{i=1}^{\rm n}\sum_{\alpha,\beta=1}^{{\rm n}-1} v^\alpha \frac{\partial \psi^i}{\partial y_\alpha} v^\beta \frac{\partial \psi^i}{\partial \psi_\beta} = \sum_{\alpha,\beta=1}^{\rm n} g_{\alpha\beta} v^\alpha v^\beta\,. \qquad \square$$

Definition A.180 (The first fundamental form)**.** Let $\Sigma \subseteq \mathbb{R}^{\rm n}$ be a $(n-1)$-dimensional manifold, and $g = [g_{\alpha\beta}]_{({\rm n}-1)\times({\rm n}-1)}$ be the metric tensor associated with the local parametrization $\{\mathcal{V}, \psi\}$ (at $p \in \Sigma$). The first fundamental form associated with the local parametrization $\{\mathcal{V}, \psi\}$ (at $p \in \Sigma$) is the scalar function ${\rm g} = \det(g)$.

The surface integral

Let $\Sigma \subseteq \mathbb{R}^{\mathrm{n}}$ be a $(n-1)$-dimensional manifold, and $\{\mathcal{V}, \psi\}$ be a global parametrization of Σ; that is, $\Sigma = \psi(\mathcal{V})$. If $f : \Sigma \to \mathbb{R}$ is a bounded continuous function, the surface integral of f over Σ, denoted by $\displaystyle\int_\Sigma f\, dS$, is defined by

$$\int_\Sigma f\, dS = \int_{\mathcal{V}} (f \circ \psi)\sqrt{\mathrm{g}}\, dx' \,, \tag{A.45}$$

where the integral on the right-hand side is the Lebesgue integral on a subset $\mathcal{V}$ of $\mathbb{R}^{\mathrm{n}-1}$ (thus dx' is the $(n-1)$-dimensional Lebesgue measure). In particular, if $f \equiv 1$, the number $\displaystyle\int_\Sigma dS \equiv \int_\Sigma 1\, dS$ is the surface area of Σ.

Since the surface integrals defined by (A.45) seems to depend on a given parametrization, before proceeding we show that the surface integral is indeed independent of the choice of the parametrizations. Suppose that $\{\mathcal{V}_1, \psi_1\}$ and $\{\mathcal{V}_2, \psi_2\}$ are two global $\mathscr{C}^1$-parametrizations of Σ at p, g_1, g_2 denote the metric tensors associated with the parametrizations $\{\mathcal{V}_1, \psi_1\}$, $\{\mathcal{V}_2, \psi_2\}$, respectively, and $\mathrm{g}_1 = \det(g_1)$, $\mathrm{g}_2 = \det(g_2)$ are corresponding first fundamental forms. Let $\Psi = \psi_2^{-1} \circ \psi_1$. Then the change of variables formula implies that

$$\int_{\mathcal{V}_2} (f \circ \psi_2)\sqrt{\mathrm{g}_2}\, dx' = \int_{\mathcal{V}_1} (f \circ \psi_2 \circ \Psi)\big(\sqrt{\mathrm{g}_2} \circ \Psi\big)|J_\Psi|\, dx' = \int_{\mathcal{V}_1} (f \circ \psi_1)\big(\sqrt{\mathrm{g}_2} \circ \Psi\big)|J_\Psi|\, dx',$$

where J_Ψ is the Jacobian of the map Ψ. By the chain rule, we find that

$$\big[D\Psi\big]^{\mathrm{T}}\big[(D\psi_2) \circ \Psi\big]^{\mathrm{T}}\big[(D\psi_2) \circ \Psi\big]\big[D\Psi\big] = \big[D\psi_1\big]^{\mathrm{T}}\big[D\psi_1\big]\,;$$

thus by the fact that $\mathrm{g}_1 = \det\big(\big[D\psi_1\big]^{\mathrm{T}}\big[D\psi_1\big]\big)$ and $\mathrm{g}_2 = \det\big(\big[D\psi_2\big]^{\mathrm{T}}\big[D\psi_2\big]\big)$, we obtain that

$$\det\big(\big[D\Psi\big]\big)^2(\mathrm{g}_2 \circ \Psi) = \mathrm{g}_1\,.$$

Since $J_\Psi = \det\big(\big[D\Psi\big]\big)$, the identity above implies that $|J_\Psi|(\sqrt{\mathrm{g}_2} \circ \Psi) = \sqrt{\mathrm{g}_1}$, so we conclude that

$$\int_{\mathcal{V}_1} (f \circ \psi_1)\sqrt{\mathrm{g}_1}\, dx' = \int_{\mathcal{V}_2} (f \circ \psi_2)\sqrt{\mathrm{g}_2}\, dx'\,. \tag{A.46}$$

Therefore, the surface integral of f over Σ is independent of the choice of parametrizations of Σ. In particular, the surface area of a regular $\mathscr{C}^1$-surface which can be parametrized by a global parametrization is also independent of the choice of parametrizations.

Next, we study the surface integral over general $(n-1)$-dimensional manifold that cannot be parametrized using a single pair $\{\mathcal{V}, \psi\}$. Let $\Sigma \subseteq \mathbb{R}^n$ be a $(n-1)$-dimensional manifold, and $\{\mathcal{V}_i, \psi_i\}_{i\in\mathcal{I}}$ be a collection of local parametrizations satisfying that for each $p \in \Sigma$ there exists $i \in \mathcal{I}$ such that $\{\mathcal{V}_i, \psi_i\}$ is a local parametrization of Σ at p. Since each $\psi_i(\mathcal{V}_i) \subseteq \Sigma$ is open (relative to Σ), there exists open set $\mathcal{U}_i \subseteq \mathbb{R}^n$ such that $\psi_i(\mathcal{V}_i) = \mathcal{U}_i \cap \Sigma$. By Proposition 5.8, there exists a partition of unity $\{\zeta_i\}_{i\in\mathcal{I}}$ of Σ subordinate to $\{\mathcal{U}_i\}_{i\in\mathcal{I}}$. Intuitively we can define the surface integral of f over Σ as follows:

$$\int_\Sigma f\, dS = \sum_{i\in\mathcal{I}} \int_\Sigma (\zeta_i f)\, dS = \sum_{i\in\mathcal{I}} \int_{\mathcal{U}_i\cap\Sigma} \zeta_i f\, dS = \int_{\mathcal{V}_i} (\zeta_j f)\circ\psi_i \sqrt{\mathrm{g}_i}\, dS\,. \tag{A.47}$$

The surface integral defined above is independent of the choice of the partition of unity.

A.7.2 Some useful identities

In this sub-section, we temporarily switch to a more general setting that the "surface" (or more precisely, manifold) under consideration is the boundary of an open set of $\mathbb{R}^n$.

Let $\Sigma \subseteq \mathbb{R}^n$ be the boundary of an open set Ω (thus an oriented surface), $\{\mathcal{V}, \psi\}$ be a local parametrization of Σ, and $\mathbf{N} : \Sigma \to \mathbb{R}^n$ be the normal vector on Σ which is compatible with the parametrization ψ; that is,

$$\det\left(\left[\,\psi_{,1} \,\vdots\, \psi_{,2} \,\vdots\, \cdots \,\vdots\, \psi_{,\mathrm{n}-1} \,\vdots\, \mathbf{N}\circ\psi\,\right]\right) > 0\,.$$

Define $\Psi(y', y_n) = \psi(y') + y_n(\mathbf{N}\circ\psi)(y')$. Then $\Psi : \mathcal{V}\times(-\varepsilon,\varepsilon) \to \mathcal{T}$ for some tubular neighborhood $\mathcal{T}$ of Σ.

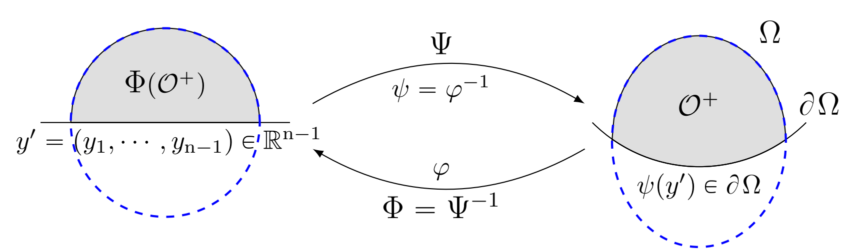


Figure A.4: The map Ψ constructed from the local parametrization $\{\mathcal{V}, \psi\}$

Since $(\nabla\Psi)\big|_{\{y_n=0\}} = \big[\,\psi_{,1} \,\vdots\, \psi_{,2} \,\vdots\, \cdots \,\vdots\, \psi_{,\mathrm{n}-1} \,\vdots\, \mathbf{N}\circ\psi\,\big]$,

$$\det(\nabla\Psi)^2\big|_{\{y_{\mathrm{n}}=0\}} = \Big[\det\big((\nabla\Psi)^{\mathrm{T}}\big)\det(\nabla\Psi)\Big]\Big|_{\{y_{\mathrm{n}}=0\}} = \det\big((\nabla\Psi)^{\mathrm{T}}\nabla\Psi\big)\Big|_{\{y_{\mathrm{n}}=0\}}$$
$$= \det\left(\begin{bmatrix} g_{11} & g_{12} & \cdots & g_{(\mathrm{n}-1)1} & 0 \\ g_{21} & g_{22} & \cdots & g_{(\mathrm{n}-1)2} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ g_{(\mathrm{n}-1)1} & g_{(\mathrm{n}-1)2} & \cdots & g_{(\mathrm{n}-1)(\mathrm{n}-1)} & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}\right) = \mathrm{g}\,.$$

Defining J as the Jacobian of the map Ψ; that is, $\mathrm{J} = \det(\nabla\Psi)$, then the identity above implies that

$$\mathrm{J} = \sqrt{\mathrm{g}} \qquad \text{on} \quad \{y_{\mathrm{n}} = 0\}\,.$$

Moreover, letting A denote the inverse of the Jacobian matrix of Ψ; that is, $\mathrm{A} = (\nabla\Psi)^{-1}$, and letting $\big[g^{\alpha\beta}\big]_{(\mathrm{n}-1)\times(\mathrm{n}-1)}$ be the inverse matrix of $\big[g_{\alpha\beta}\big]_{(\mathrm{n}-1)\times(\mathrm{n}-1)}$, we find that

$$\mathrm{A}\big|_{\{y_n=0\}} = \left[\,\sum_{\alpha=1}^{\mathrm{n}-1} g^{1\alpha}\psi_{,\alpha} \,\vdots\, \cdots\cdots\cdots\cdots \,\vdots\, \sum_{\alpha=1}^{\mathrm{n}-1} g^{(\mathrm{n}-1)\alpha}\psi_{,\alpha} \,\vdots\, \mathbf{N}\circ\psi\,\right]^{\mathrm{T}}.$$

As a consequence,

$$(\mathrm{J}\mathrm{A}^{\mathrm{T}}\mathrm{e}_{\mathrm{n}})\big|_{\{y_{\mathrm{n}}=0\}} = \sqrt{\mathrm{g}}\,(\mathbf{N}\circ\psi)\,. \tag{A.48}$$

Definition A.181 (The divergence operator)**.** Let $\boldsymbol{u} : \Omega \subseteq \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^{\mathrm{n}}$ be a vector field. The divergence of $\boldsymbol{u}$ is a scalar function defined by

$$\operatorname{div}\boldsymbol{u} = \sum_{i=1}^{\mathrm{n}} \frac{\partial \boldsymbol{u}^i}{\partial x_i}\,.$$

Definition A.182. A vector field $\boldsymbol{u} : \Omega \subseteq \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^{\mathrm{n}}$ is called *solenoidal* or *divergence-free* if $\operatorname{div}\boldsymbol{u} = 0$ in Ω.

A.7.3 The divergence theorem

Theorem A.183 (The divergence theorem)**.** *Let $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ be a bounded Lipschitz domain, and $\boldsymbol{v} \in \mathscr{C}^1(\Omega) \cap \mathscr{C}(\bar{\Omega})$. Then*

$$\int_\Omega \operatorname{div}\boldsymbol{v}\,dx = \int_{\partial\Omega} \boldsymbol{v}\cdot\mathbf{N}\,dS\,,$$

where $\mathbf{N}$ is the outward-pointing unit normal of Ω.

Proof. We prove the case that Ω is a bounded open set of class $\mathscr{C}^3$, and the general result can be obtained by approximating the Lipschitz manifold by a sequence of $\mathscr{C}^3$ manifolds.

Let $\{\mathcal{U}_m\}_{m=1}^K$ be an open cover of $\partial\Omega$ such that for each $m \in \{1, \cdots, K\}$ there exists a $\mathscr{C}^3$-parametrization $\psi_m : \mathcal{V}_m \subseteq \mathbb{R}^{\mathrm{n}-1} \to \mathcal{U}_m$ which is compatible with the orientation $\mathbf{N}$; that is,

$$\det\left([\psi_{m,1} \,\vdots\, \cdots \,\vdots\, \psi_{m,\mathrm{n}-1} \,\vdots\, \mathbf{N} \circ \psi_m]\right) > 0 \qquad \text{on} \quad \mathcal{V}_m\,.$$

Define $\vartheta_m(y', y_\mathrm{n}) = \psi_m(y') + y_\mathrm{n}(\mathbf{N} \circ \psi_m)(y')$. Then there exists $\varepsilon_m > 0$ such that $\vartheta_m : \mathcal{V}_m \times (-\varepsilon_m, \varepsilon_m) \to \mathcal{W}_m$ is a $\mathscr{C}^2$-diffeomorphism for some open set in $\mathbb{R}^\mathrm{n}$ such that $\vartheta_m : \mathcal{V}_m \times (-\varepsilon_m, 0) \to \Omega \cap \mathcal{W}_m$ while $\vartheta_m : \mathcal{V}_m \times (0, \varepsilon_m) \to \mathrm{int}(\Omega^\complement) \cap \mathcal{W}_m$.

Choose an open set $\mathcal{W}_0 \subseteq \mathbb{R}^\mathrm{n}$ such that $\overline{\mathcal{W}_0} \subseteq \Omega$ and $\bar{\Omega} \subseteq \bigcup\limits_{m=0}^{K} \mathcal{W}_m$, and define ϑ_0 as the identity map. Let $0 \leqslant \zeta_m \leqslant 1$ in $\mathscr{C}_c^\infty(\mathcal{U}_m)$ denote a partition-of-unity of $\bar{\Omega}$ subordinate to the open covering $\{\mathcal{W}_m\}_{m=0}^K$; that is,

$$\sum_{m=0}^{K} \zeta_m = 1 \quad \text{and} \quad \mathrm{spt}(\zeta_m) \subseteq \mathcal{U}_m \quad \forall\, m\,.$$

Let $\mathrm{J}_m = \det(\nabla\vartheta_m)$, $\mathrm{A}_m = (\nabla\vartheta_m)^{-1}$, and g_m denote the first fundamental form associated with $\{\mathcal{V}_m, \psi_m\}$. Using (A.48), $\sqrt{\mathrm{g}_m}(\mathbf{N} \circ \vartheta_m) = \mathrm{J}_m(\mathrm{A}_m)^\mathrm{T}\mathrm{e}_\mathrm{n}$ on $\mathcal{V}_m \times \{0\}$ for $m \in \{1, \cdots, K\}$. Therefore, making change of variable $x = \vartheta_m(y)$ in each $\mathcal{W}_m$ we find that

$$\begin{aligned}
\int_{\partial\Omega} \boldsymbol{v} \cdot \mathbf{N}\, dS &= \sum_{m=1}^{K} \int_{\partial\Omega \cap \mathcal{W}_m} \zeta_m (\boldsymbol{v} \cdot \mathbf{N})\, dS \\
&= \sum_{m=1}^{K} \sum_{i=1}^{\mathrm{n}} \int_{\mathcal{V}_m \times \{y_\mathrm{n}=0\}} (\zeta_m \circ \vartheta_m)(\boldsymbol{v}^i \circ \vartheta_m)(\mathbf{N}^i \circ \vartheta_m)\sqrt{\mathrm{g}_m}\, dy' \\
&= \sum_{m=1}^{K} \sum_{i=1}^{\mathrm{n}} \int_{\mathcal{V}_m \times \{y_\mathrm{n}=0\}} (\zeta_m \circ \vartheta_m)(\boldsymbol{v}^i \circ \vartheta_m)\mathrm{J}_m(\mathrm{A}_m)_i^\mathrm{n}\, dy' \\
&= \sum_{m=1}^{K} \sum_{i=1}^{\mathrm{n}} \int_{\mathcal{V}_m \times (-\varepsilon_m, 0)} \frac{\partial}{\partial y_\mathrm{n}}\big[(\zeta_m \circ \vartheta_m)\mathrm{J}_m(\mathrm{A}_m)_i^\mathrm{n}(\boldsymbol{v}^i \circ \vartheta_m)\big]\, dy\,.
\end{aligned}$$

On the other hand, for $\alpha \in \{1, \cdots, \mathrm{n}-1\}$ and $i \in \{1, \cdots, \mathrm{n}\}$,

$$\int_{\mathcal{V}_m \times (-\varepsilon_m, 0)} \frac{\partial}{\partial y_\alpha}\big[(\zeta_m \circ \vartheta_m)\mathrm{J}_m(\mathrm{A}_m)_i^\alpha(\boldsymbol{v}^i \circ \vartheta_m)\big]\, dy = 0\,;$$

thus the Piola identity (A.10) implies that

$$\begin{aligned}\int_{\partial\Omega} \boldsymbol{v}\cdot \mathbf{N}\,dS &= \sum_{m=1}^{K}\sum_{i,j=1}^{\mathrm{n}} \int_{\mathcal{V}_m\times(-\varepsilon_m,0)} \frac{\partial}{\partial y_j}\big[(\zeta_m\circ\vartheta_m)\mathrm{J}_m(\mathrm{A}_m)^j_i(\boldsymbol{v}^i\circ\vartheta_m)\big]\,dy\\ &= \sum_{m=1}^{K}\sum_{i,j=1}^{\mathrm{n}} \int_{\mathcal{V}_m\times(-\varepsilon_m,0)} \mathrm{J}_m(\mathrm{A}_m)^j_i(\zeta_m\circ\vartheta_m)_{,j}\,(\boldsymbol{v}^i\circ\vartheta_m)\,dy\\ &\quad+\sum_{m=1}^{K}\sum_{i,j=1}^{\mathrm{n}} \int_{\mathcal{V}_m\times(-\varepsilon_m,0)} (\zeta_m\circ\vartheta_m)\mathrm{J}_m(\mathrm{A}_m)^j_i(\boldsymbol{v}^i\circ\vartheta_m)_{,j}\,dy\,.\end{aligned}$$

Making change of variable $y=\vartheta_m^{-1}(x)$ in each $\mathcal{V}_m\times(-\varepsilon_m,0)$ again, by the fact that

$$\sum_{i,j=1}^{\mathrm{n}} (\mathrm{A}_m)^j_i(\boldsymbol{v}^i\circ\theta_m)_{,j} = (\mathrm{div}\boldsymbol{v})\circ\theta_m \qquad\text{and}\qquad \int_{\mathcal{W}_0}\mathrm{div}\big(\zeta_0\boldsymbol{v}\big)\,dx = 0\,,$$

we conclude that

$$\begin{aligned}\int_{\partial\Omega} \boldsymbol{v}\cdot\mathbf{N}\,dS &= \int_{\mathcal{W}_0}\mathrm{div}\big(\zeta_0\boldsymbol{v}\big)\,dx + \sum_{m=1}^{K}\int_{\mathcal{W}_m}(\boldsymbol{v}\cdot\nabla_x)\zeta_m\,dx + \sum_{m=1}^{K}\int_{\mathcal{W}_m}\zeta_m\mathrm{div}\boldsymbol{v}\,dx\\ &= \sum_{m=0}^{K}\int_{\mathcal{W}_m}(\boldsymbol{v}\cdot\nabla_x)\zeta_m\,dx + \sum_{m=0}^{K}\int_{\mathcal{W}_m}\zeta_m\mathrm{div}\boldsymbol{v}\,dx\\ &= \int_{\Omega}(\boldsymbol{v}\cdot\nabla_x)1\,dx + \int_{\Omega}\mathrm{div}\boldsymbol{v}\,dx = \int_{\Omega}\mathrm{div}\boldsymbol{v}\,dx\,.\end{aligned}$$

□

Letting $\boldsymbol{v}=(0,\cdots,0,f,0,\cdots,0)=f\mathrm{e}_i$, we obtain the following

Corollary A.184. *Let $\Omega\subseteq\mathbb{R}^{\mathrm{n}}$ be a bounded Lipschitz domain, and $f\in\mathscr{C}^1(\Omega)\cap\mathscr{C}(\bar\Omega)$. Then*

$$\int_{\Omega}\frac{\partial f}{\partial x_i}\,dx = \int_{\partial\Omega} f\,\mathbf{N}_i\,dS\,,$$

where $\mathbf{N}_i$ is the i-th component of the outward-pointing unit normal $\mathbf{N}$ of Ω.

Letting $\boldsymbol{v}$ be the product of a scalar function and a vector-valued function in Theorem A.183, we conclude the following

Corollary A.185. *Let $\Omega\subseteq\mathbb{R}^{\mathrm{n}}$ be a bounded Lipschitz domain, and $\boldsymbol{v}\in\mathscr{C}^1(\Omega;\mathbb{R}^{\mathrm{n}})\cap\mathscr{C}(\bar\Omega;\mathbb{R}^{\mathrm{n}})$ be a vector-valued function and $\varphi\in\mathscr{C}^1(\Omega)\cap\mathscr{C}(\bar\Omega)$ be a scalar function. Then*

$$\int_{\Omega}\varphi\,\mathrm{div}\boldsymbol{v}\,dx = \int_{\partial\Omega}(\boldsymbol{v}\cdot\mathbf{N})\varphi\,dS - \int_{\Omega}\boldsymbol{v}\cdot\nabla\varphi\,dx\,, \tag{A.49}$$

where $\mathbf{N}$ is the outward-pointing unit normal on $\partial\Omega$.

The divergence theorem on surfaces with boundary

This section is devoted to the divergence theorem on surfaces in $\mathbb{R}^3$ instead of domains of $\mathbb{R}^n$. To do so, we need to define what the divergence operator on a surface is, and this requires that we first define the vector fields on which the surface divergence operator acts.

Definition A.186. Let $\Sigma \subseteq \mathbb{R}^3$ be an open $\mathscr{C}^1$-surface; that is, Σ is of class $\mathscr{C}^1$ and $\Sigma \cap \partial\Sigma = \varnothing$. A vector field $\boldsymbol{u}$ defined on Σ is called a tangent vector field on Σ, denoted by $\boldsymbol{u} \in \mathbf{T}\Sigma$, if $\boldsymbol{u} \cdot \mathbf{N} = 0$ on Σ, where $\mathbf{N} : \Sigma \to \mathbb{S}^2$ is a unit normal vector field on Σ.

Having established (A.49), we find that the divergence operator div is the formal adjoint of the operator $-\nabla$. The following definition is motivated by this observation.

Definition A.187 (The surface gradient and the surface divergence)**.** Let $\Sigma \subseteq \mathbb{R}^n$ be a regular $\mathscr{C}^1$-surface. The surface gradient of a function $f : \Sigma \to \mathbb{R}$, denoted by $\nabla^{\partial\Omega} f$, is a vector-valued function from Σ to $\mathbf{T}_p\Sigma$ given, in a local parametrization $\{\mathcal{V}, \psi\}$, by

$$(\nabla^{\partial\Omega} f) \circ \psi = \sum_{\alpha,\beta=1}^{n-1} g^{\alpha\beta} \frac{\partial (f \circ \psi)}{\partial y_\alpha} \frac{\partial \psi}{\partial y_\beta} ,$$

where $[g^{\alpha\beta}]$ is the inverse matrix of the metric tensor $[g_{\alpha\beta}]$ associated with $\{\mathcal{V}, \psi\}$, and $\Big\{ \dfrac{\partial \psi}{\partial y_\beta} \Big\}_{\beta=1}^{2}$ are tangent vectors to Σ.

