Lecture Notes on Analysis

Steve Shkoller

February 29, 2016

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 a 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 $\mathscr{C}^0(\Omega)$ or $\mathscr{C}(\Omega)$ denote the continuous functions on Ω, and we denote by $\mathscr{C}_c^0(\Omega)$ or $\mathscr{C}_c(\Omega)$ those functions in $\mathscr{C}^0(\Omega)$ with compact support contained in Ω.

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

$$\mathscr{C}^0(\overline{\Omega}) := \{u : \Omega \to \mathbb{R} \mid u \text{ is bounded and continuous}\},$$

with norm $\|u\|_{\mathscr{C}^0(\overline{\Omega})} = \max_{x \in \overline{\Omega}} |u(x)|$. For integers $k \geqslant 0$, we let $\mathscr{C}^k(\overline{\Omega})$ denote the functions possessing partial derivatives to all orders up to k which are bounded and continuous on $\overline{\Omega}$. We use $\mathscr{C}^k_{\mathrm{loc}}(\Omega)$ to denote the functions in $\mathscr{C}^k(\overline{B})$ for all bounded balls contained in Ω.

$\mathscr{C}^k(\Omega)$ 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.

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

$\operatorname{spt}(f)$ denotes the support of a function f, and is the closure of the set $\{x \in \Omega \mid f(x) \neq 0\}$.

$\mathscr{C}_c(\Omega) = \{u \in \mathscr{C}(\Omega) \mid \operatorname{spt} u \text{ compact in } \Omega\}$.

$\mathscr{C}_c^k(\Omega) = \mathscr{C}^k(\Omega) \cap \mathscr{C}_c(\Omega)$.
$\mathscr{C}_c^\infty(\Omega) = \mathscr{C}^\infty(\Omega) \cap \mathscr{C}_c(\Omega)$. 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. [1] We will assume that all functions and sets are Lebesgue measurable. The Lebesgue measure is often denoted by μ so that $\mu(\Omega)$ denotes the length if $\mathrm{n}=1$, the area if $\mathrm{n}=2$, the volume if $\mathrm{n}=3$, and so on.

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.

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

Theorem 1.3. *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 1.4. *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.

Theorem 1.5 (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 1.6 (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 1.7. 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 1.8 (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.

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 1.9 (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 1.10 (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 1.11. *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 The L^p Space

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

1.3.1 Definitions and basic properties

Definition 1.12. Let $0 < p < \infty$ and let Ω denote and open smooth subset of $\mathbb{R}^{\mathrm{n}}$. 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.

Definition 1.13. The space $L^p(\Omega)$ is the set

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

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 $p \geqslant 1$.

The most interesting cases are $p = 1, 2, \infty$, while all of the L^p arise often in *nonlinear* estimates.

Definition 1.14. 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

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

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

1.3.2 Basic inequalities

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

Lemma 1.15. *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 1.16. *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 1.15 to obtain the desired inequality. $\square$

Theorem 1.17 (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 1.16. Let $\lambda = \frac{1}{p}$ and set

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

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 1.18. 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 1.19 (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 $\frac{1}{s} = \frac{a}{r} + \frac{1-a}{t}$*

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

Proof. By Hölder's inequality,

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

□

Theorem 1.20 (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 $q = \dfrac{p}{p-1}$,

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

from which it follows that

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

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

Corollary 1.21. *For $1 \leqslant p \leqslant \infty$, $L^p(\Omega)$ is a normed linear space.*

Example 1.22 (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 1.23 (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 1.24 (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 1.25. *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 $k \geqslant 1$,

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

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,

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

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

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

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

Example 1.27. 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 1.28. *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 $\lim_{k\to\infty} \|f_k\|_{L^p(\Omega)} = \|f\|_{L^p(\Omega)}$, then $f_k \to f$ in $L^p(\Omega)$.*

Proof. Given $a, b \geqslant 0$, convexity implies that $\left(\frac{a+b}{2}\right)^p \leqslant \frac{1}{2}(a^p + b^p)$ 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

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

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

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

Definition 1.29. With $\|f\|_{L^\infty(\Omega)} = \inf\left\{M \geqslant 0 \,\middle|\, |f(x)| \leqslant M \text{ a.e.}\right\}$, we set

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

Theorem 1.30. *$(L^\infty(\Omega), \|\cdot\|_{L^\infty(\Omega)})$ 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 1.31. 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 1.32 (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), let $\lambda \in [0,1]$ and for $f \in L^q(\Omega)$,

$$\begin{aligned}\|f\|_{L^q(\Omega)} = \Big(\int_\Omega |f|^q dx\Big)^{\frac{1}{q}} &= \Big(\int_\Omega |f|^{\lambda q}|f|^{(1-\lambda)q} d\mu\Big)^{\frac{1}{q}} \\ &\leqslant \Big(\|f\|^{\lambda q}_{L^p(\Omega)}\|f\|^{(1-\lambda)q}_{L^r(\Omega)}\Big)^{\frac{1}{q}} = \|f\|^{\lambda}_{L^p(\Omega)}\|f\|^{(1-\lambda)}_{L^r(\Omega)}.\end{aligned}$$

▫

Theorem 1.33. *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 1.34. *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 $\mathbb{R}^{\mathrm{n}}$ with $\mu(E_i) < \infty$, is dense in $L^p(\Omega)$.*

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 1.35. *Suppose that $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ 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

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

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

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

Next, let $A \subseteq \Omega$ be any measurable set, and let λ denote the Lebesgue measure. Then

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

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 1.34, 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 1.36 (Mollifiers)**.** Define $\eta \in \mathscr{C}^\infty(\mathbb{R}^n)$ by

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

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 $\operatorname{spt}(\eta_\epsilon) \subseteq \overline{B(0,\epsilon)}$.

Definition 1.37. 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 the open set $\widetilde{\Omega}$ where $\widetilde{\Omega} \subset\subset \Omega$, meaning that there exists K compact such that $\widetilde{\Omega} \subseteq K \subseteq \Omega$. We say that $\widetilde{\Omega}$ is compactly contained in Ω.

Definition 1.38 (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 1.39 (Commuting the derivative with the integral)**.** *Let $\Omega \subseteq \mathbb{R}^n$ 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

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

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 1.40 (Mollification of $L^p(\Omega)$ functions)**.** *If for $p \in [1,\infty)$, $f \in L^p_{\mathrm{loc}}(\Omega)$ 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 $L^p_{\mathrm{loc}}(\Omega)$.*

Proof. **Part (A).** This follows from repeated application of Lemma 1.39. To see that $\frac{\partial u^\epsilon}{\partial x_i}(x)$ exists and is continuous for each $x \in \Omega_\epsilon$ and $i =, ..., n$, we show that

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

From Definition 1.36, η_ϵ is a smooth function; hence, since $u \in L^1_{\rm loc}(\Omega)$, we see that $y \mapsto \frac{\partial}{\partial x_i}\eta_\epsilon(x-y)u(y) \in L^1_{\rm loc}(\Omega)$ uniformly in $x \in \omega$ for any set $\omega \subset\subset \Omega$. Application of Lemma 1.39 then shows that $u^\epsilon \in \mathscr{C}^1(\Omega_\epsilon)$. An induction argument then shows that $u^\epsilon \in \mathscr{C}^\infty(\Omega_\epsilon)$.

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

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

To see this, note that

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

so that for $\epsilon>0$ sufficiently small

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

Since by Lemma 1.35, $\mathscr{C}^0(\omega)$ is dense in $L^p(\omega)$, choose $g \in \mathscr{C}^0(\omega)$ such that $\|f-g\|_{L^p(\omega)} < \delta$; thus

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

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 1.38, is one example of the use of integral operators; the Fourier transform is another.

Definition 1.41. 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 1.49, 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 1.42. *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 1.42, we will need another well-known inequality.

Lemma 1.43 (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 1.44 (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 1.43 by setting

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

□

Proof of Theorem 1.42. According to Lemma 1.43, $|f(y)g(x)| \leqslant \frac{|f(y)|^p}{p} + \frac{|g(x)|^q}{q}$ so that

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

To improve this bound, notice that

$$\begin{aligned} &\Big|\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} k(x,y)f(y)g(x)dydx\Big| \\ &\qquad \leqslant \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|k(x,y)|}{p}dx|tf(y)|^p dy + \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\frac{|k(x,y)|}{q}dy|t^{-1}g(x)|^q dx \\ &\qquad \leqslant \frac{C_1 t^p}{p}\|f\|^p_{L^p(\Omega)} + \frac{C_2 t^{-q}}{q}\|g\|^q_{L^q(\Omega)} =: F(t)\,. \end{aligned}$$

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

Theorem 1.45 (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 1.42. Then according to Theorem 1.42, $K_k : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$ and $\|K_k\|_{\mathscr{B}(L^p(\mathbb{R}^n), L^p(\mathbb{R}^n))} \leqslant C_1$. □

Theorem 1.42 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 1.46 (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 $L^p(\Omega)$. For $\phi \in L^p(\Omega)'$, the *operator norm* of ϕ is defined by $\|\phi\|_{\mathrm{op}} = \sup_{L^p(\Omega)=1} |\phi(f)|$.