The surface divergence operator div_Σ is defined as the formal adjoint of $-\nabla^{\partial\Omega}$; that is, if $\boldsymbol{u} \in \mathbf{T}\Sigma$, then

$$-\int_\Sigma \boldsymbol{u} \cdot \nabla^{\partial\Omega} f \, dS = \int_\Sigma f \, \mathrm{div}_\Sigma \boldsymbol{u} \, dS \qquad \forall\, f \in \mathscr{C}_c^1(\Sigma; \mathbb{R}) .$$

In a local parametrization $(\mathcal{V}, \psi)$,

$$(\mathrm{div}_\Sigma \boldsymbol{u}) \circ \psi = \frac{1}{\sqrt{\mathrm{g}}} \sum_{\alpha,\beta=1}^{n-1} \frac{\partial}{\partial y_\alpha} \Big[\sqrt{\mathrm{g}} g^{\alpha\beta} \big((\boldsymbol{u} \circ \psi) \cdot \frac{\partial \psi}{\partial y_\beta} \big) \Big] ,$$

where $\mathrm{g} = \det(g)$ is the first fundamental form associated with $\{\mathcal{V}, \psi\}$.

Remark A.188. Suppose that $f : \mathcal{O} \subseteq \mathbb{R}^3 \to \mathbb{R}$ for some open set containing Σ. Then the surface gradient of f at $p \in \Sigma$ is the projection of the gradient vector $(\nabla f)(p)$

onto the tangent plane $T_p\Sigma$. In other words, let $\mathbf{N} : \Sigma \to \mathbb{R}^3$ be a continuous unit normal vector field on Σ, then

$$(\nabla^{\partial\Omega} f)(p) = (\nabla f)(p) - \big[(\nabla f)(p) \cdot \mathbf{N}(p)\big]\mathbf{N}(p) \quad (\text{or simply } \nabla^{\partial\Omega} f = \nabla f - (\nabla f \cdot \mathbf{N})\mathbf{N}) .$$

Definition A.189 (Surfaces with Boundary). An oriented $\mathscr{C}^k$-surface $\Sigma \subseteq \mathbb{R}^3$ is said to have $\mathscr{C}^\ell$-boundary $\partial\Sigma$ if there exists a collection of pairs $\{\mathcal{V}_m, \psi_m\}_{m=1}^K$, called a collection of local parametrization of $\bar{\Sigma}$, if

1. $\mathcal{V}_m \subseteq \mathbb{R}^2$ is open and $\psi_m : \mathcal{V}_m \to \mathbb{R}^3$ is one-to-one map of class $\mathscr{C}^k$ for all $m \in \{1, \cdots, K\}$;
2. $\psi_m(\mathcal{V}_m) \cap \Sigma \neq \varnothing$ for all $m \in \{1, \cdots, K\}$ and $\bar{\Sigma} \subseteq \bigcup_{m=1}^K \psi_m(\mathcal{V}_m)$;
3. $\psi_m : \mathcal{V}_m \to \psi_m(\mathcal{V}_m)$ is a $\mathscr{C}^k$-diffeomorphism if $\psi_m(\mathcal{V}_m) \subseteq \Sigma$;
4. $\psi_m : \mathcal{V}_m^+ \equiv \mathcal{V}_m \cap \{y_2 > 0\} \to \psi_m(\mathcal{V}_m) \cap \Sigma$ is a $\mathscr{C}^k$-diffeomorphism if $\mathcal{U}_m \cap \partial\Sigma \neq \varnothing$;
5. $\psi_m : \mathcal{V}_m \cap \{y_2 = 0\} \to \mathcal{U}_m \cap \partial\Sigma$ is of class $\mathscr{C}^\ell$ if $\mathcal{U}_m \cap \partial\Sigma \neq \varnothing$.

Now we are in the position of stating the divergence theorem on surfaces with boundary.

Theorem A.190. *Let $\Sigma \subseteq \mathbb{R}^3$ be an oriented $\mathscr{C}^1$-surface with $\mathscr{C}^1$-boundary $\partial\Sigma$, $\mathbf{N} : \Sigma \to \mathbb{S}^2$ be a continuous unit normal vector field on Σ, and $\mathbf{T} : \partial\Sigma \to \mathbb{S}^2$ be tangent vector on $\partial\Sigma$ such that $\mathbf{T}$ is compatible with $\mathbf{N}$ (which means $\mathbf{T} \times \mathbf{N}$ points away from Σ). Then*

$$\int_{\partial\Sigma} \boldsymbol{u} \cdot (\mathbf{T} \times \mathbf{N})\, ds = \int_{\Sigma} \operatorname{div}_\Sigma \boldsymbol{u}\, dS \qquad \forall\, \boldsymbol{u} \in \mathbf{T}\Sigma \cap \mathscr{C}^1(\Sigma; \mathbb{R}^3) \cap \mathscr{C}(\bar{\Sigma}; \mathbb{R}^3) ,$$

where $\operatorname{div}_\Sigma$ *is the surface divergence operator.*

Proof. Let $\{\mathcal{V}_m, \psi_m\}_{m=1}^K$ denote a collection of local parametrization of $\bar{\Sigma}$ such that $\psi_m(\mathcal{V}_m) \cap \partial\Sigma = \varnothing$ for $1 \leqslant m \leqslant J$, and $\psi_m(\mathcal{V}_m) \cap \partial\Sigma$ is non-empty and connected for $J+1 \leqslant m \leqslant K$. W.L.O.G., we can assume that $\mathcal{V}_m = B_m \equiv B(0, r_m)$ for some $r_m > 0$. Write $\mathcal{U}_m = \psi_m(\mathcal{V}_m)$, and let $\{g_m\}_{m=1}^K$ be the associated metric tensor, as well as the associated first fundamental form $\mathrm{g}_m = \det(g_m)$. Let $\{\zeta_m\}_{m=1}^K$ be a partition-of-unity of $\bar{\Sigma}$ subordinate to $\{\mathcal{U}_m\}_{m=1}^K$. Then

$$\begin{aligned}\int_\Sigma \mathrm{div}_\Sigma \boldsymbol{u}\, dS &= \sum_{m=1}^K \int_{\mathcal{U}_m\cap\Sigma} \zeta_m \mathrm{div}_\Sigma \boldsymbol{u}\, dS \\ &= \sum_{m=1}^J \sum_{\alpha,\beta=1}^2 \int_{B_m} (\zeta_m\circ\psi_m)\frac{\partial}{\partial y_\alpha}\Big[\sqrt{\mathrm{g}_m} g_m^{\alpha\beta}\big((\boldsymbol{u}\circ\psi_m)\cdot \frac{\partial\psi_m}{\partial y_\beta}\big)\Big] dy \\ &\quad + \sum_{m=J+1}^K \sum_{\alpha,\beta=1}^2 \int_{B_m^+} (\zeta_m\circ\psi_m)\frac{\partial}{\partial y_\alpha}\Big[\sqrt{\mathrm{g}_m} g_m^{\alpha\beta}\big((\boldsymbol{u}\circ\psi_m)\cdot \frac{\partial\psi_m}{\partial y_\beta}\big)\Big] dy\,.\end{aligned}$$

Let $\boldsymbol{n}$ denote the outward-pointing unit normal on either ∂B_m for $1 \leqslant m \leqslant J$ or ∂B_m^+ for $J+1 \leqslant m \leqslant K$. Since $\zeta_m \circ \vartheta_m = 0$ on $\partial B(0, r_m)$ for $1 \leqslant m \leqslant J$, and $\zeta_m \circ \vartheta_m = 0$ on $\{y_2 > 0\} \cap \partial B(0, r_m)$ for $J+1 \leqslant m \leqslant K$, the divergence theorem (on $\mathbb{R}^2$) implies that

$$\begin{aligned}\int_\Sigma \mathrm{div}_\Sigma \boldsymbol{u}\, dS &= -\sum_{m=1}^K \sum_{\alpha,\beta=1}^2 \int_{\psi_m^{-1}(\mathcal{U}_m\cap\Sigma)} \Big[\sqrt{\mathrm{g}_m} g_m^{\alpha\beta}\big((\boldsymbol{u}\circ\psi_m)\cdot \frac{\partial\psi_m}{\partial y_\beta}\big)\Big]\frac{\partial}{\partial y_\alpha}(\zeta_m\circ\psi_m)\, dy \\ &\quad + \sum_{m=J+1}^K \sum_{\alpha,\beta=1}^2 \int_{B_m\cap\{y_2=0\}} (\zeta_m\circ\psi_m)\boldsymbol{n}_\alpha\Big[\sqrt{\mathrm{g}_m} g_m^{\alpha\beta}\big((\boldsymbol{u}\circ\psi_m)\cdot \frac{\partial\psi_m}{\partial y_\beta}\big)\Big] dy_1 \\ &= -\sum_{m=1}^K \int_{\psi_m^{-1}(\mathcal{U}_m\cap\Sigma)} (\boldsymbol{u}\cdot\nabla_\Sigma\zeta_m)\circ\psi_m\sqrt{\mathrm{g}_m}\, dy \\ &\quad + \sum_{m=J+1}^K \int_{B_m\cap\{y_2=0\}} (\zeta_m\circ\psi_m)(\boldsymbol{u}\circ\psi_m)\cdot\Big[\sum_{\alpha,\beta=1}^2 \boldsymbol{n}_\alpha\sqrt{\mathrm{g}_m} g_m^{\alpha\beta}\frac{\partial\psi_m}{\partial y_\beta}\Big] dy_1\,.\end{aligned}$$

Since

$$\begin{aligned}&\sum_{m=1}^K \int_{\psi_m^{-1}(\mathcal{U}_m\cap\Sigma)} (\boldsymbol{u}\cdot\nabla_\Sigma\zeta_m)\circ\psi_m\sqrt{\mathrm{g}_m}\, dy \\ &\qquad = \sum_{m=1}^K \int_{\mathcal{U}_m\cap\Sigma} (\boldsymbol{u}\cdot\nabla^{\partial\Omega}\zeta_m)\, dS = \int_\Sigma \big(\boldsymbol{u}\cdot\nabla\sum_{m=1}^K\zeta_m\big)\, dS = 0\,,\end{aligned}$$

we conclude that

$$\int_\Sigma \mathrm{div}_\Sigma \boldsymbol{u}\, dS = \sum_{m=J+1}^K \int_{B_m\cap\{y_2=0\}} (\zeta_m\circ\psi_m)(\boldsymbol{u}\circ\psi_m)\cdot\Big[\sum_{\alpha,\beta=1}^2 \boldsymbol{n}_\alpha\sqrt{\mathrm{g}_m} g_m^{\alpha\beta}\frac{\partial\psi_m}{\partial y_\beta}\Big] dy_1\,.$$

On the other hand,

$$\begin{aligned}\int_{\partial\Sigma} \boldsymbol{u}\cdot(\mathbf{T}\times\mathbf{N})\,ds &= \sum_{m=J+1}^{K}\int_{\partial\Sigma\cap\mathcal{U}_m} \zeta_m \boldsymbol{u}\cdot(\mathbf{T}\times\mathbf{N})\,ds\\ &= \sum_{m=J+1}^{K}\int_{B_m\cap\{y_2=0\}} (\zeta_m\circ\psi_m)(\boldsymbol{u}\circ\psi_m)\cdot\Big[(\mathbf{T}\times\mathbf{N})\circ\psi_m\Big|\frac{\partial\psi_m}{\partial y_1}\Big|\Big]dy_1\,.\end{aligned}$$

Therefore, the theorem can be concluded as long as we can show that

$$\sum_{\alpha,\beta=1}^{2}\boldsymbol{n}_\alpha\sqrt{\mathrm{g}_m}\,g_m^{\alpha\beta}\,\frac{\partial\psi_m}{\partial y_\beta} = (\mathbf{T}\times\mathbf{N})\circ\psi_m\left|\frac{\partial\psi_m}{\partial y_1}\right| \qquad\text{on}\quad B_m\cap\{y_2=0\}\,. \tag{A.50}$$

Let $\boldsymbol{\tau}_m = \sum\limits_{\alpha,\beta=1}^{2}\boldsymbol{n}_\alpha\sqrt{\mathrm{g}_m}g_m^{\alpha\beta}\dfrac{\partial\psi_m}{\partial y_\beta}$ on $B_m\cap\{y_2=0\}$. Since $\boldsymbol{n}_\alpha = -\delta_{2\alpha}$, we find that $\boldsymbol{\tau}_m\cdot\dfrac{\partial\psi_m}{\partial y_1} = 0$ on $B_m\cap\{y_2=0\}$; thus

$$\boldsymbol{\tau}_m\cdot(\mathbf{T}\circ\psi_m) = 0 \qquad\text{on}\quad B_m\cap\{y_2=0\}\,.$$

Moreover, noting that $\boldsymbol{\tau}_m$ is a linear combination of tangent vectors $\dfrac{\partial\psi_m}{\partial y_\beta}$, we must have

$$\boldsymbol{\tau}_m\cdot(\mathbf{N}\circ\psi_m) = 0 \qquad\text{on}\quad B_m\cap\{y_2=0\}\,.$$

As a consequence,

$$\boldsymbol{\tau}_m /\!\!/ (\mathbf{T}\times\mathbf{N})\circ\psi_m \qquad\text{on}\quad B_m\cap\{y_2=0\}\,.$$

Since $(\mathbf{T}\times\mathbf{N})$ points away from Σ, while $\dfrac{\partial\psi_m}{\partial y_2}\circ\psi_m^{-1}\Big|_{\partial\Sigma}$ points toward Σ, by the fact that

$$\boldsymbol{\tau}_m\cdot\frac{\partial\psi_m}{\partial y_2} = \sum_{\alpha,\beta=1}^{2}\boldsymbol{n}_\alpha\sqrt{\mathrm{g}_m}g_m^{\alpha\beta}\frac{\partial\psi_m}{\partial y_\beta}\cdot\frac{\partial\psi_m}{\partial y_2} = -\sqrt{\mathrm{g}_m}g_m^{22} < 0\,,$$

we must have $\boldsymbol{\tau}_m\cdot(\mathbf{T}\times\mathbf{N})\circ\psi_m > 0$ on $B_m\cap\{y_2=0\}$. In other words,

$$\boldsymbol{\tau}_m = |\boldsymbol{\tau}_m|(\mathbf{T}\times\mathbf{N})\circ\psi_m \qquad\text{on}\quad B_m\cap\{y_2=0\}\,.$$

Finally, since

$$\boldsymbol{\tau}_m\cdot\boldsymbol{\tau}_m = \sum_{\alpha,\beta,\gamma,\delta=1}^{2}\mathrm{g}_m\,\boldsymbol{n}_\alpha\,\boldsymbol{n}_\gamma\,g_m^{\alpha\beta}g_m^{\gamma\delta}\,\frac{\partial\psi_m}{\partial y_\beta}\cdot\frac{\partial\psi_m}{\partial y_\delta} = \mathrm{g}_m g_m^{22} = g_{m11} = \left|\frac{\partial\psi_m}{\partial y_1}\right|^2,$$

we conclude that $\boldsymbol{\tau}_m = \left|\dfrac{\partial\psi_m}{\partial y_1}\right|(\mathbf{T}\times\mathbf{N})\circ\psi_m$ on $\{y_2=0\}$; thus (A.50) is established. □

Remark A.191. On $\partial\Sigma$, the vector $\mathbf{T}\times\mathbf{N}$ is "tangent" to Σ and points away from Σ. In other words, $\mathbf{T}\times\mathbf{N}$ can be treated as the "outward-pointing" unit "normal" of $\partial\Sigma$ which makes the divergence theorem on surfaces more intuitive.

A.7.4 The Stokes theorem

DEFINITION A.192 (The curl operator)**.** Let $\boldsymbol{u} : \Omega \subseteq \mathbb{R}^{\rm n} \to \mathbb{R}^{\rm n}$, n = 2 or n = 3, be a vector field.

1. For n = 2, the curl of u is a scalar function defined by
$$\operatorname{curl}\boldsymbol{u} = \sum_{i,j=1}^{2} \varepsilon_{3ij}\boldsymbol{u}^j_{,i}\,.$$
2. For n = 3, the curl of u is a vector-valued function defined by
$$(\operatorname{curl}\boldsymbol{u})^i = \sum_{j,k=1}^{3} \varepsilon_{ijk}\boldsymbol{u}^k_{,j}\,.$$

The function $\operatorname{curl}\boldsymbol{u}$ is also called the *vorticity* of $\boldsymbol{u}$, and is usually denoted by one single Greek letter ω.

THEOREM A.193 (The Stokes theorem)**.** *Let $\boldsymbol{u} : \Omega \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ be a smooth vector field, and Σ be a $\mathscr{C}^1$-surface with $\mathscr{C}^1$-boundary $\partial\Sigma$ in Ω. Then*
$$\int_{\partial\Sigma} \boldsymbol{u}\cdot\mathbf{T}\,ds = \int_{\Sigma} \operatorname{curl}\boldsymbol{u}\cdot\mathbf{N}\,dS\,,$$
where $\mathbf{N}$ and $\mathbf{T}$ are compatible normal and tangent vector fields.

To prove the Stokes theorem, we first establish the following

LEMMA A.194. *Let $\Omega \subseteq \mathbb{R}^3$ be a bounded Lipschitz domain, and $\boldsymbol{w} : \Omega \to \mathbb{R}^{\rm n}$ be a mooth vector-valued function. If $\Sigma \subseteq \Omega$ is an oriented $\mathscr{C}^1$-surface with normal $\mathbf{N}$, then*
$$\operatorname{curl}\boldsymbol{w}\cdot\mathbf{N} = \operatorname{div}_\Sigma(\boldsymbol{w}\times\mathbf{N}) \qquad \text{on} \quad \Sigma\,. \tag{A.51}$$

Proof. Let $\mathcal{O} \subseteq \Omega$ be a $\mathscr{C}^1$-domain such that $\Sigma \subseteq \partial\mathcal{O}$ and $\mathbf{N}$ is the outward-pointing unit normal on $\partial\mathcal{O}$. In other words, Σ is part of the boundary of $\mathcal{O}$. Since
$$(\nabla\varphi)^i = \frac{\partial\varphi}{\partial\mathbf{N}}\mathbf{N}^i + (\nabla_{\partial\mathcal{O}}\varphi)^i \qquad \text{on} \quad \partial\mathcal{O}\,,$$
by the divergence theorem we conclude that for all $\varphi \in \mathscr{C}^1(\bar{\mathcal{O}})$,
$$\begin{aligned}\int_{\partial\mathcal{O}} (\operatorname{curl}\boldsymbol{w}\cdot\mathbf{N})\varphi\,dS &= \int_{\mathcal{O}} \operatorname{curl}\boldsymbol{w}\cdot\nabla\varphi\,dx = \int_{\partial\mathcal{O}} (\mathbf{N}\times\boldsymbol{w})\cdot\nabla\varphi\,dS \\ &= \int_{\partial\mathcal{O}} (\mathbf{N}\times\boldsymbol{w})\cdot\nabla_{\partial\mathcal{O}}\varphi\,dS = \int_{\partial\mathcal{O}} \operatorname{div}_{\partial\mathcal{O}}(\boldsymbol{w}\times\mathbf{N})\varphi\,dS\,.\end{aligned}$$
Identity (A.51) is concluded since φ can be chosen arbitrarily on Σ. $\square$

Proof of the Stokes theorem. Using (A.51) and then applying the divergence theorem on surfaces with boundary (Theorem A.190), we find that

$$\int_\Sigma \operatorname{curl}\boldsymbol{u}\cdot\mathbf{N}\,dS = \int_\Sigma \operatorname{div}_\Sigma(\boldsymbol{u}\times\mathbf{N})\,dS = \int_{\partial\Sigma}(\boldsymbol{u}\times\mathbf{N})\cdot(\mathbf{T}\times\mathbf{N})\,ds = \int_{\partial\Sigma}(\boldsymbol{u}\cdot\mathbf{T})\,ds$$

in which the identity $(\boldsymbol{u}\times\mathbf{N})\cdot(\mathbf{T}\times\mathbf{N}) = \boldsymbol{u}\cdot\mathbf{T}$ is used. □

A.7.5 Reynolds' transport theorem

Let Ω_1 and Ω_2 be two Lipschitz domains of $\mathbb{R}^{\mathrm{n}}$ with outward-pointing unit normal $\mathbf{N}$ and n, respectively, and the map $\psi : \begin{cases} \Omega_1 & \to \quad \Omega_2 \\ \partial\Omega_1 & \to \quad \partial\Omega_2 \\ y & \mapsto x = \psi(y)\end{cases}$ be a diffeomorphism; that is, ψ is one-to-one and onto, and has smooth inverse. Let $f \in \mathscr{C}^1(\Omega_2)\cap\mathscr{C}(\bar{\Omega}_2)$, and $F = f\circ\psi$ which in turns belongs to $\mathscr{C}^1(\Omega_1)\cap\mathscr{C}(\bar{\Omega}_1)$. By the divergence theorem,

$$\int_{\Omega_2}\frac{\partial f}{\partial x_i}(x)dx = \int_{\partial\Omega_2}(f\mathbf{n}_i)(x)\,dS_x\,.$$

On the other hand, by the chain rule we have that

$$\frac{\partial F}{\partial y_i} = \frac{\partial(f\circ\psi)}{\partial y_i} = \sum_{j=1}^{\mathrm{n}}\Big[\frac{\partial f}{\partial x_j}\circ\psi\Big]\frac{\partial\psi^j}{\partial y_i}\,;$$

thus if $\mathrm{A} = (\nabla\psi)^{-1}$,

$$\frac{\partial f}{\partial x_i}\circ\psi = \sum_{j=1}^{\mathrm{n}}\mathrm{A}^j_i\frac{\partial F}{\partial y_j}\,. \tag{A.52}$$

Letting $\mathrm{J} = \det(\nabla\psi)$ be the Jacobian of ψ, by the change of variable $y = \psi(y)$ and the Piola identity,

$$\int_{\Omega_2}\frac{\partial f}{\partial x_i}(x)dx = \int_{\Omega_1}\frac{\partial f}{\partial x_i}\big(\psi(y)\big)\det(\nabla\psi)(y)dy = \sum_{j=1}^{\mathrm{n}}\int_{\Omega_1}\frac{\partial}{\partial y_j}(\mathrm{J}\mathrm{A}^j_i F)\,dy.$$

The divergence theorem again implies that

$$\int_{\Omega_2}\frac{\partial f}{\partial x_i}(x)dx = \sum_{j=1}^{\mathrm{n}}\int_{\Omega_1}\mathrm{J}\mathrm{A}^j_i F\mathbf{N}_j\,dS_y$$

which further implies that

$$\int_{\partial\Omega_2}(f\mathbf{n})(x)\,dS_x = \int_{\partial\Omega_1}F\frac{\mathrm{J}\mathrm{A}^{\mathrm{T}}\mathbf{N}}{|\mathrm{J}\mathrm{A}^{\mathrm{T}}\mathbf{N}|}|\mathrm{J}\mathrm{A}^{\mathrm{T}}\mathbf{N}|\,dS_y\,. \tag{A.53}$$

Let $\psi^*(dS_x)$ denote the pull-back of the surface element dS_x having the property that for any function h defined on $\partial\Omega_2 = \psi(\partial\Omega_1)$,

$$\int_{\psi(\partial\Omega_1)} h(x)\,dS_x = \int_{\partial\Omega_1} (h\circ\psi)(y)\psi^*(dS_x)\,;$$

in other words, $\psi^*(dS_x) = \sqrt{\mathrm{g}(y)}\,dS_y$ for some "Jacobian" $\sqrt{\mathrm{g}}$ of the map $\psi : \partial\Omega_1 \to \partial\Omega_2$. Therefore, (A.53) suggests that

$$\int_{\partial\Omega_2} f\mathbf{n}\,dS = \int_{\partial\Omega_1} \big[(f\mathbf{n})\circ\psi\big](y)\psi^*(dS_x) = \int_{\partial\Omega_1} (f\circ\psi)\frac{\mathrm{J}\mathrm{A}^T\mathbf{N}}{|\mathrm{J}\mathrm{A}^{\mathrm{T}}\mathbf{N}|}|\mathrm{J}\mathrm{A}^{\mathrm{T}}\mathbf{N}|\,dS_y\,.$$

Since f can be chosen arbitrarily, the equality above suggests that

$$\mathbf{n}\circ\psi = \frac{\mathrm{J}\mathrm{A}^T\mathbf{N}}{|\mathrm{J}\mathrm{A}^{\mathrm{T}}\mathbf{N}|} = \frac{\mathrm{A}^{\mathrm{T}}\mathbf{N}}{|\mathrm{A}^{\mathrm{T}}\mathbf{N}|} \tag{A.54}$$

and

$$\psi^*(dS_x) = |\mathrm{J}\mathrm{A}^{\mathrm{T}}\mathbf{N}|\,dS_y\,. \tag{A.55}$$