Theorem 1.47. *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*

$$F_g(f) = \int_\Omega f g dx .$$

Then F_g is a continuous linear functional on $L^p(\Omega)$ with operator norm $\|F_g\|_{\mathrm{op}} = \|g\|_{L^q(\Omega)}$.

Proof. The linearity of F_g again follows from the linearity of the Lebesgue integral. Since

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

with the last inequality following from Hölder's inequality, we have that $\sup_{\|f\|_{L^p(\Omega)}=1} |F_g(f)| \leqslant \|g\|_{L^q(\Omega)}$.

For the reverse inequality let $f = |g|^{q-1}\operatorname{sgn}(g)$. 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 $\|g\|_{L^q(\Omega)} = \dfrac{F_g(f)}{\|f\|_{L^p(\Omega)}} \leqslant \|F_g\|_{\mathrm{op}}$. □

Remark 1.48. Theorem 1.47 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 1.49 (Representation theorem)**.** *Suppose that $1 < p < \infty$ 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 1.50. *For $p \in (1,\infty)$ the space $L^p(\Omega,\mu)$ is reflexive; that is, $L^p(\Omega)'' = L^p(\Omega)$.*

The proof Theorem 1.49 crucially relies on the Radon-Nikodym theorem, whose statement requires the following definition.

Definition 1.51. 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 1.52 (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 Riesz representation theorem, 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 1.53 (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 $\int_\Omega F\, d\nu = \int_\Omega F h\, d\mu$.

Lemma 1.54 (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*

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

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 1.49. 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 1.34, 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 1.49 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 1.55 (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 1.56. 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 1.57 (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 $\left\{f_{k_j}\right\}_{j=1}^\infty$ such that $f_{k_j} \rightharpoonup f$ in $L^p(\Omega)$ weakly.*

Definition 1.58. 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 1.59 (Weak-* compactness for L^∞)**.** *If $\{f_k\}_{k=1}^\infty$ is a bounded sequence in $L^\infty(\Omega)$, then there exists a subsequence $\left\{f_{k_j}\right\}_{j=1}^\infty$ such that $f_{k_j} \overset{*}{\rightharpoonup} f$ in $L^\infty(\Omega)$ weak-*.*

Lemma 1.60. *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 1.61. If $p = 2$, let $\{f_k\}_{k=1}^\infty$ denote any orthonormal sequence in $L^2(\Omega)$. From Bessel's inequality

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

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

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

Example 1.62. 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 $\displaystyle\int_0^{2\pi} \sin(kx)v(x)dx \to 0$ as $k \to \infty$ for all $v \in L^2(0, 2\pi)$. By Theorem 1.40, 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 $M \geqslant 0$, $\displaystyle\max_{x\in[0,2\pi]} \Big(|v(x)| + \Big|\frac{dv}{dx}(x)\Big|\Big) \leqslant M$. Then

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

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

On the other hand,

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

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 1.62 is an example of a more general fact that periodic functions weakly converge to their average as the wavelength tends to zero (see Problem 1.12).

Example 1.63. 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 1.34, 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 1.62.

There are essentially three types of mechanisms by which a sequence $u_k \rightharpoonup u$ in $L^p(\Omega)$ but $u_k \not\to u$ 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 1.61, we see that the map $f \mapsto \|f\|_{L^p(\Omega)}$ is continuous, but not weakly continuous. It is, however, weakly lower-semicontinuous.

Theorem 1.64. *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 1.49,

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

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

Theorem 1.65. *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, Theorem **??**, by identifying f_k with an element ϕ_k of $L^q(\Omega)'$, and using Theorem 1.49 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 answer a variety of questions about convergence of sequences of functions.

Theorem 1.66 (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. □

Theorem 1.67. *Suppose that $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ is a bounded domain, and*

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

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

Proof. For notational clarity, we will use $\mu(E)$ to denote the Lebesgue measure of a set E.

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 $g \in L^q(\Omega)$, $q = \frac{p}{p-1}$. 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 1.68. The proof of Theorem 1.67 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 1.69. 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 1.70. If $u_k \rightharpoonup u$ in $L^p(\Omega)$ and $v_k \to v$ in $L^q(\Omega)$, then $\int_\Omega u_k v_k dx \to \int_\Omega u v dx$.

REMARK 1.71. 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 Ω is bounded, 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.)

Problem 1.4. Theorem 1.28 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. (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 $L^p(\Omega)$.

(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 1.6. 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 $\big\{f \in L^p(\Omega) \cap L^q(\Omega) \,\big|\, \|f\|_{L^q(\Omega)} \leqslant 1\big\}$ 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 1.7. 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 1.8. 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 1.9. 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 1.10. 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 $a_\epsilon(x) = a\big(\frac{x}{\epsilon}\big)$, 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 1.11. 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 1.12. 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 C.11) and the sets

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

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

Problem 1.13. 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 C.16) to prove that $v \in L^{\frac{pq}{p-q}}(\Omega)$.

Problem 1.14. Prove that the space $\mathscr{C}_c^0(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for any $p \in [1, \infty)$. (We use the notation $\mathscr{C}_c^0(\mathbb{R}^n)$ to denote the space of continuous functions on $\mathbb{R}^n$ with compact support.)

Problem 1.15. 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 $\Omega = \bigcup_{\alpha \in A} \Omega_\alpha$. 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 $\Omega^\complement = \overline{\left\{x \in \mathbb{R}^n \,\middle|\, u(x) \neq 0\right\}}$.

(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 1.16. 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 1.17.
(a) Let $u \in \mathscr{C}_c^0(\mathbb{R}^n)$ and $v \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. 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 $k \in \mathbb{N}$, $u \in \mathscr{C}_c^k(\mathbb{R}^n)$, 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 1.18.
(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 and open and smooth subset of $\mathbb{R}^n$. Prove that $\mathscr{C}_c^\infty(\Omega)$ is dense in $L^p(\Omega)$ for $1 \leqslant p < \infty$.

Problem 1.19. Prove that if $u \in L^1_{\mathrm{loc}}(\Omega)$ satisfies $\int_\Omega u(x)v(x)dx = 0$ for all $v \in \mathscr{C}_c^\infty(\Omega)$, then $u = 0$ a.e. in Ω.

Problem 1.20. 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 1.21. Let $u \in L^\infty(\mathbb{R}^{\mathrm{n}})$ and let η_ϵ be the mollifiers from Definition 1.36. 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 1.22.

(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

Sobolev Spaces and the Fourier Transform

2.1 Sobolev Spaces $W^{k,p}(\Omega)$ for Integers $k \geqslant 0$

2.1.1 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 $w = f e_i$, 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**:

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

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**:

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

2.1.2 Weak Derivatives

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

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

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 referblack 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 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 $v \in L^1_{\text{loc}}(\Omega)$ 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

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

where Du is the weak derivative of u.

Example 2.9. Let n = 1 and set $\Omega = (0,1)$. Define the function $f(x) = \sin \frac{1}{x}$. Then $u \in L^1(0,1)$ and $\frac{du}{dx} = -\frac{\cos(1/x)}{x^2} \in L^1_{\text{loc}}(0,1)$, 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 $u \in L^1_{\text{loc}}(\Omega)$. 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 $\lim_{\delta\to 0} \delta^{1-\alpha} \int_0^{2\pi} \varphi(x) n(x) d\theta = 0$ if $\alpha < 1$, we see that

$$\lim_{\delta\to 0} \int_{\Omega_\delta} |x|^{-\alpha} D\varphi(x)\, dx = \lim_{\delta\to 0} \alpha \int_{\Omega_\delta} |x|^{-\alpha-2} x\, \varphi(x)\, dx$$

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*

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

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

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

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

□

2.1.3 Definition of Sobolev Spaces

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

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

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 1.30), 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

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

2.1.4 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 $x \in \operatorname{spt}(u)$,

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

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

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

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

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,

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

It follows that $\partial_r^k = \sum_{|\alpha|\leqslant k} a_\alpha(\theta)D^\alpha$, 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 $kp>2$. □

2.1.5 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 $W^{k,p}_{\text{loc}}(\Omega)$ 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 1.36. Then

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

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

Proof. Theorem 1.40 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 $x \in \Omega_\epsilon$,

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

By part (D) of Theorem 1.40, $D^\alpha u^\epsilon \to v^\alpha$ in $L^p_{\mathrm{loc}}(\Omega)$. □

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 $\Omega \subseteq \mathbb{R}^n$ 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

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

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 $\mathrm{spt}(\zeta_k u) \subseteq \omega_k$. 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 $\mathrm{spt}(u^{\epsilon_k}) \subseteq \Omega_{k+4} \backslash \overline{\Omega_k}$.

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

Theorem 2.23 (Global approximation on $\overline{\Omega}$)**.** *Suppose that $\Omega \subseteq \mathbb{R}^n$ 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*

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

Proof. We employ Theorem 2.41 (which will be proven below) to obtain an extension Eu of u such that

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

Choose $v_j \in \mathscr{C}_c^\infty(\mathbb{R}^n)$ so that $v_j \to Eu$ in $W^{k,p}(\mathbb{R}^n)$, 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

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

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

2.1.6 Hölder Spaces

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

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

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.