We finish this section by the following

Theorem A.195 (Reynolds' transport theorem)**.** *Let $\Omega\subseteq\mathbb{R}^{\mathrm{n}}$ be a smooth domain, $\psi:\Omega\times[0,T]\to\mathbb{R}^{\mathrm{n}}$ be a diffeomorphism, $\Omega(t)=\psi(\Omega,t)$ and $f(x,t)$ be a function defined on $\Omega(t)$. Then*

$$\frac{d}{dt}\int_{\Omega(t)} f(x,t)dx = \int_{\Omega(t)} f_t(x,t)dx + \int_{\partial\Omega(t)} (\sigma f)(x,t)\,dS_x\,, \tag{A.56}$$

where σ is the speed of the boundary in the direction of outward pointing normal of $\partial\Omega(t)$; that is, with $\mathbf{n}$ denoting the outward-pointing unit normal of $\Omega(t)$,

$$\sigma = (\psi_t\circ\psi^{-1})\cdot\mathbf{n}\,.$$

Proof. By the change of variable formula,

$$\int_{\Omega(t)} f(x,t)dx = \int_{\Omega} f(\psi(y,t),t)\det(\nabla\psi)(y,t)\,dy.$$

Let $f(\psi(y,t),t) = F(y,t)$, $\mathrm{A} = (\nabla\psi)^{-1}$, and $\mathrm{J} = \det(\nabla\psi)$. By (A.8) and (A.52), we find that

$$\begin{aligned}\frac{d}{dt}\int_{\Omega(t)} f(x,t)dx &= \int_{\Omega}\Big[f_t(\psi(y,t),t) + \psi_t(y,t)\cdot(\nabla_x f)(\psi(y,t),t)\Big]\mathrm{J}(y,t)dy \\ &\qquad + \sum_{i,j=1}^{\mathrm{n}}\int_{\Omega} F(y,t)(\mathrm{J}\mathrm{A}^j_i\psi^i_{t,j})(y,t)\,dy \\ &= \int_{\Omega} f_t(\psi(y,t),t)dy + \sum_{i,j=1}^{\mathrm{n}}\int_{\Omega}\Big[\psi^i_t\mathrm{A}^j_i F_{,j}\,\mathrm{J} + F\mathrm{J}\mathrm{A}^j_i\psi^i_{t,j}\Big](y,t)\,dy \\ &= \int_{\Omega}(f_t\circ\psi)\mathrm{J}dy + \sum_{i,j=1}^{\mathrm{n}}\int_{\Omega}\big(\mathrm{J}\mathrm{A}^j_i\psi^i_t F\big)_{,j}\,dy,\end{aligned}$$

where the Piola identity (A.10) is used to conclude the last equality. The divergence theorem then implies that

$$\frac{d}{dt}\int_{\Omega(t)} f(x,t)dx = \int_{\Omega}(f_t\circ\psi)\mathrm{J}\,dy + \sum_{i,j=1}^{\mathrm{n}}\int_{\partial\Omega}\mathrm{J}\mathrm{A}^j_i\mathbf{N}_j\psi^i_t F\,dS_y\,.$$

As a consequence, changing back to the variable x on the right-hand side, by (A.54) and (A.55) we conclude that

$$\frac{d}{dt}\int_{\Omega(t)} f(x,t)dx = \int_{\Omega(t)} f_t(x,t)dx + \sum_{i,j=1}^{\mathrm{n}}\int_{\partial\Omega(t)}(\sigma f)(x,t)\,dS_x\,. \qquad \square$$

A.8 Exercises

In this set of exercise, the Einstein summation convention is used.

Problem A.1. Complete the following.

1. Let $\delta_{\cdot\cdot}$'s are the Kronecker deltas. Prove

$$\varepsilon_{ijk}\varepsilon_{irs} = \delta_{jr}\delta_{ks} - \delta_{js}\delta_{kr}\,. \tag{A.57}$$

2. Use (A.57) to show the following identities:

   (a) $\boldsymbol{u}\times(\boldsymbol{v}\times\boldsymbol{w}) = (\boldsymbol{u}\cdot\boldsymbol{w})\boldsymbol{v} - (\boldsymbol{u}\cdot\boldsymbol{v})\boldsymbol{w}$ if $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}$ are three 3-vectors.

   (b) $\mathrm{curl}\mathrm{curl}\boldsymbol{u} = -\Delta\boldsymbol{u} + \nabla\mathrm{div}\boldsymbol{u}$ if $\boldsymbol{u}:\Omega\to\mathbb{R}^3$ is smooth.

   (c) $\boldsymbol{u}\times\mathrm{curl}\boldsymbol{u} = \frac{1}{2}\nabla(|\boldsymbol{u}|^2) - (\boldsymbol{u}\cdot\nabla)\boldsymbol{u}$ if $\boldsymbol{u}:\Omega\to\mathbb{R}^3$ is smooth.

Problem A.2. Let $\psi(\cdot,t):\Omega\to\Omega(t)$ be a diffeomorphism as defined in Theorem A.195, and $\mathrm{J}=\det(\nabla\psi)$ and $\mathrm{A}=(\nabla\psi)^{-1}$. Complete the proof of the Piola identity, identities (A.11), (A.54) and (A.55) by the following argument:

1. Let $\boldsymbol{u}(\cdot,t):\Omega(t)\to\mathbb{R}^{\mathrm{n}}$ be a smooth vector field. Show that
$$\int_{\Omega(t)}\operatorname{div}\boldsymbol{u}\,dx=\int_{\Omega}\mathrm{JA}^j_i(\boldsymbol{u}\circ\psi)^i_{,j}\,dy\,;$$
thus by the divergence theorem,
$$\int_{\partial\Omega(t)}\boldsymbol{u}\cdot\boldsymbol{n}\,dS_x=\int_{\partial\Omega}\mathrm{JA}^j_i(\boldsymbol{u}\circ\psi)^i\mathbf{N}_j\,dS_y-\int_{\Omega}(\mathrm{JA}^j_i)_{,j}\,(\boldsymbol{u}\circ\psi)^i dy\,. \tag{A.58}$$

2. By (A.58),
$$\int_{\Omega}(\mathrm{JA}^j_i)_{,j}\,(\boldsymbol{u}\circ\psi)^i dy=0\quad\forall\,\boldsymbol{u}(\cdot,t):\Omega(t)\to\mathbb{R}^{\mathrm{n}}\text{ vanishing on }\partial\Omega(t).$$
As a consequence, the Piola identity is valid.

3. By the Piola identity, (A.58) implies that
$$\int_{\partial\Omega(t)}\boldsymbol{u}\cdot\boldsymbol{n}\,dS_x=\int_{\partial\Omega}\mathrm{JA}^j_i(\boldsymbol{u}\circ\psi)^i\mathbf{N}_j\,dS_y\quad\forall\,\boldsymbol{u}(\cdot,t):\Omega(t)\to\mathbb{R}^{\mathrm{n}}\text{ smooth.}$$
Therefore, identities (A.54) and (A.55) are also valid.

4. By identity (A.11) (which is obtained independent of), show that
$$\mathrm{J}_{,k}=\mathrm{JA}^j_i\psi^i_{,jk}\,.$$

Appendix B

Important Topics in Functional Analysis

B.1 The Hahn-Banach Theorem

Definition B.1. A vector space X is said to be a topological vector space if there is a topology τ on X so such that

(a) every point of X is a closed set, and

(b) the vector space operations (addition of vectors and multiplication with scalars) are continuous with respect to τ .

Definition B.2. The dual space of a topological vector space X is the vector space X' whose elements are the continuous linear functionals on X .

Proposition B.3. *A complex-linear functional on X is in X' if and only if its real part is continuous, and that every continuous real-linear $u : X \to \mathbb{R}$ is the real part of a unique $f \in X'$.*

Definition B.4. A map p from a real vector space V to $\mathbb{R} \cup \{\pm\infty\}$ is said to be *sub-linear over V* if

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

Theorem B.5. *Let X be a real vector space, p a sub-linear function over X, M a vector subspace of X. Suppose that T a linear functional over M and $Tx \leqslant p(x)$ on M. Then there exists a linear functional $\tilde{T}$ over X such that*

$$\tilde{T}x = Tx \qquad \forall\ x \in M\,,$$

and

$$-p(-x) \leqslant \tilde{T}x \leqslant p(x) \qquad \forall\ x \in X\,.$$

Corollary B.6. *If X is a normed space and $x_0 \in X$, there exists $T \in X'$ such that*

$$Tx_0 = \|x_0\|_X \qquad \textit{and} \qquad |Tx| \leqslant \|x\|_X \qquad \forall\ x \in X\,.$$

Theorem B.7. *Let A and B are disjoint, non-empty, convex sets in a topological vector space X.*

(a) *If A is open, then there exists $T \in X'$ and $\gamma \in \mathbb{R}$ such that*

$$Re\ Tx < \gamma \leqslant Re\ Ty$$

for every $x \in A$ and every $y \in B$.

(b) *If A is compact, B is closed, and X is locally convex, then there exist $T \in X'$, $\gamma_1, \gamma_2 \in \mathbb{R}$ such that*

$$Re\ Tx < \gamma_1 < \gamma_2 < Re\ Ty$$

for every $x \in A$ and every $y \in B$.

Theorem B.8. *Suppose M is a subspace of a locally convex space X, and $x_0 \in X$. If x_0 is not in the closure of M, then there exists $T \in X'$ such that $Tx_0 = 1$ but $Tx = 0$ for every $x \in M$.*

Theorem B.9. *If f is continuous linear functional on a subspace M of a locally convex space X, then there exists $T \in X'$ such that $T = f$ on M.*

Theorem B.10. *Suppose B is a convex, balanced, closed set in a locally convex space X, $x_0 \in X$, but $x_0 \notin B$. Then there exists $T \in X'$ such that $|Tx| \leqslant 1$ for all $x \in B$, but $Tx_0 > 1$.*

B.2 The Open Mapping and Closed Graph Theorem

Theorem B.11 (The Baire Category Theorem)**.** *Let X be a complete metric space.*

(a) *If $\{U_n\}_{n=1}^{\infty}$ is a sequence of open dense subsets of X, then $\bigcap\limits_{n=1}^{\infty} U_n$ is dense in X.*

(b) *X is not a countable union of nowhere dense sets.*

Definition B.12 (Open mapping)**.** Let X and Y be two topological vector spaces. A mapping $f : X \to Y$ is said to be open if $f(U)$ is open in Y whenever U is open in X.

Theorem B.13 (The Open Mapping Theorem)**.** *Suppose that X and Y be Banach spaces, and $T \in \mathscr{B}(X, Y)$ is surjective (i.e., onto). Then T is an open mapping.*

Theorem B.14 (A generalization of the Open Mapping Theorem)**.** *Suppose that X be a Banach space, Y be a topological vector space, and $T : X \to Y$ is linear, continuous and surjective (i.e., onto). Then T is an open mapping.*

Corollary B.15 (The Bounded Inverse Theorem)**.** *Suppose that X and Y be Banach spaces, and $T \in \mathscr{B}(X, Y)$ is bijective (i.e., one-to-one and onto), then the inverse map of T is bounded, or $T^{-1} \in \mathscr{B}(Y, X)$. Equivalently, there exist positive real numbers c and C such that*

$$c\|x\|_X \leqslant \|Tx\|_Y \leqslant C\|x\|_X \qquad \forall\ x \in X .$$

Theorem B.16 (The Closed Graph Theorem)**.** *Suppose that X and Y are Banach spaces, and $T : X \to Y$ is linear. If $G = \{(x, Tx) \mid x \in X\}$ is closed in $X \times Y$, then $T \in \mathscr{B}(X, Y)$.*

B.3 Compact Operators

Definition B.17 (Compact operators)**.** Suppose X and Y are Banach spaces and U is the open unit ball in X. A linear map $T : X \to Y$ is said to be compact if the closure of $T(U)$ is compact in Y. It is clear that T is then bounded. Thus $T \in \mathscr{B}(X, Y)$.

DEFINITION B.18. An operator $T \in \mathscr{B}(X)$ is said to be invertible if there exists $S \in \mathscr{B}(X)$ such that

$$ST = I = TS\,.$$

In this case, we write $S = T^{-1}$.

DEFINITION B.19 (Spectrum and resolvent set)**.** The specturm $\sigma(T)$ of an operator $T \in \mathscr{B}(X)$ is the set of all scalars λ such that $T - \lambda I$ is not invertible, and the resolvent set $\rho(T)$ is the complement of $\sigma(T)$ in the scalar field. Thus $\lambda \in \sigma(T)$ if and only if at least one of the following two statements is true:

(i) The range of $T - \lambda I$ is not all of X.

(ii) $T - \lambda I$ is not one-to-one.

DEFINITION B.20 (Classification of $\sigma(T)$)**.** The spectrum of $T \in \mathscr{B}(X)$ is the (disjoint) union of the following three sets:

(i) The point spectrum $\sigma_p(T) = \{\lambda \in \mathbb{C} \mid T - \lambda I \text{ is not one-to-one}\}$. If $\lambda \in \sigma_p(T)$, λ is also called an eigenvalue of T.

(ii) The continuous spectrum

$$\sigma_c(T) = \{\lambda \in \mathbb{C} \mid T - \lambda I \text{ is one-to-one, and has dense range}\}.$$

(iii) The residual spectrum

$$\sigma_r(T) = \{\lambda \in \mathbb{C} \mid T - \lambda I \text{ is one-to-one, and does not have dense range}\}.$$

PROPOSITION B.21. *The spectrum of a bounded operator $T \in \mathscr{B}(X)$ is bounded.*

THEOREM B.22. *Let X and Y be Banach spaces.*

(a) *If $T \in \mathscr{B}(X,Y)$ and $\dim R(T) < \infty$, then T is compact.*

(b) *If $T \in \mathscr{B}(X,Y)$, T is compact, and $R(T)$ is closed, then $\dim R(T) < \infty$.*

(c) *The compact operators form a closed subspace of $\mathscr{B}(X,Y)$ in its norm-topology.*

(d) *If $T \in \mathscr{B}(X)$, T is compact, and $\lambda \neq 0$, then $\dim N(T - \lambda I) < \infty$.*

(e) *If* $\dim X = \infty$, $T \in \mathscr{B}(X)$, *and* T *is compact, then* $0 \in \sigma(T)$.

(f) *If* $S \in \mathscr{B}(X)$, $T \in \mathscr{B}(X)$, *and* T *is compact, so are* ST *and* TS.

Proof. (a) and (f) are trivial and left as exercises.

(b) If $Y \equiv R(T)$ is closed, then Y is complete, so that T is an open mapping of X onto $R(X)$. Let U be the unit ball in X, then $V \equiv TU$ is open in Y. Since T is compact, V is pre-compact. Therefore, there exist $y_1, \cdots, y_m$ such that

$$\overline{V} \subseteq \bigcup_{j=1}^{m} (y_j + \frac{1}{2}V). \tag{1}$$

Let Z be the vector space spanned by $y_1, \cdots, y_m$. Then $\dim Z \leqslant m$, and Z is a closed subspace of Y. We also note that (1) implies $V \subseteq Z + \frac{1}{2}V$. Since $Z = \lambda Z$ for all $\lambda \neq 0$,

$$V \subseteq Z + \frac{1}{2}V \subseteq Z + Z + \frac{1}{4}V = Z + \frac{1}{4}V.$$

We then see that

$$V \subseteq \bigcap_{n=1}^{\infty} (Z + 2^{-n}V) = Z.$$

However, it would further implies that $kV \subseteq Z$ for all $k \in \mathbb{N}$, so $Z = Y$.

(c) Let Σ be the collection of compact operators in $\mathscr{B}(X, Y)$, U be the unit ball in X, and $T \in \overline{\Sigma}$. For every $r > 0$, there exists $S \in \Sigma$ with $\|S - T\|_{\mathscr{B}(X,Y)} < r$. Since SU is totally bounded, there exists points $x_1, \cdots, x_n$ in U such that SU is covered by the balls $B(Sx_i, r)$. Since $\|Sx - Tx\|_Y \leqslant r$ for every $x \in U$, it follows that TU is covered by the balls of $B(Tx_i, 3r)$. Thus TU is totally bounded as well, so $T \in \Sigma$.

(d) Let $Y = N(T - \lambda I)$. The restriction of T to Y is a compact operator whose range is Y. By (b), $\dim(Y) < \infty$.

(e) $T : X \to X$ cannot be an onto map since if it is onto, then $\dim R(T) = \infty$ which contradicts to (b). □

Definition B.23 (Adjoint operators)**.** The adjoint operator T^* of an operator $T \in \mathscr{B}(X, Y)$ is the unique bounded operator belonging to $\mathscr{B}(Y', X')$ satisfying

$$\langle Tx, y^* \rangle_Y = \langle x, T^* y^* \rangle_X.$$

THEOREM B.24. *Suppose X and Y are Banach spaces and $T \in \mathscr{B}(X, Y)$. Then T is compact if and only if T^* is compact.*

Proof. ($\Rightarrow$) Suppose T is compact. Let $\{y_n^*\}_{n=1}^{\infty}$ be a sequence in the unit ball of Y'. Define

$$f_n(y) = \langle y, y_n^* \rangle_Y \qquad \forall\, y \in Y .$$

Since $|f_n(y_1) - f_n(y_2)| \leqslant \|y_1 - y_2\|_Y$, $\{f_n\}_{n=1}^{\infty}$ is equi-continuous. Since $T(U)$ has compact closure in Y (as before, U is the unit ball of X), Arzela-Ascoli theorem implies that $\{f_n\}_{n=1}^{\infty}$ has a subsequence $\{f_{n_j}\}_{j=1}^{\infty}$ that converges uniformly on $T(U)$. Since

$$\|T^* y_{n_i}^* - T^* y_{n_j}^*\|_{X'} = \sup_{x \in U} |\langle Tx, y_{n_i}^* - y_{n_j}^* \rangle_Y| = \sup_{x \in U} |f_{n_i}(Tx) - f_{n_j}(Tx)| ,$$

the completeness of X' implies that $\{T^* y_{n_j}^*\}_{j=1}^{\infty}$ converges. Hence T^* is compact.
($\Leftarrow$) can be proved in the same fashion. □

DEFINITION B.25. Suppose M is a closed subspace of a topological vector space X. If there exists a closed subspace N of X such that

$$X = M + N \qquad \text{and} \qquad M \cap N = \{0\} ,$$

then M is said to be complemented in X. In this case, X is said to be the direct sum of M and N, and the notation $X = M \oplus N$ is used.

LEMMA B.26. *Let M be a closed subspace of a Banach space X.*

(a) *If $\dim M < \infty$, then M is complemented in X.*

(b) *If $\dim(X/M) < \infty$, then M is complemented in X.*

The dimension of X/M is also called the codimension of M in X.

Proof. Note that the closedness of M is only used in (b), while in (a) the closedness is implied by the finite dimensionality (so no assumption is needed).

(a) Let $\{e_1, \cdots, e_n\}$ be a basis for M. Then every $x \in M$ has a unique representation

$$x = \alpha_1(x) e_1 + \cdots + \alpha_n(x) e_n .$$

α_i is a continuous linear functional which vanishes on the span of $\{e_1, \cdots, e_{i-1}, e_{i+1}, \cdots, e_n\}$, and can be extended to a continuous linear functional that only take non-zero values in the 1-dimensional space spanned by e_i. Let N be the intersection of the null space of these extensions. Then $X = M \oplus N$.

(b) Let $\{e_1, \cdots, e_n\}$ be a basis of X/M (closedness of M is used to define the quotient space), and $\pi : X \to X/M$ be the quotient map. Pick $x_i \in X$ so that $\pi x_i = e_i$, and define N to be the span of $\{x_1, \cdots, x_n\}$.
Then $X = M \oplus N$. □

Lemma B.27. *Let M be a subspace of a normed space X. If M is not dense in X, and if $r > 1$, then there exists $x \in X$ such that*

$$\|x\|_X < r \qquad \textit{but} \qquad \|x - y\|_X \geqslant 1 \quad \forall\, y \in \mathcal{M}.$$

Proof. There exists $x_1 \in X$ whose distance from M is 1, that is,

$$\inf\{\|x_1 - y\|_X | y \in M\} = 1.$$

Choose $y_1 \in M$ such that $\|x_1 - y_1\|_X < r$, and put $x = x_1 - y_1$. □

Theorem B.28. *If X is a Banach space, $T \in \mathscr{B}(X)$, T is compact, and $\lambda \neq 0$, then $T - \lambda I$ has closed range.*

Proof. By (d) of Theorem B.22, $\dim N(T - \lambda I) < \infty$. By (a) of Lemma B.26, X is the direct sum of $N(T - \lambda I)$ and a closed subspace M. Define an operator $S \in \mathscr{B}(M, X)$ by

$$Sx = Tx - \lambda x.$$

Then S is one-to-one on M. Also, $R(S) = R(T - \lambda I)$. Similar to the proof of Lax-Milgram theorem, to show that $R(S)$ is closed, it suffices to show the existence of an $r > 0$ such that

$$r\|x\|_X \leqslant \|Sx\|_X \qquad \forall\, x \in M.$$

Suppose the contrary that for every $r > 0$, there exists $\{x_n\}$ in M such that $\|x_n\|_X = 1$, $Sx_n \to 0$, and (after passage to a subsequence) $Tx_n \to x_0$ for some $x_0 \in X$ (by the compactness of T). It follows that $\lambda x_n \to x_0$. Thus $x_0 \in M$ since M is a closed subspace, and

$$Sx_0 = \lim_{n \to \infty} (\lambda S x_n) = 0.$$

Since S is one-to-one, $x_0 = 0$. However, $\|x_n\|_X = 1$ for all n, and $x_0 = \lim_{n\to\infty} \lambda x_n$, hence $\|x_0\|_X = |\lambda| > 0$. □

COROLLARY B.29. *The continuous spectrum of a compact operator $T \in \mathscr{B}(X)$ contains at most one point, namely 0.*

THEOREM B.30. *Suppose X is a Banach space, $T \in \mathscr{B}(X)$, T is compact, $r > 0$, and E is a set of eigenvalues λ of T such that $|\lambda| > r$. Then*

(a) *for each $\lambda \in E$, $R(T - \lambda I) \neq X$, and*

(b) *E is a finite set.*

Proof. We first show that either (a) or (b) is false then there exist closed subspaces M_n of X and scalars $\lambda_n \in E$ such that

$$M_1 \subsetneq M_2 \subsetneq M_3 \subsetneq \cdots , \tag{1}$$

$$T(M_n) \subseteq M_n \qquad \text{for } n \geqslant 1 , \tag{2}$$

$$(T - \lambda_n I)(M_n) \subseteq M_{n-1} \qquad \text{for } n \geqslant 2 . \tag{3}$$

Suppose (a) is false. Then $R(T - \lambda_0 I) = X$ for some $\lambda_0 \in E$. Let $S = T - \lambda_0 I$, and define $M_n = N(S^n)$, i.e., the null space of S^n. Since λ_0 is an eigenvalue of T, there exists $x_1 \in M_1$, $x_1 \neq 0$. Since $R(S) = X$, there is a sequence $\{x_n\}_{n=1}^{\infty}$ in X such that $Sx_{n+1} = x_n$, $n = 1, 2, 3, \cdots$. Then

$$S^n x_{n+1} = x_1 \neq 0 \qquad \text{but} \qquad S^{n+1} x_{n+1} = Sx_1 = 0 .$$

Hence M_n is a proper closed subspace of M_{n+1}. It follows that (1) to (3) hold, with $\lambda_n = \lambda_0$.

Suppose (b) is false. Then E contains a sequence $\{\lambda_n\}$ of distinct eigenvalues of T. Choose corresponding eigenvectors e_n, and let M_n be the (finite-dimensional, hence closed) subspace of X spanned by $\{e_1, \cdots , e_n\}$. Since λ_n are distinct, $\{e_1, \cdots , e_n\}$ is a linearly independent set, so that M_{n-1} is a proper subspace of M_n. This gives (1). If $x \in M_n$, then

$$x = \alpha_1 e_1 + \cdots \alpha_n e_n ,$$

which shows that $Tx \in M_n$ and

$$(T - \lambda_n I)x = \alpha_1(\lambda_1 - \lambda_n)e_1 + \cdots + \alpha_{n-1}(\lambda_{n-1} - \lambda_n)e_{n-1} \in M_{n-1} .$$

Thus (2) and (3) hold.

Once we have closed subspace M_n satisfying (1) to (3), Lemma B.27 gives us vectors $y_n \in M_n$, for $n = 2, 3, 4, \cdots$, such that

$$\|y_n\|_X \leqslant 2 \quad \text{and} \quad \|y_n - x\|_X \geqslant 1 \quad \text{if } x \in M_{n-1}\,. \tag{4}$$

If $2 \leqslant m < n$, define

$$z = Ty_m - (T - \lambda_n I)y_n\,.$$

By (2) and (3), $z \in M_{n-1}$. Hence (4) shows that

$$\|Ty_n - Ty_m\|_X = \|\lambda_n y_n - z\|_X = |\lambda_n|\|y_n - \lambda_n^{-1} z\|_X \geqslant |\lambda_n| > r\,.$$

The sequence $\{Ty_n\}_{n=1}^{\infty}$ has therefore no convergent subsequences, although $\{y_n\}_{n=1}^{\infty}$ is bounded, contradicting to the compactness of T. □

Remark B.31. Let $\mathcal{H}$ denote a Hilbert space. $T \in \mathscr{B}(\mathcal{H})$ is said to be normal if $TT^* = T^*T$. A much deeper result states that a normal operator $T \in B(\mathcal{H})$ is compact if and only if it satisfies the following two conditions:

(a) $\sigma(T)$ has no limit point except possibly 0.

(b) If $\lambda \neq 0$, then $\dim N(T - \lambda I) < \infty$.

Theorem B.32 (The Fredholm Alternative)**.** *Suppose X is a Banach space, $T \in \mathscr{B}(X)$, and T is compact.*

(a) *If $\lambda \neq 0$, then the four numbers*

$$\begin{aligned} \alpha &= \dim N(T - \lambda I) & \alpha^* &= \dim N(T^* - \lambda I) \\ \beta &= \dim X/R(T - \lambda I) & \beta^* &= \dim X/R(T^* - \lambda I) \end{aligned}$$

are equal and finite.

(b) *If $\lambda \neq 0$ and $\lambda \in \sigma(T)$, then λ is an eigenvalue of T and of T^*.*

(c) *$\sigma(T)$ is compact, at most countable, and has at most one limit point, namely, 0.*

Proof. Suppose M_0 is a closed subspace of a locally convex space Y, and k is a positive integer such that $k \leqslant \dim Y/M_0$. Then there are vectors $y_1, \cdots, y_k$ in Y such that the vector space M_i generated by M_0 and $y_1, \cdots, y_i$ contains M_{i-1} as a proper subspace. Each M_i is closed, and hence by Theorem B.8, there are continuous linear functionals $T_1, \cdots, T_k$ on Y such that $T_i y_i = 1$ but $T_i y = 0$ for all $y \in M_{i-1}$. These functionals are linearly independent, so if Σ denotes the space of all continuous linear functionals on Y that annihilate M_0, then

$$\dim Y/M_0 \leqslant \dim \Sigma .$$

Let $S = T - \lambda I$. Apply this with $Y = X$, $M_0 = R(S)$. Since $R(S)$ is closed, $\Sigma = R(S)^{\perp} = N(S^*)$, so $\beta \leqslant \alpha^*$. Next, take $Y = X'$ with its weak*-topology, and $M_0 = R(S^*)$. A result from functional analysis states that $R(S^*)$ is weak*-closed. Since Σ consists of all weak*-continuous linear functional on X' that annihilate $R(S^*)$, Σ is isomorphic to ${}^{\perp}R(S^*) = N(S)$, hence $\beta^* \leqslant \alpha$.

Next we show that $\alpha \leqslant \beta$, and the same proof can be used to show that $\alpha^* \leqslant \beta^*$, so the proof of (a) (and hence (b) and (c)) is complete. Assume the contrary that $\alpha > \beta$. By (d) of Theorem B.22, $\alpha < \infty$. By Lemma B.26, there exists closed subspaces E and F such that $\dim F = \beta$ and

$$X = N(S) \oplus E = R(S) \oplus F .$$

Every $x \in X$ has unique representation $x = x_1 + x_2$, with $x_1 \in N(S)$, $x_2 \in E$. Define $\pi : X \to N(S)$ by $\pi x = x_1$. It is easy to see (by the closed graph theorem B.16) that π is continuous.

Since we assume that $\dim N(S) > \dim F$, there is a linear mapping ϕ of $N(S)$ onto F such that $\phi x_0 = 0$ for some $x_0 \neq 0$. Define

$$\Phi x = Tx + \phi\pi x \qquad \forall\ x \in X .$$

Then $\Phi \in \mathscr{B}(X)$. Since $\dim R(\phi\pi) < \infty$, $\phi\pi$ is a compact operator, hence so is Φ.

Observe that $\Phi - \lambda I = S + \phi\pi$. If $x \in E$, then $\pi x = 0$, so $(\Phi - \lambda I)x = Sx$; hence

$$(\Phi - \lambda I)(E) = R(S) .$$

If $x \in N(S)$, then $\pi x = x$, $(\Phi - \lambda I)x = \phi x$; hence

$$(\Phi - \lambda I)(N(S)) = \phi(N(S)) = F .$$

Therefore, $R(\Phi - \lambda I) = R(S) + F = X$. Moreover, λ is an eigenvalue of Φ (with x_0 as a corresponding eigenvector), and since Φ is compact, Theorem B.30 states that $R(\Phi - \lambda I)$ cannot be all of X. $\square$

Corollary B.33. *The residual spectrum of a compact operator $T \in \mathscr{B}(X)$ contains at most one point, namely 0. Moreover, $\sigma(T) = \sigma_p(T) \cup \{0\}$.*

Corollary B.34. *Suppose that $\mathcal{H}$ is a Hilbert space, and $T \in \mathscr{B}(\mathcal{H})$ is compact. Then $R(T - \lambda I) = N(T^* - \lambda I)^\perp$ for all $\lambda \neq 0$.*

Remark B.35. If $T \in \mathscr{B}(X)$ is compact, then the injectivity of $T - \lambda I$ implies the invertibility of $T - \lambda I$ if $\lambda \neq 0$.

Remark B.36. A much deeper result states that the spectrum of a bounded operator $T \in \mathscr{B}(X)$ is also compact.

B.3.1 Symmetric operators on Hilbert Spaces

Let $\mathcal{H}$ be a Hilbert space, and $T \in \mathscr{B}(\mathcal{H})$. By Riesz representation theorem, given a continuous linear functional $y^* \in \mathcal{H}'$, there exists $y \in \mathcal{H}$ such that

$$\langle h, y^* \rangle_{\mathcal{H}} = (h, y)_{\mathcal{H}} \qquad \forall\ h \in \mathcal{H}.$$

In particular, let $h = Tx$, and suppose the representation of T^*y^* is z, then

$$(x, z)_{\mathcal{H}} = \langle x, T^*y^* \rangle_{\mathcal{H}} = \langle Tx, y^* \rangle_{\mathcal{H}} = (Tx, y)_{\mathcal{H}} \qquad \forall\ h \in \mathcal{H}.$$

The element $z \in \mathcal{H}$ is denoted by $T'y$. In this case, T' is also called the adjoint operator of T (and T' can be thought as the representation of T^*).

Definition B.37 (Symmetry)**.** The operator $T \in \mathscr{B}(\mathcal{H})$ is called symmetric if $T = T'$.

Lemma B.38. *Suppose that $T \in \mathscr{B}(\mathcal{H})$ be symmetric, and*

$$m \equiv \inf_{\|u\|_{\mathcal{H}}=1} (Tu, u)_{\mathcal{H}}, \qquad M \equiv \sup_{\|u\|_{\mathcal{H}}=1} (Tu, u)_{\mathcal{H}}.$$

Then $\sigma(T) \subseteq [m, M]$, and $m, M \in \sigma(T)$.

Proof. Let $\lambda > M$. Then $\mathcal{L} : \mathcal{H} \to \mathcal{H}'$ defined by

$$\langle \mathcal{L}u, \varphi \rangle_{\mathcal{H}} = (\lambda u - Tu, \varphi)_{\mathcal{H}} \qquad \forall\, \varphi \in \mathcal{H}$$

is bounded and coercive: the boundedness is trivial, and the coercivity follows from that

$$\langle \mathcal{L}u, u \rangle_H = (\lambda u - Tu, u)_{\mathcal{H}} \geqslant (\lambda - M)\|u\|^2_{\mathcal{H}} .$$

Therefore, by the Lax-Milgram theorem, $\mathcal{L} : H \to \mathcal{H}'$ is one-to-one and onto, so is $\lambda I - T$ since $\lambda I - T$ is the representation of $\mathcal{L}$. Therefore, $\lambda \notin \sigma(T)$. Similarly, $\lambda \notin \sigma(T)$ if $\lambda < m$. So $\sigma(T) \subseteq [m, M]$.

Let $[u, v] = (Mu - Tu, v)_{\mathcal{H}}$. The proof of the Schwarz inequality (Proposition **??**) implies that

$$\big|[u, v]\big| \leqslant \big|[u, u]\big|^{1/2}\big|[v, v]\big|^{1/2} .$$

Taking the supremum over all v such that $\|v\|_{\mathcal{H}} = 1$, then

$$\|Mu - Tu\|_{\mathcal{H}} \leqslant C(Mu - Tu, u)^{1/2}_{\mathcal{H}} \qquad \forall\, u \in \mathcal{H} \tag{B.1}$$

for some constant C.

Let $\{u_k\}_{k=1}^{\infty}$ be such that $\|u_k\|_{\mathcal{H}} = 1$, and $(Tu_k, u_k)_{\mathcal{H}} \to M$. Then (B.1) implies $\|Mu_k - Tu_k\|_{\mathcal{H}} \to 0$ as $k \to \infty$. If $M \notin \sigma(T)$, $MI - T$ is invertible and has a bounded inverse (by the bounded inverse theorem), so

$$u_k = (MI - T)^{-1}(Mu_k - Tu_k) \to 0 \text{ in } \mathcal{H}$$

which contradicts to $\|u_k\|_{\mathcal{H}} = 1$ for all k. Hence $M \in \sigma(T)$. Similarly, $m \in \sigma(T)$. □

Theorem B.39. *Let $\mathcal{H}$ be a separable Hilbert space, and suppose that $T \in B(\mathcal{H})$ is compact and symmetric. Then there exists a countable orthonormal basis of $\mathcal{H}$ consisting of eigenvectors of T.*

Proof. Let $\{\lambda_k\}_{k=1}^{\infty}$ be the sequence of distinct eigenvalues of T, $\lambda_k \neq 0$. Set $\lambda_0 = 0$, and $\mathcal{H}_k = N(T - \lambda_k I)$ for $k \geqslant 0$. Then $\dim \mathcal{H}_k < \infty$ if $k > 0$. Moreover, if $x_i \in \mathcal{H}_i$ and $x_j \in \mathcal{H}_j$, then

$$\lambda_i(x_i, x_j)_{\mathcal{H}} = (Tx_i, x_j)_{\mathcal{H}} = (x_i, Tx_j)_{\mathcal{H}} = \overline{\lambda_j}(x_i, x_j)_{\mathcal{H}} \;\Rightarrow\; (x_i, x_j)_{\mathcal{H}} = 0 \text{ if } i \neq j .$$

Therefore, the subspaces $\mathcal{H}_i$ and $\mathcal{H}_j$ are orthogonal.

Let $\tilde{\mathcal{H}}$ be the smallest subspace of $\mathcal{H}$ consisting of all these $\mathcal{H}_i$, $i = 0, 1, \cdots$. Then

$$\tilde{\mathcal{H}} = \Big\{ \sum_{k=0}^{m} c_k u_k \;\Big|\; m \in \mathbb{N} \cup \{0\}, u_k \in \mathcal{H}_k, a_k \in \mathbb{R} \Big\}.$$

We note that $T(\tilde{\mathcal{H}}) \subseteq \tilde{\mathcal{H}}$, and this further implies that $T(\tilde{\mathcal{H}}^{\perp}) \subseteq \tilde{\mathcal{H}}^{\perp}$ since

$$(Tu, v)_{\mathcal{H}} = (u, Tv)_{\mathcal{H}} = 0 \qquad \forall\, u \in \tilde{\mathcal{H}}^{\perp}, v \in \tilde{\mathcal{H}}.$$

The operator $\tilde{T} \equiv T|_{\tilde{\mathcal{H}}^{\perp}}$, the restriction of T to $\tilde{\mathcal{H}}^{\perp}$, is also compact and symmetric. In addition, $\sigma(\tilde{T}) = \{0\}$, since any nonzero eigenvalue of $\tilde{T}$ would be an eigenvalue of T as well. According to the previous lemma, $(\tilde{T}u, u)_{\mathcal{H}} = 0$ for all $u \in \tilde{H}^{\perp}$. But then if $u, v \in \tilde{H}^{\perp}$,

$$2(\tilde{T}u, v)_{\mathcal{H}} = (\tilde{T}(u+v), (u+v))_{\mathcal{H}} - (\tilde{T}u, u)_{\mathcal{H}} - (\tilde{T}v, v)_{\mathcal{H}} = 0$$

Hence $\tilde{T} = 0$ on $\tilde{\mathcal{H}}^{\perp}$. As a consequence, $\tilde{\mathcal{H}}^{\perp} \subseteq N(T) \subseteq \tilde{\mathcal{H}}$, so $\tilde{\mathcal{H}}^{\perp} = \{0\}$. Thus $\tilde{\mathcal{H}}$ is dense in $\mathcal{H}$.

An orthonormal basis of $\mathcal{H}$ then can be obtained by choosing an orthonormal basis for each subspace $\mathcal{H}_k$, $k = 0, 1, \cdots$. Note that the separability of $\mathcal{H}$ implies that $\mathcal{H}_0$ has a countable orthonormal basis, and these basis vectors are all eigenvectors corresponding to $\lambda_0 = 0$. $\square$

B.4 The Peetre-Tartar Theorem

The following theorem due to Peetre and Tartar can be used to derive various Poincaré type inequalities, and sometimes is useful to guarantee the existence of solutions to certain PDEs.

Theorem B.40 (Peetre-Tartar). *Let X, Y, Z be three Banach spaces, $A \in \mathscr{B}(X, Y)$ and K is a compact operator in $\mathscr{B}(X, Z)$ such that*

$$C_1 \|u\|_X \leqslant \|Au\|_Y + \|Ku\|_Z \leqslant C_2 \|u\|_X \qquad \forall\, u \in X \tag{B.2}$$

for some positive constants C_1 and C_2. Then

1. *The dimension of* $\mathrm{Ker}(A)$ *is finite, the mapping A is an isomorphism from $X/\mathrm{Ker}(A)$ on $R(A)$, and $R(A)$ is a closed subspace of Y. We recall that $R(A)$ is the range of A.*

2. *There exists a constant C_0 such that if F is a Banach space and $L_1 \in \mathscr{B}(X,F)$ which vanishes on* $\mathrm{Ker}(A)$, *then*

$$\|L_1 u\|_F \leqslant C_0 \|L_1\|_{\mathscr{B}(X,F)} \|Au\|_Y \qquad \forall\, u \in X\,. \tag{B.3}$$

3. *If G is a Banach space and $L_2 \in \mathscr{B}(X,G)$ satisfies*

$$L_2 u \neq 0 \qquad \forall\, u \in \mathrm{Ker}(A)\backslash\{0\}\,, \tag{B.4}$$

then

$$C_3 \|u\|_X \leqslant \|Au\|_Y + \|L_2 u\|_G \leqslant C_4 \|u\|_X \qquad \forall\, u \in X \tag{B.5}$$

for some positive constants C_3 and C_4.

Proof. 1. Because of (B.2), we find that

$$C_1 \|u\|_X \leqslant \|Ku\|_Z \leqslant C_2 \|u\|_X \qquad \forall\, u \in \mathrm{Ker}(A)\,. \tag{B.6}$$

Let $\{u_n\}_{n=1}^\infty$ be a bounded sequence in $\mathrm{Ker}(A) \subseteq X$. Since K is compact, there exists a subsequence $\{u_{n_k}\}_{k=1}^\infty$ such that $\{Ku_{n_k}\}_{k=1}^\infty$ converges in Z. Using (B.6) we find that $\{u_{n_k}\}_{k=1}^\infty$ converges in X. In other words, the identity map $\iota : \mathrm{Ker}(A) \to X$ is compact; thus (b) of Theorem B.22 implies that $\mathrm{Ker}(A)$ is finite dimensional.

Consider the quotient space $M = X/\mathrm{Ker}(A)$ which is a Banach space with quotient norm

$$\big\|[u]\big\|_M = \inf_{u \in [u]} \|u\|_X \qquad \forall\, [u] \in M \text{ or } u \in X\,.$$

We remark that the infimum above is in fact minimum since $\mathrm{Ker}(A)$ is finite dimensional. In the following, we let $\widetilde{u}$ denotes an element in X such that $\|[u]\|_M = \|\widetilde{u}\|_X$. Equip $R(A)$ with norm $\|\cdot\|_Y$. Then $(R(A), \|\cdot\|_Y)$ is a topological vector space. Since $A : M \to R(A)$ is bounded surjective, the open mapping theorem (Theorem B.14) implies that A is an open mapping; thus

$$\big\|[u]\big\|_M \leqslant C \|A[u]\|_Y \qquad \forall\, [u] \in M$$

which further implies that $R(A)$ is closed. In fact, if $\{A[u_n]\}_{n=1}^\infty$ is a convergent sequence in $R(A)$, then $\{[u_n]\}_{n=1}^\infty$ is Cauchy in M; thus $\{[u_n]\}_{n=1}^\infty$ converges to a limit $[u] \in M$ and $\{A[u_n]\}_{n=1}^\infty$ converges to $A[u]$ in Y.

Finally, the injectivity of A further suggests that

$$\|A^{-1}v\|_M \leqslant C\|AA^{-1}v\|_M = C\|v\|_Y \qquad \forall\, v \in R(A)\,.$$

Therefore, $A^{-1} \in \mathscr{B}(R(A), M)$.

2. Since L_1 vanishes on $\mathrm{Ker}(A)$, we find that

$$L_1 u = L_1 \widetilde{u} = L_1 A^{-1} A[u] \qquad \forall\, u \in X\,;$$

thus for all $u \in X$,

$$\begin{aligned}\|L_1 u\|_F &\leqslant \|L_1\|_{\mathscr{B}(X,F)} \|A^{-1}A[u]\|_X \leqslant \|L\|_{\mathscr{B}(X,F)} \|A^{-1}\|_{\mathscr{B}(R(A),M)} \|A[u]\|_Y \\ &\leqslant \|L\|_{\mathscr{B}(X,F)} \|A^{-1}\|_{\mathscr{B}(R(A),M)} \|Au\|_Y\end{aligned}$$

which concludes (B.3) by letting $C_0 = \|A^{-1}\|_{\mathscr{B}(R(A),M)}$.

3. Since $L_2 \in \mathscr{B}(X, G)$, it suffices to show that there exists $C > 0$ such that

$$\|u\|_X \leqslant C\big[\|Au\|_Y + \|L_2 u\|_G\big] \qquad \forall\, u \in X\,.$$

Suppose the contrary that there exists $\{u_n\}_{n=1}^{\infty}$ such that $\|u_n\|_X = 1$ while $\|Au_n\|_Y + \|L_2 u_n\|_G \leqslant \dfrac{1}{n}$ for all $n \in \mathbb{N}$. Since $K \in \mathscr{B}(X, Z)$ is compact, there exists a subsequence $\{u_{n_k}\}_{k=1}^{\infty}$ such that $\{Ku_{n_k}\}_{k=1}^{\infty}$ converges in Z. Moreover, $\{Au_n\}_{n=1}^{\infty}$ converges to 0; thus using (B.2) we find that $\{u_{n_k}\}_{k=1}^{\infty}$ is Cauchy in X. Suppose that $\lim\limits_{k\to\infty} u_{n_k} = u$. Then by the continuity of A and L_2, we must have $Au = L_2 u = 0$; thus by condition (B.4) we conclude that $u = 0$ which contradicts to that
$\|u\|_X = \lim\limits_{k\to\infty} \|u_{n_k}\|_X = 1$. □

Example B.41. Let Ω be a bounded domain, $E_1 = H^1(\Omega)$, $E_2 = E_3 = L^2(\Omega)$, A be the gradient operator, and K be the identity map. The Rellich theorem implies that the assumptions in the Peetre-Tartar theorem are valid.

1. The kernel of A is the collection of all constants; that is, $\mathrm{Ker}(A) = \mathbb{R}$. Therefore, 1 of the Peetre-Tartar theorem suggests that the gradient operator is an isomorphism from $H^1(\Omega)/\mathbb{R}$ to $L^2(\Omega)$. In other words, one has

$$\|u\|_{H^1(\Omega)} \leqslant C\|Du\|_{L^2(\Omega)} \qquad \forall\, u \in H^1(\Omega)/\mathbb{R}$$

which is the Poincaré inequality.

2. If $F = H^1(\Omega)/\mathbb{R}$, and L_1 is defined by

$$L_1 u = u - \frac{1}{|\Omega|}\int_\Omega u dx \,,$$

then 2 of the Peetre-Tartar theorem implies the Poincaré inequality (2.33).

3. Let $G = L^2(\partial\Omega)$ and $k \in L^\infty(\partial\Omega)$. If $k \neq 0$ on a portion of $\partial\Omega$, and define L_2 by

$$L_2 u = ku \qquad \forall\, u \in \mathrm{Ker}(A)\backslash\{0\}\,,$$

then the trace estimate implies that $L_2 \in \mathscr{B}(E_1, G)$; thus the use of part 3 of the Peetre-Tartar theorem leads to the Poincaré inequality (2.33).

Example B.42. Let $\Omega \subseteq \mathbb{R}^3$ be a bounded smooth domain, $E_1 = H^1(\Omega;\mathbb{R}^3)$, $E_2 = E_3 = L^2(\Omega;\mathbb{R}^3)$, A be the gradient operator, and K be the identity map. Recall that

1. Let $G = L^2(\partial\Omega)$, and L_2 be defined by

$$L_2 \boldsymbol{u} = \boldsymbol{u}\cdot\mathbf{N} \qquad \forall\, \boldsymbol{u} \in E_1\,.$$

Then L_2 clearly belongs to $\mathscr{B}(E_1, G)$ because of the trace estimate. Moreover, since Ω is bounded and smooth, $\mathbf{N} : \partial\Omega \to \mathbb{S}^1$ is onto; thus

$$L_2 \boldsymbol{u} \neq 0 \qquad \forall\, \boldsymbol{u} \in \mathrm{Ker}(A)\backslash\{\mathbf{0}\}\,.$$

Therefore, 3 of the Peetre-Tartar theorem implies that

$$\|\boldsymbol{u}\|_{H^1(\Omega)} \leqslant C\big[\|D\boldsymbol{u}\|_{L^2(\Omega)} + \|\boldsymbol{u}\cdot\mathbf{N}\|_{L^2(\partial\Omega)}\big] \qquad \forall\, \boldsymbol{u} \in H^1(\Omega)$$

which, in particular, implies the Poincaré inequality.

2. As in 2, letting $G = L^2(\partial\Omega;\mathbb{R}^3)$ and L_2 be defined by $L_2\boldsymbol{u} = \boldsymbol{u}\times\mathbf{N}$ can be used to conclude that

$$\|\boldsymbol{u}\|_{H^1(\Omega)} \leqslant C\big[\|D\boldsymbol{u}\|_{L^2(\Omega)} + \|\boldsymbol{u}\times\mathbf{N}\|_{L^2(\partial\Omega)}\big] \qquad \forall\, \boldsymbol{u} \in H^1(\Omega)$$

which further shows that the Poincaré inequality holds.

Appendix C

The Laplace and Poisson Equations

Definition C.1. On regions $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$, the Laplace operator Δ, also called the Laplacian, is defined as

$$\Delta = \sum_{i=1}^{\mathrm{n}} \frac{\partial^2}{\partial x_i^2}\,.$$

Definition C.2. A $\mathscr{C}^2$-function u is called a harmonic function if $\Delta u = 0$.