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

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

Lemma 2.30. *Let $B_r \subseteq \mathbb{R}^{\rm n}$ 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 $x \in B_r$,*

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

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

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

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,

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

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,

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

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 $u \in \mathscr{C}^1_c(\mathbb{R}^{\mathrm{n}})$. 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,

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

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

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 $p > \mathrm{n}$,

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

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

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

where

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

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

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

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

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 2.5, 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. $\square$

2.1.8 The Gagliardo-Nirenberg-Sobolev Inequality

In the previous section, we consideblack 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 $u \in \mathscr{C}^1_c(\mathbb{R}^2)$.

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

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|u(x)|^{\frac{\mathrm{n}}{\mathrm{n}-1}}\,dx_1\,dx_2 \leqslant \Big(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|Du|\,dx_1dy_2\Big)^{\frac{1}{\mathrm{n}-1}}\int_{-\infty}^{\infty}\prod_{\substack{i=1\\ i\neq 2}}^{\mathrm{n}} I_i^{\frac{1}{\mathrm{n}-1}}\,dx_2\,,$$

where

$$I_1=\int_{-\infty}^{\infty}|Du|dy_1\,,\quad I_i=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|Du|dx_1dy_i \text{ for } i=3,\cdots,\mathrm{n}\,.$$

Applying Hölder's inequality, we find that

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|u(x)|^{\frac{\mathrm{n}}{\mathrm{n}-1}}\,dx_1dx_2\\
&\qquad \leqslant \Big(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|Du|\,dx_1dy_2\Big)^{\frac{1}{\mathrm{n}-1}}\Big(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|Du|dy_1\,dx_2\Big)^{\frac{1}{\mathrm{n}-1}}\times\\
&\qquad\qquad \times\prod_{i=3}^{\mathrm{n}}\Big(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|Du|\,dx_1\,dx_2dy_i\Big)^{\frac{1}{\mathrm{n}-1}}\,.
\end{aligned}$$

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 $1\leqslant q<\infty$,*

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

Proof. 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. Define $U(r,\theta) := g(r)u(r,\theta)$ or equivalently, $U(y) = g(|x-y|)u(y)$. 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 obtain:

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

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

Using Young's inequality from Theorem 1.46, 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

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

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

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

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

Example 2.37. 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

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

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

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

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

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

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

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

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

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

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 2.38 (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*

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

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

Theorem 2.39 (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*

$$\|u\|_{\mathscr{C}^{k-1-[\frac{\mathrm{n}}{p}],1+[\frac{\mathrm{n}}{p}]-\frac{\mathrm{n}}{p}}(\mathbb{R}^{\mathrm{n}})} \leqslant C\|u\|_{W^{k,p}(\mathbb{R}^{\mathrm{n}})} \text{ for a constant } C = C(k,p,\mathrm{n})\,. \tag{2.18}$$

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

Theorem 2.40 (Interpolation inequality for $W^{k,p}(\mathbb{R}^{\mathrm{n}})$)**.** *Let* n *be a given positive integer, and* p,q,r,j,k,ℓ,α *satisfy the relations*

$$\begin{aligned} &j \leqslant k \lneqq \ell\,,\ \frac{1}{p}-\frac{k}{\mathrm{n}} = \alpha\Big(\frac{1}{q}-\frac{\ell}{\mathrm{n}}\Big) + (1-\alpha)\Big(\frac{1}{r}-\frac{j}{\mathrm{n}}\Big),\\ &p,q,r \geqslant 1,\ 0<\alpha\leqslant 1,\ \frac{1}{p}-\frac{k}{\mathrm{n}} > \frac{1}{q}-\frac{\ell}{\mathrm{n}} \geqslant \frac{1}{p}-\frac{k+1}{\mathrm{n}}. \end{aligned} \tag{2.19}$$

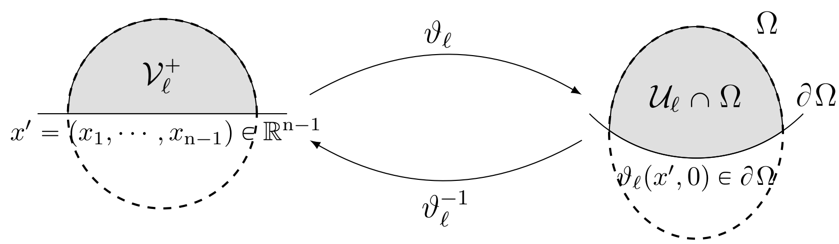
Then there exists a generic constant $C = C(p,q,r,j,k,\ell)$ *such that*

$$\|u\|_{W^{k,p}(\mathbb{R}^{\mathrm{n}})} \leqslant C\|u\|^{\alpha}_{W^{\ell,q}(\mathbb{R}^{\mathrm{n}})}\|u\|^{1-\alpha}_{W^{j,r}(\mathbb{R}^{\mathrm{n}})} \quad \forall\, u \in W^{\ell,q}(\mathbb{R}^{\mathrm{n}}) \cap W^{j,r}(\mathbb{R}^{\mathrm{n}})\,. \tag{2.20}$$