C.1 The Fundamental Solution

When $\Omega = \mathbb{R}^{\mathrm{n}}$, the Laplace operator has radial symmetry, and we may search for harmonic functions on $\mathbb{R}^{\mathrm{n}}$ which depend only upon the radial component. Letting

$$r = |x|\,,$$

we look for a harmonic function u satisfying

$$u(x) = v(r)\,.$$

Since $\dfrac{\partial r}{\partial x_i} = \dfrac{x_i}{r}$, by the chain rule,

$$\frac{\partial u}{\partial x_i} = \frac{dv}{dr}\frac{\partial r}{\partial x_i} = v'(r)\frac{x_i}{r}\,, \quad \frac{\partial^2 u}{\partial x_i^2} = \frac{\partial}{\partial x_i}\Big[v'(r)\frac{x_i}{r}\Big] = v''(r)\frac{x_i^2}{r^2} + v'(r)\Big[\frac{1}{r} - \frac{x_i^2}{r^3}\Big]\,.$$

Therefore,

$$\Delta u = \sum_{i=1}^{\mathrm{n}} \frac{\partial^2 u}{\partial x_i^2} = v''(r) + \frac{\mathrm{n}-1}{r}v'(r)\,.$$

Hence $\Delta u = 0$ if and only if

$$v'' + \frac{\mathrm{n}-1}{r} v' = 0 \,.$$

If we consider solutions away from $r = 0$ and suppose that $v'(r) \neq 0$, then

$$\big[\log v'(r)\big]' = \frac{\mathrm{n}-1}{r} \,.$$

This is a simple ordinary differential equation which we can directly integrate to find that for $r > 0$,

$$v(r) = \begin{cases} b \log r + c & (\mathrm{n} = 2) \\ \dfrac{b}{r^{\mathrm{n}-2}} + c & (\mathrm{n} \geqslant 3) \,, \end{cases}$$

where b and c are constants. These radially symmetric functions, harmonic away from the origin, provide us with the singular integral kernels on $\mathbb{R}^{\mathrm{n}}$ and explicit representations for the solutions to the Poisson equation $-\Delta u = f$, at least when the forcing function f is "nice" enough.

DEFINITION C.3. The function

$$\Phi(x) = \begin{cases} -\dfrac{1}{2\pi} \log |x| & (\mathrm{n} = 2) \\ \dfrac{1}{(\mathrm{n}-2)\omega_{\mathrm{n}-1} |x|^{\mathrm{n}-2}} & (\mathrm{n} \geqslant 3) \,, \end{cases}$$

defined for $x \in \mathbb{R}^{\mathrm{n}}$, $x \neq 0$, is the fundamental solution of Laplace's equation, where $\omega_{\mathrm{n}-1}$ denotes the surface area of the unit ball in $\mathbb{R}^{\mathrm{n}}$ and is defined by

$$\omega_{\mathrm{n}-1} = \frac{2\pi^{\mathrm{n}/2}}{\Gamma(\mathrm{n}/2)} \,,$$

where Γ denotes the Gamma function.

Notation. We will use that notation $F_{,i}$ to denote $\dfrac{\partial F}{\partial x_i}$, while $F_{,ij}$ denotes $\dfrac{\partial^2 F}{\partial x_i \partial x_j}$ and similarly for higher-order partial derivatives.

A direct computation shows that

$$\Phi_{,i}(x) = \frac{-x_i}{\omega_{\mathrm{n}-1}} |x|^{-\mathrm{n}} \,,$$

$$\Phi_{,ij}(x) = \frac{1}{\omega_{\mathrm{n}-1}} \Big[-|x|^2 \delta_{ij} + \mathrm{n} x_i x_j \Big] |x|^{-\mathrm{n}-2} \,,$$

and we have the following derivative estimates:

$$|\Phi_{,i}(x)| \leqslant \frac{1}{\omega_{\mathrm{n}-1}}|x|^{1-\mathrm{n}}\,, \tag{C.1a}$$

$$|\Phi_{,ij}(x)| \leqslant \frac{\mathrm{n}}{\omega_{\mathrm{n}-1}}|x|^{-\mathrm{n}}\,, \tag{C.1b}$$

$$|D^\alpha\Phi(x)| \leqslant C(\mathrm{n},|\alpha|)|x|^{2-\mathrm{n}-|\alpha|}\,, \tag{C.1c}$$

where D^α and $|\alpha|$ are the multi-index notation defined by the following

Definition C.4 (Multi-index)**.** An n-dimensional multi-index is a vector $\alpha = (\alpha_1, \cdots, \alpha_\mathrm{n})$ of non-negative integers. $|\alpha|$ is defined as the sum of α_k and $\alpha!$ is defined as the product of $\alpha_k!$, i.e.,

$$|\alpha| = \sum_{k=1}^{\mathrm{n}} \alpha_k \quad \text{and} \quad \alpha! = \prod_{k=1}^{\mathrm{n}} \alpha_k!\,.$$

The differential operator D_x^α is defined by

$$D_x^\alpha = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_\mathrm{n}}}{\partial x_\mathrm{n}^{\alpha_\mathrm{n}}}\,.$$

When the (spatial) variable is specified, we simply use D^α to denote D_x^α.

C.1.1 Uniform and Hölder continuous functions

For $\Omega \subseteq \mathbb{R}^\mathrm{n}$ open, a function $u : \Omega \to \mathbb{R}$ is *Lipschitz continuous* if

$$|u(x) - u(y)| \leqslant C|x-y| \quad \forall \;\; x, y \in \Omega\,, \tag{C.2}$$

where C is a constant that depends on Ω but not on the function u itself. The inequality (C.2) provides a uniform modulus of continuity. The standard example of functions which are Lipschitz continuous but not differentiable is given by $u(x) = |x|$. It is interesting to refine this functional framework to be able to discern the regularity of functions $u(x) = |x|^\alpha$ for positive $\alpha \leqslant 1$. We wish to understand how "cuspy" the graph of u is near the origin, for example. To do so, we replace the difference quotient bound in (C.2) with the following inequality:

$$|u(x) - u(y)| \leqslant C|x-y|^\alpha \qquad \forall\, x, y \in \Omega\,. \tag{C.3}$$

Functions which satisfy the inequality (C.3) are termed *Hölder continuous with exponent* α.

DEFINITION C.5 (Continuous functions and compact support)**.** For $\Omega \subseteq \mathbb{R}^n$, we let $\mathscr{C}^0(\Omega)$ denote the collection of continuous functions on Ω, and we denote by $\mathscr{C}^0_c(\Omega)$ the collection of those functions in $\mathscr{C}^0(\Omega)$ with compact support contained in Ω.

DEFINITION C.6 (Bounded continuous functions)**.** For $\Omega \subseteq \mathbb{R}^n$ we set

$$\mathscr{C}^0(\bar{\Omega}) := \left\{ u : \Omega \to \mathbb{R} \,\middle|\, u \text{ is bounded and continuous} \right\},$$

with norm $\|u\|_{\mathscr{C}^0(\bar{\Omega})} = \max_{x \in \bar{\Omega}} |u(x)|$. For integers $k \geqslant 0$, we let $\mathscr{C}^k(\bar{\Omega})$ denote the collection of functions possessing partial derivatives to all orders up to k which are bounded and continuous on $\bar{\Omega}$. We use $\mathscr{C}^k_{\mathrm{loc}}(\Omega)$ to denote the functions in $\mathscr{C}^k(\bar{B})$ for all bounded balls $\bar{B}$ contained in Ω.

DEFINITION C.7 (Hölder continuous functions)**.** For $0 < \alpha \leqslant 1$, we set

$$\mathscr{C}^{0,\alpha}(\bar{\Omega}) := \left\{ u \in \mathscr{C}^0(\bar{\Omega}) \,\middle|\, \|u\|_{\mathscr{C}^0(\bar{\Omega})} + [u]_{\mathscr{C}^{0,\alpha}(\bar{\Omega})} < \infty \right\},$$

where

$$[u]_{\mathscr{C}^{0,\alpha}(\bar{\Omega})} = \sup_{\substack{x,y \in \Omega \\ x \neq y}} \frac{|u(x) - u(y)|}{|x - y|^\alpha}.$$

The norm of u in $\mathscr{C}^{0,\alpha}(\bar{\Omega})$ is $\|u\|_{\mathscr{C}^{0,\alpha}(\bar{\Omega})} = \|u\|_{\mathscr{C}^0(\bar{\Omega})} + [u]_{\mathscr{C}^{0,\alpha}(\bar{\Omega})}$.

THEOREM C.8. *The space $\mathscr{C}^{0,\alpha}(\bar{\Omega})$ endowed with the norm $\|\cdot\|_{\mathscr{C}^{0,\alpha}(\bar{\Omega})}$ is a Banach space.*

We leave the proof as an exercise for the reader.

We will denote $\mathscr{C}^{0,\alpha}_c(\bar{\Omega}) = \mathscr{C}^{0,\alpha}(\bar{\Omega}) \cap \mathscr{C}^0_c(\Omega)$.

C.1.2 The Poisson equation $-\Delta u = f$ in $\mathbb{R}^n$

Our objective, here, is to produce explicit solutions to the Poisson equation $-\Delta u = f$ in $\mathbb{R}^n$. We will show that convolution between the fundamental solution Φ and the "forcing function" f is a solution to this problem.

DEFINITION C.9. We set

$$u(x) = \int_{\mathbb{R}^n} \Phi(x - y) f(y)\, dy \tag{C.4}$$

whenever $\Phi(x - \cdot) f(\cdot) \in L^1(\mathbb{R}^n)$.

Lemma C.10. *Suppose that f is bounded and integrable with compact support. Then if u is given by* (C.4), $u \in \mathscr{C}^1(\mathbb{R}^{\mathrm{n}})$ *and for any* $x \in \mathbb{R}^{\mathrm{n}}$ *and* $i = 1, ..., \mathrm{n}$,

$$u,_i(x) = \int_{\mathbb{R}^{\mathrm{n}}} \Phi,_i(x-y) f(y)\, dy\,.$$

Proof. Since f is bounded with compact support, the integral

$$\mathrm{I}_i(x) \equiv \int_{\mathbb{R}^{\mathrm{n}}} \Phi,_i(x-y) f(y)\, dy$$

is well-defined and continuous (in x). It suffices to show that it is the derivative of $u(x) \equiv \displaystyle\int_{\mathbb{R}^{\mathrm{n}}} \Phi(x-y) f(y)\, dy\,.$

Let $\rho : (0,\infty) \to \mathbb{R}$ be a smooth, monotone increasing function such that

$$\rho(z) = \begin{cases} 1 & \text{if } \ z \in (2,\infty)\,, \\ 0 & \text{if } \ z \in (0,1)\,, \end{cases} \qquad \text{and} \quad |\rho'| \leqslant 2\,,$$

and define

$$u_\epsilon(x) = \int_{\mathbb{R}^{\mathrm{n}}} \rho\Big(\frac{|x-y|}{\epsilon}\Big) \Phi(x-y) f(y)\, dy\,.$$

Note that since ρ is uniformly bounded and $\Phi \in L^1(\mathbb{R}^{\mathrm{n}})$, by the dominated convergence theorem, $u_\epsilon \to u$ uniformly as $\epsilon \to 0$ on compact subsets. Furthermore, as $D_x\Big[\rho\big(\frac{|x-y|}{\epsilon}\big)\Phi(x-y)\Big] f(y)$ is also integrable for $\epsilon > 0$, we may differentiate under the integral to find that

$$\begin{aligned} u_{\epsilon},_i(x) &= \int_{\mathbb{R}^{\mathrm{n}}} \frac{\partial}{\partial x_i}\Big[\rho\Big(\frac{|x-y|}{\epsilon}\Big)\Phi(x-y)\Big] f(y)\, dy \\ &= \int_{\mathbb{R}^{\mathrm{n}}} \frac{x_i - y_i}{\epsilon|x-y|} \rho'\Big(\frac{|x-y|}{\epsilon}\Big)\Phi(x-y) f(y)\, dy \\ &\quad + \int_{\mathbb{R}^{\mathrm{n}}} \rho\Big(\frac{|x-y|}{\epsilon}\Big)\Phi,_i(x-y) f(y)\, dy\,. \end{aligned}$$

Therefore,

$$\begin{aligned} |u_{\epsilon},_i(x) - \mathrm{I}_i(x)| \leqslant{}& \int_{\mathbb{R}^{\mathrm{n}}} \Big|\frac{x_i - y_i}{\epsilon|x-y|}\rho'\Big(\frac{|x-y|}{\epsilon}\Big)\Big| \big|\Phi(x-y)\big|\big|f(y)\big| dy \\ &+ \int_{\mathbb{R}^{\mathrm{n}}} \Big|1 - \rho\Big(\frac{|x-y|}{\epsilon}\Big)\Big|\big|\Phi,_i(x-y)\big|\big|f(y)\big| dy\,. \end{aligned}$$

Note that $|\rho| \leqslant 1$ and $|\rho'| \leqslant 2$. Moreover, since $\mathrm{spt}(1-\rho) \subseteq [0,2]$, it follows that $\mathrm{spt}\big(1 - \rho\big(\frac{\cdot}{\epsilon}\big)\big) \subseteq [0, 2\epsilon]$. Similarly, since $\mathrm{spt}(\rho') \subseteq [1,2]$, we see that $\mathrm{spt}\big(\rho'\big(\frac{\cdot}{\epsilon}\big)\big) \subseteq$

$[\epsilon, 2\epsilon]$. As a consequence, for the case $\mathrm{n} \geqslant 3$ (the case $\mathrm{n} = 2$ is left as an exercise for the reader),

$$
\begin{aligned}
&|u_{\epsilon,i}(x) - \mathrm{I}_i(x)| \\
&\quad \leqslant \frac{2}{\epsilon}\int_{\epsilon<|x-y|<2\epsilon} \frac{\|f\|_{L^\infty(\Omega)}}{\omega_{\mathrm{n}-1}}|x-y|^{-n+2}dy + \int_{|x-y|<2\epsilon} \frac{\|f\|_{L^\infty(\Omega)}}{(\mathrm{n}-2)\omega_{\mathrm{n}-1}}|x-y|^{1-n}dy \\
&\quad \leqslant \frac{\|f\|_{L^\infty(\Omega)}}{\omega_{n-1}}\Big[\frac{2}{\epsilon}\int_{|z|=1}\int_{\epsilon}^{2\epsilon} r dr dS_z + \int_{|z|=1}\int_0^{2\epsilon} dr dS_z\Big] \\
&\quad = \|f\|_{L^\infty(\Omega)}\Big[\frac{2}{\epsilon}\int_{\epsilon}^{2\epsilon} r dr + \int_0^{2\epsilon} dr\Big] \to 0 \quad \text{as } \epsilon \to 0\,;
\end{aligned}
$$

hence $u_{\epsilon,i} \to \mathrm{I}_i$ uniformly as $\epsilon \to 0$.

Finally, by the uniform convergence of u_ϵ to u and $u_{\epsilon,i}$ to I_i as $\epsilon \to 0$, we conclude that

$$
u(x) = \lim_{\epsilon\to 0} u_\epsilon(x) = \lim_{\epsilon\to 0}\Big[u_\epsilon(x_0) + \int_{x_0}^{x} \frac{\partial u_\epsilon}{\partial x_i}dx_i\Big] = u(x_0) + \int_{x_0}^{x} \mathrm{I}_i dx_i\,;
$$

thus $u_{,i} = \mathrm{I}_i$. □

Remark C.11. Given that $D\Phi$ is integrable near the origin, it is possible to compute the first partial derivatives of u by taking the limit of a sequence of difference quotients of u. On the other hand, since $D^2\Phi$ is not integrable near the origin, analysis of second partial derivatives of u require some sort of limiting process, wherein the singular behavior at $|x| = 0$ is either excised or regularized.

For example, we might consider removing a small ball near the origin, and defining an approximation to u as follows:

$$
\tilde{u}_\epsilon(x) \equiv \int_{\mathbb{R}^{\mathrm{n}}\setminus B(x,\epsilon)} \Phi(x-y)f(y)\,dy\,,
$$

which makes $\tilde{u}_\epsilon$ a differentiable function. However, as the domain of integration also depends upon x, differentiation of $\tilde{u}_\epsilon$ becomes a bit complicated, requiring a change of variables. To avoid this procedure, one alternative is the introduction of the cut-off function ρ (introduced in the above proof), which has the similar affect of removing the singular region, without any difficulties in differentiation.

Theorem C.12. *Suppose that $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$, $f \in \mathscr{C}_c^{0,\alpha}(\bar{\Omega})$ with $0 < \alpha \leqslant 1$, and suppose that u is given by* (C.4). *Then $u \in \mathscr{C}^2(\mathbb{R}^{\mathrm{n}})$, and for any $x \in \mathbb{R}^{\mathrm{n}}$, and $i, j = 1, ..., \mathrm{n}$,*

$$
u_{,ij}(x) = \int_{\Omega_0} \Phi_{,ij}(x-y)\big(f(y) - f(x)\big)dy - f(x)\int_{\partial\Omega_0} \Phi_{,i}(x-y)\mathrm{N}_j\,dS_y\,, \tag{C.5}
$$

where Ω_0 *is any bounded, smooth domain containing* Ω. *In particular,*

$$-\Delta u = f \quad in \quad \mathbb{R}^{\rm n}. \tag{C.6}$$

Remark C.13. Before starting the proof of Lemma C.12, we explain why a formula like (C.5) is well-defined. Note that as $D^\alpha\Phi$ is not integrable for $|\alpha| \geqslant 2$, $\Phi,_{ij}(x-\cdot)f(\cdot) \notin L^1(\mathbb{R}^{\rm n})$ even if $f \in \mathscr{C}_c^\infty(\mathbb{R}^{\rm n})$. Nevertheless, the integral $\displaystyle\int_{\mathbb{R}^{\rm n}} \Phi,_{ij}(x-y)\big(f(y)-f(x)\big)dy$ is well-defined for all $x \in \mathbb{R}^{\rm n}$ due to the Hölder continuity of f. In particular, it is essential that second-derivatives of Φ are multiplying the difference $[f(y)-f(x)]$ – the Hölder continuity of f "cancels" the singular nature of $D^2\Phi$ near the origin, at least enough so that the integral converges.

The presence of the boundary integral on the right-hand side of (C.5) is necessary in order to cancel the effect of the subtraction of $f(x)$ from $f(y)$.

Proof of Theorem C.12. To see that (C.6) follows from (C.5), notice that $\Delta u = u,_{ii}$ and that according to (C.5),

$$u,_{ii}(x) = \int_{\Omega_0} \Phi,_{ii}(x-y)\big(f(y)-f(x)\big)dy - f(x)\int_{\partial\Omega_0} \Phi,_i(x-y){\rm N}_i dS_y\,.$$

Since $\Phi,_{ii}(x-y) = 0$ if $x \neq y$, and $\Phi,_{ii}(x-\cdot)\big(f(\cdot)-f(x)\big) \in L^1(\mathbb{R}^{\rm n})$, the first integral on the right-hand side vanishes. Thus

$$u,_{ii}(x) = -f(x)\int_{\partial\Omega_0} \Phi,_i(x-y){\rm N}_i dS_y\,. \tag{C.7}$$

Choose $R > 0$ sufficiently large (for example, $R = \operatorname{diam}(\Omega)+1$) so that $\Omega_0 \subset\subset B(x,R)$, and let ${\rm A} = B(x,R) - \Omega_0$. Since $\Delta\Phi(x-\cdot) = 0$ in A, the divergence theorem implies that

$$0 = \int_{\rm A} \Delta\Phi(x-y)\,dx = \int_{\partial B(x,R)} D\Phi(x-y)\cdot n(y)\,dS_y + \int_{\partial\Omega_0} D\Phi(x-y)\cdot n(y)\,dS_y\,,$$

where n denotes the outward-pointing unit normal to $\partial{\rm A}$. Since $n = -{\rm N}$ on $\partial\Omega_0$, substitution into (C.7) shows that

$$\begin{aligned} u,_{ii}(x) &= -f(x)\int_{\partial B(x,R)} \Phi,_i(x-y){\rm N}_i dS_y \\ &= -f(x)\int_{\partial B(x,R)} \frac{-(x_i-y_i)}{\omega_{\rm n-1}}|x-y|^{-{\rm n}}\frac{y_i-x_i}{R}\,dS_y \\ &= -f(x)\frac{1}{\omega_{\rm n-1}}\int_{\partial B(x,R)} R^{1-{\rm n}} dS_y = -f(x)\,. \end{aligned}$$

Next, we establish (C.5). Following the proof of Lemma C.10, we define

$$v^i_\epsilon(x) = \int_{\mathbb{R}^n} \rho\Big(\frac{|x-y|}{\epsilon}\Big)\Phi,_i(x-y)f(y)\,dy\,.$$

The derivative of this integrand has an $L^1(\mathbb{R}^n)$ dominating function, so the dominated convergence theorem allows to differentiate under the integral. We thus find that

$$\begin{aligned}
v^i_\epsilon,_j(x) =& \int_{\Omega_0}\Big[\frac{\partial}{\partial x_j}\rho\Big(\frac{|x-y|}{\epsilon}\Big)\Big]\Phi,_i(x-y)f(y)\,dy + \int_{\Omega_0}\rho\Big(\frac{|x-y|}{\epsilon}\Big)\Phi,_{ij}(x-y)f(y)\,dy \\
=& \int_{\Omega_0}\Big[\frac{\partial}{\partial x_j}\rho\Big(\frac{|x-y|}{\epsilon}\Big)\Big]\Phi,_i(x-y)\big(f(y)-f(x)\big)dy \\
&- f(x)\int_{\Omega_0}\Big[\frac{\partial}{\partial y_j}\rho\Big(\frac{|x-y|}{\epsilon}\Big)\Big]\Phi,_i(x-y)\,dy \\
&+ \int_{\Omega_0}\rho\Big(\frac{|x-y|}{\epsilon}\Big)\Phi,_{ij}(x-y)\big(f(y)-f(x)\big)dy \\
&+ f(x)\int_{\Omega_0}\rho\Big(\frac{|x-y|}{\epsilon}\Big)\Phi,_{ij}(x-y)\,dy\,.
\end{aligned}$$

We will show that $v^i_\epsilon,_j$ converges to the right-hand side of (C.5) uniformly. Since $\Phi,_{ij}(x-y) = -\dfrac{\partial}{\partial y_j}\Phi,_i(x-y)$, integration by parts shows that

$$\begin{aligned}
&\int_{\Omega_0}\rho\Big(\frac{|x-y|}{\epsilon}\Big)\Phi,_{ij}(x-y)\,dy \\
&\qquad = \int_{\Omega_0}\Big[\frac{\partial}{\partial y_j}\rho\Big(\frac{|x-y|}{\epsilon}\Big)\Big]\Phi,_i(x-y)\,dy - \int_{\partial\Omega_0}\rho\Big(\frac{|x-y|}{\epsilon}\Big)\Phi,_i(x-y)\mathrm{N}_j dS_y\,,
\end{aligned}$$

and hence that

$$\begin{aligned}
v^i_\epsilon,_j(x) =& \int_{\Omega_0}\Big[\frac{\partial}{\partial x_j}\rho\Big(\frac{|x-y|}{\epsilon}\Big)\Big]\Phi,_i(x-y)\big(f(y)-f(x)\big)dy \\
&+ \int_{\Omega_0}\rho\Big(\frac{|x-y|}{\epsilon}\Big)\Phi,_{ij}(x-y)\big(f(y)-f(x)\big)dy \\
&- f(x)\int_{\partial\Omega_0}\rho\Big(\frac{|x-y|}{\epsilon}\Big)\Phi,_i(x-y)\mathrm{N}_j dS_y\,.
\end{aligned}$$

Following the proof of Lemma C.10, by the Hölder continuity of f,

$$\begin{aligned}
&\Big|\int_{\Omega_0}\Big[\frac{\partial}{\partial x_j}\rho\Big(\frac{|x-y|}{\epsilon}\Big)\Big]\Phi_{,i}(x-y)\big(f(y)-f(x)\big)dy\Big| \\
&\qquad\leqslant \int_{B(x,R)}\Big|\frac{\partial}{\partial x_j}\rho\Big(\frac{|x-y|}{\epsilon}\Big)\Big|\big|\Phi_{,i}(x-y)\big|\big|f(y)-f(x)\big|dy \\
&\qquad\leqslant \int_{\epsilon<|x-y|<2\epsilon}\Big|\frac{x_i-y_i}{\epsilon|x-y|}\rho'\Big(\frac{|x-y|}{\epsilon}\Big)\Big|\big|\Phi_{,i}(x-y)\big|\big|f(y)-f(x)\big|dy \\
&\qquad\leqslant \int_{\epsilon<|x-y|<2\epsilon}\frac{2[f]_{\mathscr{C}^{0,\alpha}(\overline{\Omega})}}{\epsilon\omega_{n-1}}|x-y|^{1-\mathrm{n}+\alpha}dy \\
&\qquad= \frac{2[f]_{\mathscr{C}^{0,\alpha}(\overline{\Omega})}}{\epsilon\omega_{\mathrm{n}-1}}\int_{|z|=1}\int_{\epsilon}^{2\epsilon} r^{\alpha}drdS_z \to 0 \quad \text{as } \epsilon\to 0\,.
\end{aligned}$$

Here we note that since R can be chosen independent of $x\in\Omega$, the convergence above is in fact uniform in x. Similarly, by (C.1b),

$$\begin{aligned}
&\int_{\Omega_0}\Big|1-\rho\Big(\frac{|x-y|}{\epsilon}\Big)\Big|\Phi_{,ij}(x-y)\big(f(y)-f(x)\big)dy \\
&\qquad\qquad\leqslant \frac{\mathrm{n}[f]_{\mathscr{C}^{0,\alpha}(\overline{\Omega})}}{\omega_{\mathrm{n}-1}}\int_{|z|=1}\int_0^{2\epsilon} r^{\alpha-1}drdS_z \to 0 \quad \text{as } \epsilon\to 0\,;
\end{aligned}$$

and again the convergence above is uniform in $x\in\Omega$. Consequently, $v^i_{\epsilon,j}$ converges to the right-hand side of (C.5) uniformly as $\epsilon\to 0$.