2.1.9 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 centeblack 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

$$\begin{aligned} \vartheta_\ell : \mathcal{V}_\ell \to \mathcal{U}_\ell \ \text{ is a } \mathscr{C}^1 \text{ diffeomorphism}\,, \\ \vartheta_\ell(\mathcal{V}_\ell^+) = \mathcal{U}_\ell \cap \Omega\,, \\ \vartheta_\ell(\mathcal{V}_\ell \cap \{x_{\mathrm{n}} = 0\}) = \mathcal{U}_\ell \cap \partial\Omega\,. \end{aligned} \tag{2.21}$$



2.1.10 Sobolev Extensions and Traces

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

Theorem 2.41. *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 2.42. *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, $B(z,r) \cup \partial\Omega \subseteq \{x_{\rm n} = 0\}$. Let $0 \leqslant \xi \in \mathscr{C}_c^\infty(B(z,r))$ such that $\xi = 1$ on $B(z,r/2)$. Set $\Gamma = \partial\Omega \cup B(z,r/2)$, $B^+(z,r) = B(z,r) \cap \Omega$, and let $dx_h = dx_1 \cdots dx_{{\rm n}-1}$. Then

$$\begin{aligned}
\int_\Gamma |u|^p\,dx_h \leqslant \int_{\{x_{\rm n}=0\}} \xi |u|^p\,dx_h &= -\int_{B^+(z,r)} \frac{\partial}{\partial x_{\rm n}}(\xi|u|^p)\,dx \\
&\leqslant -\int_{B^+(z,r)} \frac{\partial\xi}{\partial x_{\rm n}}|u|^p\,dx - p\int_{B^+(z,2\delta)} \xi|u|^{p-2}u\frac{\partial u}{\partial x_{\rm n}}dx \\
&\leqslant C\int_{B^+(z,r)} |u|^p\,dx + C\||u|^{p-1}\|_{L^{\frac{p}{p-1}}(B^+(z,r))}\left\|\frac{\partial u}{\partial x_{\rm n}}\right\|_{L^p(B^+(z,r))} \\
&\leqslant C\int_{B^+(z,r)} (|u|^p + |Du|^p)\,dx\,. \qquad (2.22)
\end{aligned}$$

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 $z \in \partial\Omega \cup U_\ell$ for some $\ell \in \{1,...,K\}$ and consider the $\mathscr{C}^1$-chart ϑ_ℓ defined in (2.21). Define the function $U = u \circ \vartheta_\ell$; then $U : \mathcal{V}_\ell^+ \to \mathbb{R}$. Setting $\Gamma = \mathcal{V}_\ell \cup \{x_{\rm n} = 0\}$, we see from the inequality (2.22) 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

$$\int_{\partial\Omega} |u|^p dS \leqslant C\int_\Omega (|u|^p + |Du|^p)dx\,. \qquad (2.23)$$

The inequality (2.23) 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 (2.23),

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

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 2.41. 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.

Integration by parts for functions in $H^1(\Omega)$

Finally, we finish this section by stating the following theorem which is the generalization of (2.3) and the divergence theorem.

Theorem 2.43. *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 2.42 implies that $u_k \to u$ and $v_k \to v$ in $L^2(\partial\Omega)$. Therefore, the divergence theorem implies that

$$\begin{aligned}\int_\Omega \frac{\partial u}{\partial x_i} v\, dx &= \lim_{k\to\infty} \int_\Omega \frac{\partial u_k}{\partial x_i} v_k\, dx = \lim_{k\to\infty}\Big[\int_{\partial\Omega} u_k v_k \mathrm{N}_i\, dS - \int_\Omega \frac{\partial u_k}{\partial x_i} v_k\, dx\Big] \\ &= \int_{\partial\Omega} uv\mathrm{N}_i\, dS - \int_\Omega \frac{\partial u}{\partial x_i} v\, dx\,. \end{aligned}$$ □

2.1.11 The subspace $W_0^{1,p}(\Omega)$

Definition 2.44. We let $W_0^{1,p}(\Omega)$ denote the closure of $\mathscr{C}_c^\infty(\Omega)$ in $W^{1,p}(\Omega)$.

Theorem 2.45. *Suppose that $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ is bounded with $\mathscr{C}^1$-boundary, and that $u \in W^{1,p}(\Omega)$. Then*

$$u \in W_0^{1,p}(\Omega) \;\; \textit{if and only if} \;\; \tau u = 0 \;\; \textit{on} \;\; \partial\Omega\,.$$

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 which satisfy

$$\vartheta_\ell : B(0, r_\ell) \subseteq \mathbb{R}^{\mathrm{n}-1} \to \mathcal{U}_\ell \cap \partial\Omega \;\; \text{is a } \mathscr{C}^1\text{-diffeomorphism}\,.$$

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 $\sum_{\ell=0}^K \xi_\ell = 1$. 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 $\sqrt{\zeta_\ell} \in \mathscr{C}_c^\infty(\mathbb{R}^{\rm n})$ 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 $u^{(\ell)} \in W^{1,p}(\mathbb{R}^{\rm n}_+)$ and $\tau u^{(\ell)} = 0$ for all $\ell \in \{1, \cdots, K\}$. By definition of the trace, for each ℓ there exists a sequence $\{u_k^{(\ell)}\}_{k=1}^\infty \subseteq \mathscr{C}^\infty(\overline{\mathbb{R}^{\rm n}_+})$ such that $u_k^{(\ell)} \to u^{(\ell)}$ in $W^{1,p}(\mathbb{R}^{\rm n}_+)$ and $\tau u_k^{(\ell)} = u_k^{(\ell)}|_{\mathbb{R}^{\rm n-1}} \to 0$ in $L^p(\mathbb{R}^{\rm n-1})$ as $k \to \infty$. Note that for $x_h \in \mathbb{R}^{\rm n-1}$ and $x_{\rm n} \geqslant 0$,

$$u_k^{(\ell)}(x_h, x_{\rm n}) = u_k^{(\ell)}(x_h, 0) + \int_0^{x_{\rm n}} u_{k,{\rm n}}^{(\ell)}(x_h, t)\, dt\,;$$

thus Hölder's inequality implies that

$$\begin{aligned} &\int_{\mathbb{R}^{\rm n-1}} \big|u_k^{(\ell)}(x_h, x_{\rm n})\big|^p\, dx_h \\ &\qquad \leqslant C\Big[\int_{\mathbb{R}^{\rm n-1}} \big|u_k^{(\ell)}(x_h, 0)\big|^p\, dx_h + x_{\rm n}^{p-1}\int_0^{x_{\rm n}}\int_{\mathbb{R}^{\rm n-1}} \big|Du_k^{(\ell)}(x_h, t)\big|^p\, dx_h dt\Big]. \end{aligned}$$

Passing to the limit as $k \to \infty$, by the fact that $u_k^{(\ell)} \to u^{(\ell)}$ in $W^{1,p}(\mathbb{R}^{\rm n}_+)$ and $\tau u_k^{(\ell)} \to 0$ in $L^p(\mathbb{R}^{\rm n-1})$ we find that

$$\int_{\mathbb{R}^{\rm n-1}} \big|u^{(\ell)}(x_h, x_{\rm n})\big|^p\, dx_h \leqslant C x_{\rm n}^{p-1}\int_0^{x_{\rm n}}\int_{\mathbb{R}^{\rm n-1}} \big|Du^{(\ell)}(x_h, t)\big|^p\, dx_h dt\,. \tag{2.24}$$

Let $\chi \in \mathscr{C}^\infty(\mathbb{R}_+)$ satisfy

$$\chi = 1 \;\text{ on }\; [0, 1]\,, \quad \chi = 0 \;\text{ on }\; [2, \infty)\,, \quad \text{and} \quad 0 \leqslant \chi \leqslant 1\,.$$

Define $\chi_k(x) = \chi(kx_{\rm n})$ for $x \in \mathbb{R}^{\rm n}_+$ as well as $v_k^{(\ell)} = (1 - \chi_k)u^{(\ell)}$. Then using (2.24),