It remains to show that $v^i_{\epsilon,j}\to u_{,ij}$ as $\epsilon\to 0$. From Lemma C.10, $v^i_\epsilon\to u_{,i}$ uniformly as $\epsilon\to 0$, so using the fundamental theorem of calculus (as in the proof of Lemma C.10), we indeed see that $u_{,ij}$ must be equal to the right-hand side of (C.5). □

C.2 A Representation Formula

For a point $y\in\Omega$, the function $\Phi(x-y)$ is harmonic if $x\neq y$. Therefore, letting $v(x)=\Phi(x-y)$ in Green's second identity (2.2),

$$\int_{\Omega_k}\Phi(x-y)\Delta u(x)\,dx=\int_{\partial\Omega_k}\Phi(x-y)\frac{\partial u}{\partial \mathrm{N}}(x)\,dS_x-\int_{\partial\Omega_k}u(x)\frac{\partial\Phi}{\partial\mathrm{N}}(x-y)\,dS_x$$

where $\Omega_k=\Omega\backslash B(y,\frac{1}{k})$. For k big enough, $\partial\Omega_k=\partial\Omega\cup\partial B(y,\frac{1}{k})$. Therefore,

$$\begin{aligned}
\int_{\partial B(y,\frac{1}{k})}u(x)\frac{\partial\Phi}{\partial\mathrm{N}}(x-y)\,dS_x&=\int_{\partial\Omega}\Phi(x-y)\frac{\partial u}{\partial\mathrm{N}}(x)\,dS_x+\int_{\partial B(y,\frac{1}{k})}\Phi(x-y)\frac{\partial u}{\partial\mathrm{N}}(x)\,dS_x \\
&\quad-\int_{\partial\Omega}u(x)\frac{\partial\Phi}{\partial\mathrm{N}}(x-y)\,dS_x-\int_{\Omega\backslash B(y,\frac{1}{k})}\Phi(x-y)\Delta u(x)\,dx\,,
\end{aligned}$$

here we have to note that the unit normal on $\partial B(y, \frac{1}{k})$ points to the center y. For $x \in \partial B(y, \frac{1}{k})$,

$$\Phi(x-y) = \begin{cases} \dfrac{1}{2\pi} \log k & \text{if n} = 2 \\ \dfrac{k^{\mathrm{n}-2}}{(n-2)\omega_{\mathrm{n}-1}} & \text{if n} \geqslant 3 \end{cases},$$

$$\frac{\partial \Phi}{\partial \mathrm{N}}(x-y) = \nabla\Phi(x-y) \cdot k(y-x) = \frac{k^{\mathrm{n}-1}}{\omega_{\mathrm{n}-1}}.$$

Therefore, as $k \to \infty$,

$$\int_{\partial B(y,\frac{1}{k})} u(x) \frac{\partial \Phi}{\partial \mathrm{N}}(x-y)\, dS_x \to u(y) \quad \text{and} \quad \int_{\partial B(y,\frac{1}{k})} \Phi(x-y) \frac{\partial u}{\partial \mathrm{N}}(x)\, dS_x \to 0.$$

Moreover, $\displaystyle\int_{\Omega_k} \Phi(x-y)\Delta u(x)\, dx \to \int_{\Omega} \Phi(x-y)\Delta u(x)\, dx$ as $k \to \infty$. As a consequence, $u \in \mathscr{C}^2(\overline{\Omega})$ satisfies

$$u(x) = \int_{\partial\Omega} \Phi(y-x) \frac{\partial u}{\partial \mathrm{N}}(y)\, dS_y - \int_{\partial\Omega} u(y) \frac{\partial \Phi}{\partial \mathrm{N}}(y-x)\, dS_y - \int_{\Omega} \Phi(y-x)\Delta u(y)\, dy. \quad \text{(C.8)}$$

REMARK C.14. The integral $\displaystyle\int_{\Omega} \Phi(x-y) f(y)\, dy$ is called the *Newtonian potential* with density f.

Given the formula (C.8) it is tempting to believe that the equation

$$\begin{aligned} -\Delta u &= f && \text{in} \quad \Omega, \\ u &= g && \text{on} \quad \partial\Omega, \\ \frac{\partial u}{\partial \mathrm{N}} &= h && \text{on} \quad \partial\Omega \end{aligned}$$

has a solution

$$u(x) = \int_{\partial\Omega} \Phi(x-y) h(y)\, dS_y - \int_{\partial\Omega} g(y) \frac{\partial \Phi}{\partial \mathrm{N}}(y-x)\, dS_y + \int_{\Omega} \Phi(y-x) f(y)\, dy.$$

This is, in fact, not the case, and we shall examine this in great detail below.

C.3 Properties of Harmonic Functions

C.3.1 Mean-value property

Now we consider harmonic functions on an open set $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$.

Theorem C.15. *If $u \in \mathscr{C}^2(\Omega)$ is harmonic, then*

$$u(x) = \fint_{\partial B(x,r)} u(y)\, dS_y = \fint_{B(x,r)} u(y)\, dy \tag{C.9}$$

for each ball $B(x,r) \subseteq \Omega$.

Proof. We begin by proving the first equality in (C.9). By the divergence theorem,

$$\begin{aligned} 0 &= \int_{\partial B(x,r)} \frac{\partial u}{\partial \mathrm{N}}(y)\, dS_y = \int_{\partial B(0,1)} r^{\mathrm{n}-1} \frac{\partial}{\partial r} u(x+rw)\, dS_w \\ &= r^{\mathrm{n}-1} \frac{\partial}{\partial r} \int_{\partial B(0,1)} u(x+rw)\, dS_w = r^{\mathrm{n}-1} \frac{\partial}{\partial r} \Big[r^{1-\mathrm{n}} \int_{\partial B(x,r)} u(y)\, dS_y \Big]. \end{aligned}$$

Therefore,

$$r^{1-\mathrm{n}} \int_{\partial B(x,r)} u(y)\, dS_y = \lim_{r\to 0^+} r^{1-\mathrm{n}} \int_{\partial B(x,r)} u(y)\, dS_y = \omega_{\mathrm{n}-1} u(x)\,. \tag{C.10}$$

The second equality in (C.9) is obtained by integrating (C.10) in r. □

Exercise: If $u \in \mathscr{C}^2(\Omega)$ satisfying $\Delta u \geqslant 0$ (or $\Delta u \leqslant 0$), show that

$$\begin{aligned} u(x) &\leqslant \fint_{\partial B(x,r)} u(y)\, dS_y \quad \Big(\text{or} \quad u(x) \geqslant \fint_{\partial B(x,r)} u(y)\, dS_y\Big), \\ u(x) &\leqslant \fint_{B(x,r)} u(y)\, dy \qquad \Big(\text{or} \quad u(x) \geqslant \fint_{B(x,r)} u(y)\, dy\Big). \end{aligned}$$

A function $u \in \mathscr{C}^2(\Omega)$ is called a sub-harmonic/super-harmonic function if $\Delta u \geqslant 0/\leqslant 0$.

Theorem C.16 (Converse to mean-value property)**.** *If $u \in \mathscr{C}^2(\Omega)$ satisfies* (C.9) *for each $B(x,r) \subseteq \Omega$, then u is harmonic.*

Proof. If $\Delta u \neq 0$, there exists some ball $B(x,r) \subseteq \Omega$ on which $\Delta u > 0$ (or perhaps $\Delta u < 0$) in $B(x,r)$. But this would imply that $0 = \fint_{B(x,r)} \Delta u(y)\, dy > 0$, a contradiction. □

Theorem C.17. *If $u \in L^1_{\mathrm{loc}}(\Omega)$ satisfies the mean-value property* (C.9) *for each $B(x,r) \subseteq \Omega$, then $u \in \mathscr{C}^\infty(\Omega)$.*

Proof. Let η be the standard mollifier defined in Definition 1.38, and let $u^\epsilon = \eta_\epsilon * u$ in Ω_ϵ. Then

$$\begin{aligned} u^\epsilon(x) &= \int_\Omega \eta_\epsilon(x-y)u(y)\,dy = \frac{1}{\epsilon^{\mathrm{n}}}\int_{B(x,\epsilon)} \eta\Big(\frac{x-y}{\epsilon}\Big)u(y)\,dy \\ &= \frac{1}{\epsilon^{\mathrm{n}}}\int_0^\epsilon \int_{\partial B(0,1)} \eta(\frac{rw}{\epsilon})u(x+rw)r^{\mathrm{n}-1}dS_w dr \\ &= \frac{1}{\epsilon^{\mathrm{n}}}\omega_{\mathrm{n}-1}u(x)\int_0^\epsilon \eta(\frac{rw}{\epsilon})r^{\mathrm{n}-1}dr = u(x)\,. \end{aligned}$$

Therefore, $u^\epsilon = u$ in Ω_ϵ. □

COROLLARY C.18. *A harmonic function is a $\mathscr{C}^\infty$-function.*

COROLLARY C.19. *The limit of a uniformly convergent sequence of harmonic functions is harmonic.*

C.3.2 Maximum principles

THEOREM C.20 (Strong maximum principle)**.** *Suppose that $u \in \mathscr{C}^2(\Omega) \cap \mathscr{C}^0(\overline{\Omega})$ is harmonic within a bounded domain Ω.*

(1) *Then $\max\limits_{\overline{\Omega}} u = \max\limits_{\partial\Omega} u$.*

(2) *Furthermore, if Ω is connected and there exists a point $x_0 \in \Omega$ such that $u(x_0) = \max\limits_{\overline{\Omega}} u$, then $u = u(x_0)$ within Ω.*

Proof. Let $M = \max\limits_{\overline{\Omega}} u$ (the boundedness of Ω implies that such an M exists), and $A = \{x \in \Omega \mid u(x) = M\}$. The continuity of u implies that A is closed (relative to Ω). If A is empty, then the maximum of u is attained on $\partial\Omega$, so we may assume that A is not empty.

Let $x \in A$, and $r > 0$ be such that $B(x,r) \subseteq \Omega$. According to (C.9)

$$M = u(x) = ⨍_{B(x,r)} u(y)\,dy \leqslant M\,.$$

Equality can only hold if $u(y) = u(x)$ for all $y \in B(x,r)$; thus, $B(x,r) \subseteq A$, and hence A is also open in Ω. It must be that $A = \Omega$ since it is the only subset of Ω which is both open and closed in Ω. □

Remark C.21. In the statement of Theorem C.20, it suffices to assume that $\Delta u \geqslant 0$.

Exercise: Let u be sub-harmonic/super-harmonic in Ω. Then

(1) Then $\max\limits_{\overline{\Omega}} u = \max\limits_{\partial\Omega} u / \min\limits_{\overline{\Omega}} u = \min\limits_{\partial\Omega} u$.

(2) If Ω is connected and there exists a point $x_0 \in \Omega$ such that $u(x_0) = \max\limits_{\overline{\Omega}} u / u(x_0) = \min\limits_{\overline{\Omega}} u$, then $u = u(x_0)$ within Ω.

Corollary C.22 (Uniqueness). *Let $g \in \mathscr{C}^0(\partial\Omega)$, $f \in \mathscr{C}^0(\Omega)$. Then there exists at most one solution $u \in \mathscr{C}^2(\Omega) \cap \mathscr{C}^0(\overline{\Omega})$ to the boundary-value problem*

$$\begin{aligned} -\Delta u &= f \qquad \textit{in} \quad \Omega\,, \\ u &= g \qquad \textit{on} \quad \partial\Omega\,. \end{aligned}$$

Corollary C.23 (Comparison). *For $u, v \in \mathscr{C}^2(\Omega) \cap \mathscr{C}^0(\bar{\Omega})$, if $\Delta u \geqslant 0$, $\Delta v = 0$ in Ω, and $u = v$ on $\partial\Omega$, then $u \leqslant v$ in Ω. (For this reason, we call u subharmonic.*

C.3.3 The Harnack inequality

Theorem C.24. *Let u be a non-negative harmonic function in Ω. Then for any bounded sub-domain $\Omega' \subset\subset \Omega$, there exists a constant C depending only on n, Ω' and Ω such that*

$$\max_{\Omega'} u \leqslant C \min_{\Omega'} u\,.$$

Proof. Let $y \in \Omega$, $B(y, 4R) \subseteq \Omega$. Then for any two points x_1, $x_2 \in B(y, R)$, we have

$$\begin{aligned} u(x_1) &= \fint_{B(x_1,R)} u(x)\,dx \leqslant \frac{1}{|B(0,R)|} \int_{B(x_1,2R)} u(x)\,dx\,, \\ u(x_2) &= \fint_{B(x_2,3R)} u(x)\,dx \geqslant \frac{1}{|B(0,3R)|} \int_{B(x_1,2R)} u(x)\,dx\,. \end{aligned}$$

Consequently, $\max\limits_{B(y,R)} u \leqslant 3^{\mathrm{n}} \min\limits_{B(y,R)} u$. The general case follows from connecting any two points in Ω' by an arc Γ so that $\Gamma \subseteq \bigcup\limits_{k=1}^{m} B(x_k, 4R) \subseteq \Omega$ for some $x_k \in \Gamma$ and $R > 0$. The number m only depends on Ω' and Ω, so the constant C depends only on n, Ω' and Ω. □

C.3.4 Local estimates

THEOREM C.25. *Assume u is harmonic in Ω and $U \subset\subset \Omega$. Then for any multi-index α we have*

$$\sup_{x\in U} |D^\alpha u(x)| \leqslant \Big(\frac{\mathrm{n}|\alpha|}{d}\Big)^{|\alpha|} \sup_{x\in\Omega} |u(x)|\,, \tag{C.11}$$

where $d = dist(U, \partial\Omega)$.

Proof. Suppose that $\alpha = e_i = (\underbrace{0,\cdots,0}_{(i-1)\text{ zeros}}, 1, 0, \cdots, 0)$. Since $D^\alpha u$ is harmonic,

$$\begin{aligned} D^\alpha u(x) &= \frac{1}{|B(x,r)|}\int_{B(x,r)} D^\alpha u(y)\,dy = \frac{1}{|B(x,r)|}\int_{B(x,r)} \mathrm{div}(ue_i)\,dy \\ &= \frac{1}{|B(x,r)|}\int_{\partial B(x,r)} u(y)\mathrm{N}_i(y)\,dS_y \leqslant \frac{|\partial B(x,r)|}{|B(x,r)|} \max_{y\in B(x,r)} |u(y)|\,. \end{aligned}$$

As a consequence, with U_r denoting the set $\bigcup\limits_{x\in \mathrm{U}} B(x,r)$,

$$\max_{x\in\overline{\mathrm{U}}} |D^\alpha u(x)| \leqslant \frac{\mathrm{n}}{r} \max_{x\in\overline{\mathrm{U}_r}} |u(x)|\,.$$

The general result follows from applying the above inequality $|\alpha|$ times with $r = \dfrac{d}{|\alpha|}$.
□

REMARK C.26. Inequality (C.11) is also called the *gradient estimate*for harmonic functions.

C.3.5 Regularity of weakly harmonic functions

THEOREM C.27. *For $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$, suppose that $u \in L^1_{\mathrm{loc}}(\Omega)$ and satisfies*

$$\int_\Omega u(x)\Delta\phi(x)\,dx = 0 \quad \forall\, \phi \in \mathscr{C}^2_c(\Omega)\,.$$

Then u is harmonic in Ω.

Proof. Without loss of generality, we may assume that $u \in L^1(\Omega)$.

Choose $\epsilon > 0$ sufficiently small so that $\phi^\epsilon := \eta_\epsilon * \phi \in \mathscr{C}^\infty_c(\Omega)$, where η_ϵ denotes the standard mollifiers given in Definition 1.38. By assumption,

$$0 = \int_\Omega u(x)\Delta\phi^\epsilon(x)\,dx = \int_\Omega u^\epsilon(x)\Delta\phi(x)\,dx\,. \tag{C.12}$$

Since u^ϵ is smooth, we can integrate by parts to find that

$$\int_\Omega \Delta u^\epsilon(x)\phi(x)\,dx \qquad \forall\,\phi \in \mathscr{C}_c^2(\Omega)\,.$$

Thus, $\Delta u^\epsilon = 0$ in Ω, so u^ϵ is harmonic.

By Young's inequality the sequence u^ϵ is uniformly bounded in $L^1(\Omega)$:

$$\|u^\epsilon\|_{L^1(\Omega)} \leqslant \|\eta_\epsilon\|_{L^1(\Omega)}\|u\|_{L^1(\Omega)} = \|u\|_{L^1(\Omega)}\,.$$

Since u^ϵ is harmonic,

$$u^\epsilon(x) = \frac{1}{|B(0,R)|}\int_{B(x,R)} u^\epsilon(y)\,dy\,,$$

which implies that

$$|u^\epsilon(x)| \leqslant \frac{1}{|B(0,R)|}\|u^\epsilon\|_{L^1(\Omega)}\,.$$

The local gradient estimate (C.11) then provides the inequality

$$\sup_{x\in U}|D^\alpha u(x)| \leqslant C\sup_{\Omega}|u^\epsilon| \leqslant C\|u\|_{L^1(\Omega)}\,.$$

We have therefore shown that $D^\alpha u^\epsilon$ is bounded and equi-continuous on any $U \subset\subset \Omega$. By the Arzela-Ascoli Theorem, there exists some subsequence ϵ' such that $u^{\epsilon'} \to v \in \mathscr{C}^2(U)$.

On the other hand, $u^{\epsilon'} \to u$ in $L^1(U)$ so $u = v$ on U. By pushing ∂U closer to $\partial\Omega$, we see that u is smooth in Ω. Hence for all $\phi \in \mathscr{C}_c^2(\Omega)$,

$$0 = \int_\Omega u(x)\Delta\phi(x)\,dx = \int_\Omega \Delta u(x)\phi(x)\,dx\,,$$

so that u is harmonic. □

C.3.6 Liouville's theorem

Theorem C.28. *Suppose $u : \mathbb{R}^n \to \mathbb{R}$ is harmonic and bounded. Then u is constant.*

Proof. By gradient estimate (C.11), the derivatives $D^\alpha u(x)$ have to vanish for all $x \in \mathbb{R}^n$ and multi-index α. In particular, this implies that u is constant along any lines, so u is constant. □

Corollary C.29. *Let* $\mathrm{n} \geqslant 3$ *and* $f \in \mathscr{C}^{0,\alpha}(\mathbb{R}^{\mathrm{n}})$ *with compact support in* $\mathbb{R}^{\mathrm{n}}$ *. Then any bounded solution of*

$$-\Delta u = f \quad \textit{in} \quad \mathbb{R}^{\mathrm{n}}$$

has the form

$$u(x) = \int_{\mathbb{R}^{\mathrm{n}}} \Phi(x-y) f(y)\, dy + C$$

for some constant C *.*

C.3.7 Analyticity

Theorem C.30. *Assume* u *is harmonic in* Ω *. Then* u *is analytic in* Ω *; that is, for each* $x_0 \in \Omega$ *,*

$$u(x) = \sum_{\alpha} \frac{D^\alpha u(x_0)}{\alpha!} (x - x_0)^\alpha$$

within some ball $B(x_0, r)$ *.*

C.4 Green's Functions

For fixed $x \in \Omega$, suppose that there is a harmonic function $\tilde{\Phi}^x(y)$ such that $\Delta_y \tilde{\Phi}^x(y) = 0$ for all $y \in \Omega$ and

$$\tilde{\Phi}^x(y) = \Phi(y-x) \qquad \forall\, y \in \partial\Omega\,.$$

Then by Green's second identity,

$$\int_{\Omega} \tilde{\Phi}^x(y) \Delta u(y)\, dy = \int_{\partial\Omega} \Big[\tilde{\Phi}^x(y) \frac{\partial u}{\partial \mathrm{N}}(y) - u(y) \frac{\partial \tilde{\Phi}^x}{\partial \mathrm{N}}(y) \Big] dS_y\,. \tag{C.13}$$

By (C.8) and (C.13), we obtain that if $u \in \mathscr{C}^2(\Omega) \cap \mathscr{C}^0(\overline{\Omega})$, then

$$u(x) = -\int_{\partial\Omega} u(y) \frac{\partial G}{\partial \mathrm{N}}(x-y)\, dS_y - \int_{\Omega} G(x,y) \Delta u(y)\, dy\,, \tag{C.14}$$

where $G(x,y) = \Phi(y-x) - \tilde{\Phi}^x(y)$. The function $G(x,y)$ is called the Green's function for the domain Ω.

Theorem C.31 (Symmetry of Green's function)**.** *For all* x *,* $y \in \Omega$ *,* $x \neq y$ *, we have*

$$G(x,y) = G(y,x)\,.$$

Proof. Fix $x, y \in \Omega$, $x \neq y$. Define $v(z) = G(x,z)$ and $w(z) = G(y,z)$. The goal is to show that $v(y) = w(x)$.

Letting $\Omega_\epsilon = \Omega - (B(x,\epsilon) \cup B(y,\epsilon))$, and applying Green's second identity (2.2), we obtain that

$$\int_{\partial B(x,\epsilon)} \Big[\frac{\partial v}{\partial \mathrm{N}} w - \frac{\partial w}{\partial \mathrm{N}} v\Big] dS_z = \int_{\partial B(y,\epsilon)} \Big[\frac{\partial w}{\partial \mathrm{N}} v - \frac{\partial v}{\partial \mathrm{N}} w\Big] dS\,,$$

where N denotes the inward-pointing unit normal on $\partial B(x,\epsilon) \cup \partial B(y,\epsilon)$. Passing $\epsilon \to 0$, the left-hand side converges to $w(x)$ while the right-hand side converges to $v(y)$. □

C.4.1 The case $\Omega = \mathbb{R}^{\mathrm{n}-1} \times \mathbb{R}_+$

For $x \in \mathbb{R}^{\mathrm{n}}_+ \equiv \mathbb{R}^{\mathrm{n}-1} \times \mathbb{R}_+$, let $\tilde{x} = (x_1, \cdots, x_{\mathrm{n}-1}, -x_{\mathrm{n}})$. Then $\tilde{\Phi}^x(y) = \Phi(y - \tilde{x})$ is harmonic in $\mathbb{R}^{\mathrm{n}}_+$, and $\tilde{\Phi}^x(y) = \Phi(y - x)$ for all $y \in \partial\mathbb{R}^{\mathrm{n}}_+$. Therefore, the Green's function is given by

$$G(x,y) = \Phi(y - x) - \Phi(y - \tilde{x})\,,$$

the outward unit normal to $\partial\mathbb{R}^{\mathrm{n}}_+$ is $\mathrm{N} = -\mathrm{e}_n = (0, ..., 1)$ so $\dfrac{\partial G}{\partial \mathrm{N}} = -\dfrac{\partial G}{\partial y_{\mathrm{n}}}$, and

$$\frac{\partial G}{\partial \mathrm{N}}(x,y)\Big|_{y_{\mathrm{n}}=0} = \frac{-1}{\omega_{\mathrm{n}-1}} \frac{2x_{\mathrm{n}}}{|x-y|^{\mathrm{n}}}\,,$$

and Green's representation formula (C.14) suggests that the (bounded) solution to

$$\Delta u = 0 \qquad \text{in} \quad \mathbb{R}^{\mathrm{n}}_+\,, \tag{C.15a}$$

$$u = f \qquad \text{on} \quad \partial\mathbb{R}^{\mathrm{n}}_+ \tag{C.15b}$$

is

$$u(x) = -\int_{\partial\mathbb{R}^{\mathrm{n}}_+} f(y) \frac{\partial G}{\partial \mathrm{N}}(x,y)\, dS_y = \frac{2x_{\mathrm{n}}}{\omega_{\mathrm{n}-1}} \int_{\mathbb{R}^{\mathrm{n}-1}} \frac{f(y)}{|x-y|^{\mathrm{n}}} dy \quad (\forall\, x \in \mathbb{R}^{\mathrm{n}}_+)\,. \tag{C.16}$$

With $K(x,y) = \dfrac{2x_{\mathrm{n}}}{\omega_{\mathrm{n}-1}|x-y|^{\mathrm{n}}}$, the equation (C.16) can be written as

$$u(x) = \int_{\partial\mathbb{R}^{\mathrm{n}}_+} K(x,y) f(y)\, dS_y\,.$$

The function K is termed the Poisson kernel for $\mathbb{R}^{\mathrm{n}}_+$, and (C.16) is called the Poisson integral formula.

It remains to verify that (C.16) provides a solution $u \in \mathscr{C}^2(\mathbb{R}^n_+) \cap \mathscr{C}^0(\overline{\mathbb{R}^n_+})$ with prescribed boundary condition.

Theorem C.32. *Assume that* $f \in \mathscr{C}^0(\partial\mathbb{R}^n_+) \cap L^\infty(\partial\mathbb{R}^n_+)$, *and* u *is given by* (C.16). *Then* $u \in \mathscr{C}^2(\mathbb{R}^n_+) \cap \mathscr{C}^0(\overline{\mathbb{R}^n_+})$ *satisfies* (C.15a) *and* (C.15b).

Proof. Since $K(x,y)$ is smooth if $x \neq y$, (C.16) indeed shows that u is smooth. In particular, $u \in \mathscr{C}^2(\mathbb{R}^n_+)$.

Next, we note that the Poisson kernel is normalized; namely,

$$\int_{\partial\mathbb{R}^n_+} K(x,y)\, dS_y = \frac{2x_n}{\omega_{n-1}} \int_{\mathbb{R}^{n-1}} \frac{1}{|x-y|^n} dy = 1 \qquad \forall\, x \in \mathbb{R}^n_+\,. \tag{C.17}$$

The identity (C.17) is an immediate consequence of (C.14) with $u \in \mathscr{C}^2(\mathbb{R}^n_+) \cap \mathscr{C}^0(\overline{\mathbb{R}^n_+})$ taken to be $u(x) = 1$.

Let $z \in \partial\mathbb{R}^n_+$. Since $f \in \mathscr{C}^0(\partial\mathbb{R}^n_+)$, given $\epsilon > 0$, there exists $\delta > 0$ such that

$$|f(x) - f(z)| < \frac{\epsilon}{2} \qquad \text{whenever} \qquad |x - z| < 2\delta\,.$$

Then if $|x - z| < \delta$, $x \in \mathbb{R}^n_+$,

$$\begin{aligned}
|u(x) - f(z)| &= \Big| \int_{\partial\mathbb{R}^n_+} K(x,y)\big(f(y) - f(z)\big) dS_y \Big| \\
&\leqslant \int_{\partial\mathbb{R}^n_+ \cap B(z,2\delta)} K(x,y) |f(y) - f(z)| dS_y + 2 \max_{\partial\mathbb{R}^n_+} |f| \int_{\partial\mathbb{R}^n_+ \setminus B(z,2\delta)} K(x,y)\, dS_y \\
&\leqslant \frac{\epsilon}{2} + \underbrace{2 \max_{\partial\mathbb{R}^n_+} |f| \int_{\partial\mathbb{R}^n_+ \setminus B(z,2\delta)} K(x,y)\, dS_y}_{\equiv \mathrm{I}}\,.
\end{aligned}$$

If $y \in \partial\mathbb{R}^n_+ \setminus B(z, 2\delta)$, $|y - z| > 2\delta$. So if $|x - z| < \delta_1 \leqslant \delta$,

$$|y - z| \leqslant |y - x| + |x - z| \leqslant |y - x| + \frac{1}{2}|y - z|$$

which implies $|y - x| \geqslant \frac{1}{2}|y - z|$. As a consequence,

$$\mathrm{I} \leqslant \frac{2^{n+2}}{\omega_{n-1}} \max_{\partial\mathbb{R}^n_+} |f| \int_{\partial\mathbb{R}^n_+ \setminus B(z,2\delta)} \frac{x_n}{|y - z|^n} dS_y \leqslant \frac{C x_n}{\delta}$$

for some constant $C < \infty$. Choose δ_1 even smaller so that $2C\delta_1 < \epsilon\delta$, then $|u(x) - f(z)| < \epsilon$ whenever $|x - z| < \delta_1$. This proves $u \in \mathscr{C}^0(\overline{\Omega})$. $\square$

C.4.2 The case $\Omega = B(0,1)$ or $\Omega = B(0,R)$

For $x \in B(0,1)$, let $\tilde{x} = \dfrac{x}{|x|^2}$, and $\tilde{\Phi}^x(y) = \Phi(|x|(y-\tilde{x}))$. Then $\tilde{\Phi}^x(y)$ is harmonic in Ω, i.e., $\Delta_y \tilde{\Phi}^x(y) = 0$ for all $x, y \in \Omega$. Moreover, $\tilde{\Phi}^x(y) = \Phi(y-x)$ for all $y \in \partial B(0,1)$. Therefore, the Green's function for the unit ball is

$$G(x,y) = \Phi(x-y) - \Phi(|x|(y-\tilde{x}))\,.$$

Using (C.14), we find that the solution to

$$\begin{aligned} \Delta u &= 0 \qquad \text{in} \quad B(0,1)\,, \\ u &= f \qquad \text{on} \quad \partial B(0,1) \end{aligned}$$

is

$$u(x) = -\int_{\partial B(0,1)} f(y) \frac{\partial G}{\partial \mathrm{N}}(x,y)\, dS_y = \frac{1-|x|^2}{\omega_{\mathrm{n}-1}} \int_{\partial B(0,1)} \frac{f(y)}{|x-y|^{\mathrm{n}}} dS_y\,.$$

since

$$\frac{\partial G}{\partial \mathrm{N}}(x,y)\Big|_{y\in\partial B(0,1)} = \frac{-1}{\omega_{\mathrm{n}-1}} \frac{1-|x|^2}{|x-y|^{\mathrm{n}}}\,.$$

By a change of variables, the solution to

$$\Delta u = 0 \qquad \text{in} \quad B(0,R)\,, \tag{C.18a}$$

$$u = f \qquad \text{on} \quad \partial B(0,R)\,, \tag{C.18b}$$

is then

$$u(x) = \frac{R^2-|x|^2}{\omega_{\mathrm{n}-1}R} \int_{\partial B(0,R)} \frac{f(y)}{|x-y|^{\mathrm{n}}} dS_y\,. \tag{C.19}$$

Similar to Theorem C.32, we have that

Theorem C.33. *Assume that $f \in \mathscr{C}^0(\partial B(0,R))$, and u is given by* (C.19). *Then $u \in \mathscr{C}^2(B(0,R)) \cap \mathscr{C}^0(\overline{B(0,R)})$ satisfies* (C.18a) *and* (C.18b).

The function

$$\mathscr{K}(x,y) = \frac{R^2-|x|^2}{\omega_{\mathrm{n}-1}R} \frac{1}{|x-y|^{\mathrm{n}}}$$

is the *Poisson kernel* for the ball $B(0,R)$.

C.5 Perron's Method and Solutions to the Poisson Equation[1]

In this section, we prove the existence of solutions to

$$-\Delta u = f \qquad \text{in} \quad \Omega\,, \tag{C.20a}$$
$$u = g \qquad \text{on} \quad \partial\Omega\,. \tag{C.20b}$$

using Perron's method under the assumption that $f \in \mathscr{C}^{0,\alpha}(\overline{\Omega})$ and $g \in \mathscr{C}^0(\partial\Omega)$.

First, we extend f to whole $\mathbb{R}^n$ with compact support so that the extension, still denoted by f, belongs to $\mathscr{C}^{0,\alpha}(\mathbb{R}^n)$. Let $\varphi = \Phi * f$, and $v = u - \varphi$. By Lemma C.12, $-\Delta\varphi = f$, v is harmonic. So, v solves

$$-\Delta v = 0 \qquad \text{in} \quad \Omega\,, \tag{C.21a}$$
$$v = g - \varphi \equiv \psi \qquad \text{on} \quad \partial\Omega\,. \tag{C.21b}$$

As long as we know how to solve the Dirichlet problem (C.21), we obtain a solution to (C.20) by summing φ and the solution to (C.21). Therefore, we concentrate on how (C.21) is solved.

First we generalize the notion of sub-harmonic function. Recall that a function $w \in \mathscr{C}^2(\Omega)$ is sub-harmonic if $\Delta w \geqslant 0$ in Ω, and

$$w(\xi) \leqslant \fint_{\partial B(\xi,\rho)} w(y)\,dS_y \equiv M_w(\xi,\rho)\,. \tag{C.22}$$

A function $w \in \mathscr{C}^0(\Omega)$ is called sub-harmonic, if for each $\xi \in \Omega$, (C.22) holds for all $\rho > 0$ such that $B(\xi,\rho) \subset\subset \Omega$. Let $\sigma(\Omega)$ denote the space of all sub-harmonic functions on Ω, and given $\psi \in \mathscr{C}^0(\partial\Omega)$, define

$$\sigma_\psi(\Omega) = \Big\{ w \in \sigma(\Omega) \cap \mathscr{C}^0(\overline{\Omega}) \;\Big|\; w \leqslant \psi \text{ on } \partial\Omega \Big\}\,.$$

$\sigma_\psi(\Omega)$ is non-empty since the constant function $w \equiv c$ belongs to $\sigma_\psi(\Omega)$ if $c \leqslant \inf\limits_{x\in\partial\Omega} \psi(x)$.

For $u \in \mathscr{C}^0(\Omega)$ and $B(\xi,\rho) \subset\subset \Omega$, we define $u_{\xi,\rho}$, the *harmonic lifting* of u, as the function in $\mathscr{C}^0(\Omega)$ for which

$$\Delta_x u_{\xi,\rho}(x) = 0 \qquad \text{in} \quad B(\xi,\rho)\,,$$
$$u_{\xi,\rho}(x) = u(x) \qquad \text{in} \quad \Omega\backslash B(\xi,\rho)\,.$$

[1] The reader may skip this section on the first reading

Claim: For $u \in \sigma(\Omega)$ and $B(\xi,\rho) \subset\subset \Omega$, $u(x) \leqslant u_{\xi,\rho}(x)$ for all $x \in \Omega$, and $u_{\xi,\rho} \in \sigma(\Omega)$.

Proof. It suffices to show that

$$u_{\xi,\rho}(x) \leqslant M_{u_{\xi,\rho}}(x,r) \qquad \forall\, r > 0 \text{ such that } B(x,r) \subset\subset \Omega\,. \tag{C.23}$$

If $B(x,r) \subseteq B(\xi,\rho)$, since $u_{\xi,\rho}$ is harmonic in $B(\xi,\rho)$, (C.23) holds because of the mean-value property for the harmonic functions. If $B(x,r) \cap B(\xi,\rho) = \varnothing$, then $u_{\xi,\rho} = u$, so (C.23) holds because of (C.22). Other than these two cases, let w be the harmonic function satisfying $w = u_{\xi,\rho}$ on $\partial B(x,r)$. On $\partial B(x,r) \cap B(\xi,\rho)$, $w = u_{\xi,\rho} \geqslant u$, and on $\partial B(x,r) \backslash B(\xi,\rho)$, $w = u_{\xi,\rho} = u$, $w \geqslant u$ on $\partial B(x,r)$, which by the maximum principle implies that $w \geqslant u$ in $B(x,r)$. Apply the maximum principle once again to the domain $B(x,r) \cap B(\xi,\rho)$ and $B(x,r) \backslash B(\xi,\rho)$, we conclude that $w \geqslant u_{\xi,\rho}$ in $B(x,r)$ and hence (C.23) holds. □

Claim: Let $u_1, \cdots, u_k \in \sigma_\psi(\Omega)$, and $v = \max\{u_1, \cdots, u_k\}$. Then $v \in \sigma_\psi(\Omega)$.

Proof. Given $\xi \in \Omega$, for all sufficient small ρ such that $B(\xi,\rho) \subset\subset \Omega$,

$v(\xi) = \max\{u_1(\xi), \cdots, u_k(\xi)\} \leqslant \max\{M_{u_1}(\xi,\rho), \cdots, M_{u_k}(\xi,\rho)\} \leqslant M_v(\xi,\rho)$. □

Claim: For all given $\psi \in \mathscr{C}^0(\partial\Omega)$, the function $w_\psi(x) \equiv \sup\limits_{w \in \sigma_\psi(\Omega)} w(x)$ is well-defined and is harmonic in Ω.

Proof. w_ψ defined above is well-defined due to the maximum principle for sub-harmonic functions, and is clearly in $\sigma_\psi(\Omega)$ since

$$\sup_{w \in \sigma_\psi(\Omega)} ⨍_{\partial B(\xi,\rho)} w(y)\, dS_y \leqslant ⨍_{\partial B(\xi,\rho)} \sup_{w \in \sigma_\psi(\Omega)} w(y)\, dS_y\,.$$

It suffices to show that w_ψ has the mean-value property. Suppose the contrary, then

$$w_\psi(\xi) < ⨍_{\partial B(\xi,\rho)} w_\psi(y)\, dS_y$$

for some $\xi \in \Omega$ and some ρ such that $B(\xi,\rho) \subset\subset \Omega$. By the previous two claims, the function $(w_\psi)_{\xi,\rho}$ belongs to $\sigma_\psi(\Omega)$, and $w_\psi(\xi) < (w_\psi)_{\xi,\rho}(\xi)$. Then $w_\psi(\xi) \neq \sup\limits_{w \in \sigma_\psi(\Omega)} w(\xi)$.
□

If (C.21) has a solution, it has to be equal to w_ψ defined above since the solution itself belongs to $\sigma_\psi(\Omega)$. In other words, w_ψ is the only candidate for the solution. In order to make sure that w_ψ solves (C.21), we need to make sure that w_ψ satisfies the boundary condition.

Definition C.34 (Barrier Property)**.** A domain is said to have the barrier property if for each $\eta \in \partial\Omega$, there exists a function, called a barrier function, $Q_\eta \in \sigma(\Omega) \cap C(\overline{\Omega})$ for which

$$Q_\eta(\eta) = 0\,, \quad Q_\eta(x) < 0 \quad \text{for } x \in \partial\Omega\,,\, x \neq \eta\,.$$

The barrier property can be verified for a large class of domains Ω. For example, if Ω is *strictly convex* in the sense that through each point $\eta \in \partial\Omega$ there passes a hyperplane π_η having only η in common with $\overline{\Omega}$, then Ω has the barrier property.

As long as Ω has the barrier property, for each $y \in \partial\Omega$, there exists a sub-harmonic function $w \in \sigma_\psi(\Omega)$ and $w(y) = \psi(y)$ (for example, consider $w(x) = \psi(y) + Q_y(x)$). The only thing it remains to be proved is that $w_\psi \in \mathscr{C}^0(\overline{\Omega})$, or

$$\lim_{\substack{x \in \Omega \\ x \to \eta}} w_\psi(x) = \psi(\eta)\,.$$

Claim: If Ω has the barrier property, then for $\eta \in \partial\Omega$,

$$\liminf_{\substack{x \in \Omega \\ x \to \eta}} w_\psi(x) \geqslant \psi(\eta)\,.$$

Proof. For $\epsilon > 0$ and $K > 0$, the function $v(x) = \psi(\eta) - \epsilon + KQ_\eta(x)$ belongs to $\sigma(\Omega) \cap \mathscr{C}^0(\overline{\Omega})$, and satisfies

$$v(x) \leqslant \psi(\eta) - \epsilon \quad \forall\, x \in \partial\Omega\,, \qquad v(\eta) = \psi(\eta) - \epsilon\,.$$

Since $\psi \in \mathscr{C}^0(\partial\Omega)$, there exists $\delta > 0$ such that $|\psi(x) - \psi(\eta)| < \epsilon$ whenever $|x - \eta| < \delta$, $x \in \partial\Omega$; thus $v(x) \leqslant \psi(x)$ if $|x - \eta| < \delta$. If $|x - \eta| \geqslant \delta$, we can choose K large enough so that $v(x) \leqslant \psi(x)$ since Q_η has negative upper bound on $|x - \eta| \geqslant \delta$. Therefore, $v \in \sigma_\psi(\Omega)$ (if K is large enough). By the definition of w_ψ, $v(x) \leqslant w_\psi(x)$ for all $x \in \Omega$; hence

$$\psi(\eta) - \epsilon = \liminf_{\substack{x \in \Omega \\ x \to \eta}} v(x) \leqslant \liminf_{\substack{x \in \Omega \\ x \to \eta}} w_\psi(\eta)\,. \tag{C.24}$$

We then conclude the claim since (C.24) holds for all $\epsilon > 0$ and all $\eta \in \partial\Omega$. $\square$

Claim: If Ω has the barrier property, then for $\eta \in \partial\Omega$,

$$\lim_{\substack{x\in\Omega\\ x\to\eta}} w_\psi(x) = \psi(\eta)\,.$$

Proof. It suffices to show that

$$\limsup_{\substack{x\in\Omega\\ x\to\eta}} w_\psi(x) \leqslant \psi(\eta)\,. \tag{C.25}$$

This is done by considering $-w_{-\psi}(x)$ which is defined in Ω by

$$-w_{-\psi}(x) = -\sup_{w\in\sigma_{-\psi}(\Omega)} w(x) = \inf_{-v\in\sigma_{-\psi}(\Omega)} v(x)\,.$$

For all $w \in \sigma_\psi(\Omega)$ and $-v \in \sigma_{-\psi}(\Omega)$, $w \leqslant \psi \leqslant v$ on $\partial\Omega$, and $w - v \in \sigma(\Omega) \cap \mathscr{C}^0(\overline{\Omega})$; therefore by the maximum principle for the sub-harmonic functions,

$$w \leqslant v \quad \forall\, x \in \Omega \quad \Rightarrow \quad w_\psi(x) \leqslant -w_{-\psi}(x) \quad \forall\, x \in \Omega\,.$$

By previous claim,

$$\liminf_{\substack{x\in\Omega\\ x\to\eta}} w_{-\psi}(x) \geqslant -\psi(\eta) \;\Rightarrow\; \limsup_{\substack{x\in\Omega\\ x\to\eta}} w_\psi(x) \leqslant \limsup_{\substack{x\in\Omega\\ x\to\eta}} -w_{-\psi}(x) \leqslant \psi(\eta)\,.$$ □

THEOREM C.35. *If the domain Ω has the barrier property, and $f \in \mathscr{C}^{0,\alpha}(\overline{\Omega})$, then there exists a unique solution $u \in \mathscr{C}^2(\Omega) \cap \mathscr{C}^0(\overline{\Omega})$ of the Dirichlet problem*

$$\begin{aligned} -\Delta u &= f \qquad \text{in} \quad \Omega \\ u &= g \qquad \text{on} \quad \partial\Omega\,, \end{aligned}$$

for arbitrary continuous boundary value g.

C.6 Exercises

PROBLEM C.1. Let $\Omega = B(0, \frac{1}{2}) \subseteq \mathbb{R}^2$ denote the open ball of radius $\frac{1}{2}$ centered at the origin. For $x = (x_1, x_2) \in \Omega$, let

$$u(x_1, x_2) = x_1 x_2 \left[\log\left(\log \frac{1}{\sqrt{x_1^2 + x_2^2}} \right) - \log\log 2 \right].$$

(a) Show that $u \in \mathscr{C}^1(\bar{\Omega})$;

(b) Show that $\Delta u \in C(\Omega)$, but that $u \notin \mathscr{C}^2(\Omega)$.

PROBLEM C.2. Let $u \in \mathscr{C}^2_{\text{loc}}(\mathrm{D}) \cap C(\overline{\mathrm{D}})$ be a solution to the problem

$$\begin{aligned} -\Delta u &= 1 \qquad \text{in} \quad \mathrm{D} \equiv (-1,1) \times (-1,1) \subseteq \mathbb{R}^2 , \\ u &= 0 \qquad \text{on} \quad \partial \mathrm{D} . \end{aligned}$$

Show that u cannot belong to $\mathscr{C}^2(\overline{\mathrm{D}})$.

PROBLEM C.3. Find a solution to the Dirichlet problem $\Delta u = 0$ in the square

$$\left\{(x, y) \in \mathbb{R}^2 \,\middle|\, -1 \leqslant x \leqslant 1 , -1 \leqslant y \leqslant 1\right\}$$

satisfying the boundary conditions

$$\begin{aligned} u(x,y) &= \cos\left(\frac{3\pi}{2} x\right), \quad \text{on } y = \pm 1 ,\ -1 \leqslant x \leqslant 1 , \\ u(x,y) &= \cos\left(\frac{3\pi}{2} y\right), \quad \text{on } x = \pm 1 ,\ -1 \leqslant y \leqslant 1 . \end{aligned}$$

PROBLEM C.4. Let u, v be smooth harmonic functions, such that

$$u(tx) \equiv t^a u(x) , \qquad v(tx) = t^b v(x)$$

for all $x \in \mathbb{R}^{\mathrm{n}}$, $t > 0$, with constants $a \neq b$. Use Green's identity to show that

$$\int_{\partial B(0,1)} uv dS = 0 .$$

PROBLEM C.5. Let u be a harmonic function in the unit ball $B_1 \equiv B(x_0, 1) \subseteq \mathbb{R}^{\mathrm{n}}$. Prove the following gradient estimate:

$$|\nabla u(x_0)| \leqslant \mathrm{n}\Big[\sup_{x \in B_1} u(x) - u(x_0)\Big]$$

Hint: Note that all the derivatives $\partial_{x_i} u$ are harmonic in B_1, so that by the mean value and divergence theorems,

$$\frac{\partial u}{\partial x_i}(x_0) = \frac{1}{|B_1|} \int_{B_1} \frac{\partial u}{\partial x_i}(x)\, dx = \frac{1}{|B_1|} \int_{\partial B_1} u \mathrm{N}_i dS ,$$

where N_i is the i-th component of the unit normal N to ∂B_1. Obviously, for $x \in \partial B_1$, we have $\mathrm{N}(x) = x - x_0$.

Problem C.6. Show that there are no functions $u \in \mathscr{C}^2_{\mathrm{loc}}(\mathbb{R}^2_+) \cap \mathcal{C}_{\mathrm{loc}}(\overline{\mathbb{R}^2_+})$ satisfying the properties

$$u \geqslant 0\,, \quad \Delta u = 0 \ \text{ in } \ \mathbb{R}^2_+ \equiv \{x = (x_1, x_2) \in \mathbb{R}^2 \mid x_2 > 0\}\,, \quad u(x_1, 0) = x_1^2\,.$$

Hint: Suppose there exists such a function u. For arbitrary $R > 0$, compare u with the solution $v \in \mathscr{C}^2_{\mathrm{loc}}(\mathbb{R}^2_+) \cap C(\overline{\mathbb{R}^2_+})$ to the problem

$$\Delta v = 0 \text{ in } \mathbb{R}^2_+\,, \quad v(x_1, 0) = \zeta(x_1) x_1^2\,,$$

where

$$\zeta \in \mathscr{C}^\infty_c(-2R, 2R)\,, \quad 0 \leqslant \zeta \leqslant 1\,, \quad \text{and} \quad \zeta \equiv 1 \text{ on } [-R, R]\,.$$

Problem C.7. Let $u(x_1, x_2) \in \mathscr{C}^2(\overline{\Omega})$, where $\Omega \equiv \big\{x = (x_1, x_2) \in \mathbb{R}^2 \;\big|\; x_1 > 0\,, x_2 > 0\big\}$, such that

$$\Delta u = 0 \quad \text{in} \quad \Omega\,, \qquad u = 0 \quad \text{on} \quad \partial\Omega\,.$$

In addition, let $|u(x)| \leqslant c_1 + c_2|x|$ in Ω with some constants c_1 and c_2. Show that $u \equiv 0$ in Ω.
Hint: Extend the domain where u is defined to $\mathbb{R}^2$ and use the gradient estimates to show that u is in fact a linear function.

Problem C.8. Let $u(x) = u(x_1, x_2)$ be a bounded solution of the Laplace equation

$$\Delta u = 0 \qquad \text{in} \quad \mathbb{R}^2_+$$

with boundary condition

$$u(x_1, 0) = g(x_1) = \frac{|x_1|}{1 + x_1^2}\,.$$

Show that the gradient ∇u is unbounded on $\mathbb{R}^2_+$.

Problem C.9. Find the Green function for the domain $\Omega = \big\{x \in \mathbb{R}^{\mathrm{n}} \big| |x|^2 < 1\,, x_{\mathrm{n}} > 0\big\}$ and the corresponding Green's representation formula for the solution of the Dirichlet problem

$$\begin{aligned} \Delta u &= 0 \qquad \text{in} \quad \Omega\,, \\ u &= f \qquad \text{on} \quad \partial\Omega\,. \end{aligned}$$

Appendix D

L^p-Estimates for Solutions of Elliptic Equations

D.1 The Riesz Potential

Let $\mu \in (0,1]$, and $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ be bounded. We define the operator V_μ on $L^1(\Omega)$ by the Riesz potential

$$(V_\mu f)(x) = \int_\Omega |x-y|^{\mathrm{n}(\mu-1)} f(y)dy\,.$$

Note that $\mu = \dfrac{2}{\mathrm{n}}$ corresponds to the case of Newtonian potential (introduced in Chapter C).

The first observation is that when $f \equiv 1$, then

$$V_\mu 1 = \int_\Omega |x-y|^{\mathrm{n}(\mu-1)}dy \leqslant \int_{B(x,R)} |x-y|^{\mathrm{n}(\mu-1)}dy \leqslant \frac{\omega_{\mathrm{n}-1}R^{\mathrm{n}\mu}}{\mathrm{n}\mu}\,, \tag{D.1}$$

where R is chosen so that $|\Omega| = |B(x,R)| = \dfrac{\omega_{\mathrm{n}-1}}{\mathrm{n}} R^{\mathrm{n}}$.

Theorem D.1. *The operator V_μ maps $L^p(\Omega)$ continuously into $L^q(\Omega)$ for any q, $1 \leqslant q \leqslant \infty$ satisfying*

$$0 \leqslant \delta = \delta(p,q) = \frac{1}{p} - \frac{1}{q} < \mu\,.$$

Furthermore, for any $f \in L^p(\Omega)$,

$$\|V_\mu f\|_{L^q(\Omega)} \leqslant C\|f\|_{L^p(\Omega)}\,,$$

where C depends only on μ, δ and $|\Omega|$.

Proof. Let $r^{-1} = 1 - \delta$ and $h(x-y) = |x-y|^{\mathrm{n}(\mu-1)}$. Then $h(x-\cdot) \in L^r(\Omega)$ and by the same type of estimates as (D.1),

$$\|h\|_{L^r(\Omega)} \leqslant C(\mu, \delta, |\Omega|)\,.$$

By writing $h(x-y)|f(y)| = h(x-y)^{\frac{r}{q}} h(x-y)^{r(1-\frac{1}{p})} |f(y)|^{\frac{p}{q}} |f(y)|^{p\delta}$, we find that

$$|V_\mu f(x)| \leqslant \Big[\int_\Omega h^r(x-y)|f(y)|^p dy\Big]^{\frac{1}{q}} \Big[\int_\Omega h^r(x-y) dy\Big]^{1-\frac{1}{p}} \Big[\int_\Omega |f(y)|^p dy\Big]^{\delta};$$

hence

$$\|V_\mu f\|_{L^q(\Omega)} \leqslant C(\mu, \delta, |\Omega|) \|f\|_{L^p(\Omega)}\,. \qquad \square$$

D.2 Marcinkiewicz Interpolation Theorem

Definition D.2. Let f be a measurable function on a domain Ω in $\mathbb{R}^{\mathrm{n}}$. The distribution function $\mu = \mu_f$ of f is defined by

$$\mu_f(t) = \big|\big\{x \in \Omega \,\big|\, |f(x)| > t\big\}\big|\,.$$

Some properties of the distribution function:

1. μ_f is non-decreasing in $[0, \infty)$.
2. μ_f is right continuous; that is, $\lim\limits_{t \to t_0^+} f(t) = f(t_0)$ for all $t_0 \in [0, \infty)$.
3. **(Layer cake representation)** If $u \in L^p(\Omega)$, then

$$\|f\|_{L^p(\Omega)} = \begin{cases} \Big[p\displaystyle\int_0^\infty t^{p-1}\mu_f(t)dt\Big]^{\frac{1}{p}} & \text{if } p \in [1, \infty)\,, \\ \inf\big\{t \in [0, \infty) \,\big|\, \mu_f(t) = 0\big\} & \text{if } p = \infty\,. \end{cases} \tag{D.2}$$

4. **(Chebyshev's inequality)** For any $p > 0$ and $f \in L^p(\Omega)$,

$$t^p \mu_f(t) \leqslant \int_{\{|f|>t\}} |f(x)|^p dx \leqslant \|f\|^p_{L^p(\Omega)}\,. \tag{D.3}$$

Definition D.3. The space weak-$L^p(\Omega)$ consists of measurable functions f such that

$$\mu_f(t) \leqslant Ct^{-p}$$

for some constant C (depending on f).

Theorem D.4. *Let T be a linear mapping from $L^q(\Omega) \cap L^r(\Omega)$ into itself, $1 \leqslant q < r < \infty$ and suppose that there are constants M_1 and M_2 such that*

$$\mu_{Tf}(t) \leqslant \Big(\frac{M_1\|f\|_{L^q(\Omega)}}{t}\Big)^q, \quad \mu_{Tf}(t) \leqslant \Big(\frac{M_2\|f\|_{L^r(\Omega)}}{t}\Big)^r \tag{D.4}$$

for all $f \in L^q(\Omega) \cap L^r(\Omega)$ and $t > 0$. Then T extends as a bounded linear mapping from $L^p(\Omega)$ into itself for any p in between q and r, and

$$\|Tf\|_{L^p(\Omega)} \leqslant CM_1^{\alpha}M_2^{1-\alpha}\|f\|_{L^p(\Omega)} \tag{D.5}$$

for all $f \in L^q(\Omega) \cap L^r(\Omega)$, where

$$\frac{1}{p} = \frac{\alpha}{q} + \frac{1-\alpha}{r}$$

and C depends only on p, q and r.

Remark D.5. Condition (D.4) is the same as saying that $T : L^q(\Omega) \to$ weak-$L^q(\Omega)$ and $T : L^r(\Omega) \to$ weak-$L^r(\Omega)$ are bounded.

Proof. For $f \in L^q(\Omega) \cap L^r(\Omega)$ and any $s > 0$, we write $f = g_s + h_s$, where $g_s = f\chi_{\{|f|>s\}}$ and $h_s = f\chi_{\{|f|\leqslant s\}}$. Then $|Tf| \leqslant |Tf_1| + |Tf_2|$, and hence

$$\mu_{Tf}(t) \leqslant \mu_{Tf_1}(t/2) + \mu_{Tf_2}(t/2) \leqslant \Big(\frac{2M_1\|g_s\|_{L^q(\Omega)}}{t}\Big)^q + \Big(\frac{2M_2\|h_s\|_{L^r(\Omega)}}{t}\Big)^r .$$

By (D.2) with $s = t/A$ (for some A to be determined later),

$$\begin{aligned}
\|Tf\|^p_{L^p(\Omega)} &\leqslant p(2M_1)^q \int_0^\infty t^{p-1-q}\Big[\int_{\{|f|>t/A\}} |f(x)|^q dx\Big]dt \\
&\quad + p(2M_2)^r \int_0^\infty t^{p-1-r}\Big[\int_{\{|f|\leqslant t/A\}} |f(x)|^r dx\Big]dt \\
&= p(2M_1)^q A^{p-q} \int_0^\infty s^{p-1-q}\Big[\int_{\{|f|>s\}} |f(x)|^q dx\Big]ds \\
&\quad + p(2M_2)^r A^{p-r} \int_0^\infty s^{p-1-r}\Big[\int_{\{|f|\leqslant s\}} |f(x)|^r dx\Big]ds .
\end{aligned}$$

Now, by the Fubini Theorem,

$$\int_0^\infty s^{p-1-q}\Big[\int_{\{|f|>s\}}|f(x)|^q dx\Big]ds = \int_\Omega\Big[\int_0^{|f(x)|} s^{p-1-q}ds\Big]|f(x)|^q dx = \frac{1}{p-q}\|f\|^p_{L^p(\Omega)}$$

and similarly we have $\displaystyle\int_0^\infty s^{p-1-q}\Big[\int_{\{|f|\leqslant s\}}|f(x)|^r dx\Big]ds = \frac{1}{r-p}\|f\|^p_{L^p(\Omega)}$. Therefore, for all $A > 0$,

$$\|Tf\|^p_{L^p(\Omega)} \leqslant \Big[\frac{p}{p-q}(2M_1)^q A^{p-q} + \frac{p}{r-p}(2M_2)^r A^{p-r}\Big]\|f\|^p_{L^p(\Omega)} .$$

Minimizing the bracket on the right-hand side, we find that the minimum of the right-hand side is attained when $A = 2M_1^{\frac{q}{r-q}}M_2^{\frac{r}{r-q}}$ which implies the desired inequality. □

D.3 Calderon-Zygmund Inequality

Theorem D.6. *Let $f \in L^p(\Omega)$, $1 < p < \infty$, and let w be the Newtonian potential with density f defined in Remark* C.14. *Then $w \in W^{2,p}(\Omega)$, $-\Delta w = f$ a.e., and*

$$\|D^2 w\|_{L^p(\Omega)} \leqslant C\|f\|_{L^p(\Omega)} , \tag{D.6}$$

where C depends only on n *and p . Furthermore, if $p = 2$, we have*

$$\|D^2 w\|_{L^2(\mathbb{R}^{\mathrm{n}})} = \|f\|_{L^2(\Omega)} . \tag{D.7}$$

Proof. We first note that (D.6) and Theorem D.1 implies $w \in W^{2,p}(\Omega)$.

1. We prove first the case that $p = 2$. Suppose that $f \in \mathscr{C}_c^\infty(\mathbb{R}^{\mathrm{n}})$, then $w \in \mathscr{C}^\infty(\mathbb{R}^{\mathrm{n}})$ and $-\Delta w = f$. Therefore, if $\mathrm{spt}(f) \subseteq B(0,R)$,

$$\int_{B(0,R)} |\Delta w(x)|^2 dx = \int_{B(0,R)} |f(x)|^2 dx .$$

   On the other hand, integrating by parts implies that

$$\begin{aligned}\int_{B(0,R)} |D^2 w(x)|^2 dx &= \sum_{i,j=1}^{\mathrm{n}} \int_{B(0,R)} |w_{x_i x_j}(x)|^2 dx \\ &= -\sum_{i=1}^{\mathrm{n}} \int_{B(0,R)} w_{x_i}(x)\Delta w_{x_i}(x)dx + \int_{\partial B(0,R)} w_{x_i}(x) w_{x_i x_j}(x)\mathrm{N}_j dS \\ &= \int_{B(0,R)} |\Delta w(x)|^2 dx + \int_{\partial B(0,R)} w_{x_i}(x) w_{x_i x_j}(x)\mathrm{N}_j dS ,\end{aligned}$$

   where in the last equality we use the fact that $\Delta w = 0$ on $\partial B(0,R)$ to avoid having another boundary integral. By (C.1a) and (C.1b), $|\nabla w| = \mathcal{O}(R^{1-\mathrm{n}})$ and

$|D^2w| = \mathcal{O}(R^{-\mathrm{n}})$; hence the boundary integral on the right-hand side approaches zero as $R \to \infty$. So (D.7) is established for the case $f \in \mathscr{C}_c^\infty(\mathbb{R}^{\mathrm{n}})$.

Now suppose that $f \in L^2(\Omega)$. Choose $f_k \in \mathscr{C}_c^\infty(\Omega) \subseteq \mathscr{C}_c^\infty(\mathbb{R}^{\mathrm{n}})$ such that $f_k \to f$ in $L^2(\Omega)$ as $k \to \infty$. Then $w_k = \Phi * f_k \in H^1(\Omega)$ and $\|w_k - w_\ell\|_{H^1(\Omega)} \leqslant C\|f_k - f_\ell\|_{L^2(\Omega)}$ because of Theorem D.1. Moreover, $\|D^2(w_k - w_\ell)\|_{L^2(\mathbb{R}^{\mathrm{n}})} = \|f_k - f_\ell\|_{L^2(\Omega)}$ because of (D.7). Therefore, w_k is a Cauchy sequence in $H^2(\Omega)$ hence converges in $H^2(\Omega)$. Again by Theorem D.1 $\|w_k - w\|_{H^1(\Omega)} \to 0$ as $k \to \infty$, so the H^2-limit must be w as well; thus

$$\|D^2w\|_{L^2(\Omega)} \leqslant C\|f\|_{L^2(\Omega)} . \tag{D.8}$$

Since $w_k \to w$ in $H^2(\Omega)$, $\lim\limits_{k\to\infty}\langle \Delta w_k, \varphi\rangle = \langle \Delta w, \varphi\rangle$ thus $-\Delta w = f$ a.e.

Finally, we explain why equality in (D.7) still holds if $f \in L^2(\Omega)$. First, by (D.7) D^2w_k is a Cauchy sequence in $L^2(\mathbb{R}^{\mathrm{n}})$, so it converges to some element $u \in L^2(\mathbb{R}^{\mathrm{n}})$. The restriction of u to Ω is D^2w sine $w_k \to w$ in $H^2(\Omega)$. It is unclear if the limit is still the Hessian of w outside Ω . Nevertheless, for all bounded $V \supset\supset \Omega$,

$$w(x) = \int_\Omega \Phi(x-y)f(y) = \int_V \Phi(x-y)f(y)dy .$$

if we extend f to be zero outside Ω . By doing so, we know that $w_k \to w$ in $H^2(V)$, thus $u = D^2w$ in V . Since V is arbitrary, $u = D^2w$ in $\mathbb{R}^{\mathrm{n}}$, and (D.7) follows.

2. For fixed i and j , we define the linear operator $T : L^2(\Omega) \to L^2(\Omega)$ by $Tf = w_{x_ix_j}$. Then (D.7) and the Chebyshev inequality (D.3) imply

$$\mu_{Tf}(t) \leqslant \Big(\frac{\|f\|_{L^2(\Omega)}}{t}\Big)^2 \qquad \forall\ t > 0 , f \in L^2(\Omega) .$$

We next show that

$$\mu_{Tf}(t) \leqslant \frac{C\|f\|_{L^1(\Omega)}}{t} \qquad \forall\ t > 0 , f \in L^1(\Omega) , \tag{D.9}$$

then (D.6) (for the case that $1 \leqslant p \leqslant 2$) follows from the Marcinkiewicz interpolation theorem.

Extend f to be zero outside Ω, and for fixed $t>0$ choose a cube $K_0 \supseteq \Omega$ such that

$$\int_{K_0} |f(x)|dx \leqslant t|K_0|\,.$$

Subdivide K_0 into $2^{\rm n}$ congruent subcubes with disjoint interiors. Those subcubes K satisfying $\displaystyle\int_K |f|dx \leqslant t|K|$ are subdivided again and the process is repeated indefinitely. Let $\mathcal{P}$ denote the collection of subcubes K that satisfy $\displaystyle\int_K |f(x)| > t|K|$, and for each $K \in \mathcal{P}$ let $\widetilde{K}$ denote the subcube whose subdivision gives K. Then for each $K \in \mathcal{P}$,

$$t < \frac{1}{|K|}\int_K |f(x)|dx \leqslant \frac{2^{\rm n}}{|\widetilde{K}|}\int_K |f(x)|dx \leqslant \frac{2^{\rm n}}{|\widetilde{K}|}\int_{\tilde{K}} |f(x)|dx \leqslant 2^{\rm n} t\,, \tag{D.10}$$

and by Lebesgue differentiation theorem $|f| \leqslant t$ a.e. on $G \equiv K_0 - \bigcup\limits_{K\in\mathcal{P}} K$. The function f then can be expressed as the sum of two functions g and b (that stand for good and bad parts of f), where g and b are defined by

$$g(x) = \begin{cases} f(x) & \text{if } x \in G\,, \\ \dfrac{1}{|K|}\displaystyle\int_K |f(x)|dx & \text{if } x \in K \text{ for some } K \in \mathcal{P}\,, \end{cases}$$

and $b = f - g$. It is then easy to see that $|g| \leqslant 2^{\rm n} t$ a.e., $b = 0$ if $x \in G$, and $\displaystyle\int_K bdx = 0$.

3. To prove (D.9), we note that

$$\mu_{Tf}(t) \leqslant \mu_{Tg}(t/2) + \mu_{Tb}(t/2)\,, \tag{D.11}$$

so we concentrate on the upper bounds of the right-hand side. Clearly $g \in L^2(\Omega)$, so by $|g| \leqslant 2^{\rm n} t$ a.e.,

$$\mu_{Tg}(t/2) \leqslant \frac{4\|g\|^2_{L^2(\Omega)}}{t^2} \leqslant \frac{4\cdot 2^{\rm n} t}{t^2}\int_\Omega |g(x)|dx \leqslant \frac{2^{\rm n+2}}{t}\int_\Omega |f(x)|dx\,.$$

Let $b_K = b\chi_K$. For each $K \in \mathcal{P}$, there exists $b_m \in \mathscr{C}_c^\infty(K)$ converging to b_K in $L^2(\Omega)$ and satisfying $\displaystyle\int_K b_m dx = 0$. Then if $x \notin K$, we have

$$\begin{aligned} Tb_m(x) &= \int_K \Phi_{x_ix_j}(x-y)b_m(y)dy \\ &= \int_K \Big[\Phi_{x_ix_j}(x-y) - \Phi_{x_ix_j}(x-\bar{y})\Big]b_m(y)dy\,, \end{aligned}$$

where $\bar{y}$ denotes the center of K. By the mean value theorem,

$$\begin{aligned}&|\Phi_{x_ix_j}(x-y)-\Phi_{x_ix_j}(x-\bar{y})|\\&\qquad=|(y-\bar{y})\cdot\nabla\Phi_{x_ix_j}(x-\hat{y})|\leqslant C|y-\bar{y}|\mathrm{dist}(x,K)^{-1-\mathrm{n}};\end{aligned}$$

hence by letting $\delta=\mathrm{diam}(K)$,

$$|Tb_m(x)|\leqslant C\delta\mathrm{dist}(x,K)^{-1-\mathrm{n}}\int_K|b_m(y)|dy\,.$$

Since $\mathrm{dist}(x,K)\geqslant c|x-\bar{y}|$ for all $x\notin B(\bar{y},\delta)$ for some fixed constant $c>0$,

$$\begin{aligned}&\int_{K_0-B(\bar{y},\delta)}|Tb_m(x)|dx\\&\qquad\leqslant C\delta\int_{\mathbb{R}^\mathrm{n}\backslash B(\bar{y},\delta)}|x-\bar{y}|^{-1-\mathrm{n}}dx\int_K|b_m(y)|dy\leqslant C\int_K|b_m(x)|dx\,.\end{aligned}$$

Passing m to infinity we conclude that

$$\int_{K_0-B(\bar{y}_K,\delta_K)}|Tb_K(x)|dx\leqslant C\int_K|b_K(x)|dx\,.$$

Therefore, if $F^*=\bigcup_{K\in\mathcal{P}}B(\bar{y}_K,\delta_K)$ and $G^*=K_0-F^*$,

$$\begin{aligned}\int_{G^*}|Tb(x)|dx&=\int_{G^*}\Big|T\sum_{K\in\mathcal{P}}b_K(x)\Big|dx\leqslant\sum_{K\in\mathcal{P}}\int_{G^*}|Tb_K(x)|dx\\&\leqslant C\sum_{K\in\mathcal{P}}\int_K|b_K(x)|dx\leqslant C\int_\Omega|b(x)|dx\leqslant C\int_\Omega|f(x)|dx\,.\end{aligned}$$

By the Chebyshev inequality (D.3),

$$\Big|\Big\{x\in G^*\,\Big|\,|Tb(x)|>t/2\Big\}\Big|\leqslant\frac{C\|f\|_{L^1(\Omega)}}{t}\,.$$

Moreover, by (D.10),

$$\begin{aligned}|B(\bar{y}_K,\delta_K)|&=\frac{\omega_{\mathrm{n}-1}}{\mathrm{n}}\delta_K^\mathrm{n}=\frac{\omega_{\mathrm{n}-1}}{\mathrm{n}}2^\mathrm{n}\mathrm{n}^{\frac{\mathrm{n}}{2}}\Big(\frac{\delta_K}{2\sqrt{\mathrm{n}}}\Big)^\mathrm{n}\\&=\frac{\omega_{\mathrm{n}-1}}{\mathrm{n}}2^\mathrm{n}\mathrm{n}^{\frac{\mathrm{n}}{2}}|K|\leqslant\frac{\omega_{\mathrm{n}-1}2^\mathrm{n}\mathrm{n}^{\frac{\mathrm{n}}{2}}}{\mathrm{n}t}\int_K|f(x)|dx\,;\end{aligned}$$

thus $|F^*|\leqslant\dfrac{C\|f\|_{L^1(\Omega)}}{t}$. As a consequence, $\mu_{Tg}(t/2)\leqslant\dfrac{C\|f\|_{L^1(\Omega)}}{t}$ thus (D.9) follows from (D.11).

4. By Marcinkiewicz interpolation theorem,

$$\|Tf\|_{L^p(\Omega)} \leqslant C(\mathrm{n},p)\|f\|_{L^p(\Omega)} \qquad \forall\ f \in L^p(\Omega)\,, p \in (1,2]\,.$$

Inequality (D.6) can be extended to the case $p > 2$ by duality: suppose that $f, g \in \mathscr{C}_c^\infty(\Omega)$, then

$$\begin{aligned}\int_\Omega Tf(x)g(x)dx &= \int_\Omega w(x)g_{x_ix_j}(x)dx = \int_\Omega\int_\Omega \Phi(x-y)f(y)g_{x_ix_j}(x)dydx \\ &= \int_\Omega\int_\Omega \Phi(x-y)f(y)g_{x_ix_j}(x)dxdy = \int_\Omega f(x)Tg(x)dx\,;\end{aligned}$$

hence by denoting $p' = \dfrac{p}{p-1} < 2$,

$$\left|\int_\Omega Tf(x)g(x)dx\right| \leqslant \|f\|_{L^p(\Omega)}\|Tg\|_{L^{p'}(\Omega)} \leqslant C\|f\|_{L^p(\Omega)}\|g\|_{L^{p'}(\Omega)}\,.$$

The above inequality holds for all $g \in \mathscr{C}_c^\infty(\Omega)$, and by the density argument, it holds for all $g \in L^{p'}(\Omega)$, so for all $f \in \mathscr{C}_c^\infty(\Omega)$,

$$\|Tf\|_{L^p(\Omega)} \leqslant C\|f\|_{L^p(\Omega)}\,. \tag{D.12}$$

The argument leads to (D.8) then implies that the equality above holds
for all $f \in L^p(\Omega)$. $\square$