$$\begin{aligned} &\int_{\mathbb{R}^{\rm n}_+} \big|Dv_k^{(\ell)}(x) - Du^{(\ell)}(x)\big|^p\, dx \\ &\qquad \leqslant C\Big[\int_{\mathbb{R}^{\rm n}_+} \big|\chi_k(x)\big|^p \big|Du^{(\ell)}(x)\big|^p\, dx + \int_{\mathbb{R}^{\rm n}_+} \big|D\chi_k(x)\big|^p \big|u^{(\ell)}(x)\big|^p\, dx\Big] \\ &\qquad \leqslant C\Big[\int_{\mathbb{R}^{\rm n}_+} \big|\chi_k(x)\big|^p \big|Du^{(\ell)}(x)\big|^p\, dx + k^p\int_0^{\frac{2}{k}}\int_{\mathbb{R}^{\rm n-1}} \big|u^{(\ell)}(x_h, t)\big|^p\, dx_h dt\Big] \\ &\qquad \leqslant C\Big[\int_{\mathbb{R}^{\rm n}_+} \big|\chi_k(x)\big|^p \big|Du^{(\ell)}(x)\big|^p\, dx + \int_0^{\frac{2}{k}}\int_{\mathbb{R}^{\rm n-1}} \big|u^{(\ell)}(x_h, t)\big|^p\, dx_h dt\Big] \end{aligned}$$

which converges to 0 as $k \to \infty$. In other words, $\big\{Dv_k^{(\ell)}\big\}_{k=1}^\infty$ converges to $Du^{(\ell)}$ in $L^p(\mathbb{R}^{\rm n}_+)$. It is also clear that $\big\{v_k^{(\ell)}\big\}_{k=1}^\infty$ converges to $u^{(\ell)}$ in $L^p(\mathbb{R}^{\rm n}_+)$ since

$$\big\|v_k^{(\ell)} - u^{(\ell)}\big\|_{L^p(\mathbb{R}^{\rm n}_+)} = \big\|\chi_k u^{(\ell)}\big\|_{L^p(\mathbb{R}^{\rm n}_+)} \leqslant \big\|u^{(\ell)}\big\|_{L^p(\mathbb{R}^{\rm n-1}\times[0,\frac{2}{k}))}\,.$$

Define $u_k = \zeta_0 u + \sum_{\ell=1}^K \sqrt{\zeta_\ell}(v_k^{(\ell)} \circ \vartheta_\ell^{-1})$. 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 the Sobolev embedding theorems for bounded domains Ω.

Theorem 2.46 (Gagliardo-Nirenberg inequality for $W^{1,p}(\Omega)$)**.** *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*

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

Proof. Choose $\widetilde{\Omega} \subseteq \mathbb{R}^{\mathrm{n}}$ bounded such that $\Omega \subset\subset \widetilde{\Omega}$, and let Eu denote the Sobolev extension of u to $\mathbb{R}^{\mathrm{n}}$ such that $Eu = u$ a.e., $\operatorname{spt}(Eu) \subseteq \widetilde{\Omega}$, and $\|Eu\|_{W^{1,p}(\mathbb{R}^{\mathrm{n}})} \leqslant C\|u\|_{W^{1,p}(\Omega)}$. Then by the Gagliardo-Nirenberg inequality,

$$\|u\|_{L^{\frac{\mathrm{n}p}{\mathrm{n}-p}}(\Omega)} \leqslant \|Eu\|_{L^{\frac{\mathrm{n}p}{\mathrm{n}-p}}(\mathbb{R}^{\mathrm{n}})} \leqslant C\|D(Eu)\|_{L^p(\mathbb{R}^{\mathrm{n}})} \leqslant C\|Eu\|_{W^{1,p}(\mathbb{R}^{\mathrm{n}})} \leqslant C\|u\|_{W^{1,p}(\Omega)}\,. \quad ▫$$

By following the proof of Theorem 2.35, we have the following generalization for integers $k \geqslant 1$:

Theorem 2.47 (Gagliardo-Nirenberg-Sobolev inequality for $W^{k,p}(\Omega)$)**.** *Suppose that $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ is open and bounded with $\mathscr{C}^1$-boundary, and $1 \leqslant kp < \mathrm{n}$. Then there exists a generic constant $C = C(k, p, \mathrm{n}, \Omega)$ such that*

$$\|u\|_{L^{\frac{\mathrm{n}p}{\mathrm{n}-kp}}(\Omega)} \leqslant C\|u\|_{W^{k,p}(\Omega)} \qquad \forall\, u \in W^{k,p}(\Omega)\,. \tag{2.25}$$

In fact, 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 2.4.2. 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 2.48 (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 $1 \leqslant q \leqslant \frac{\mathrm{n}p}{\mathrm{n}-p}$,*

$$\|u\|_{L^q(\Omega)} \leqslant C\|Du\|_{L^p(\Omega)} \qquad \forall\, u \in W_0^{1,p}(\Omega)\,. \tag{2.26}$$

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 2.49. *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 $1 \leqslant q < \infty$,*

$$\|u\|_{L^q(\Omega)} \leqslant C\sqrt{q}\|Du\|_{L^2(\Omega)} \qquad \forall\, u \in H_0^1(\Omega)\,. \tag{2.27}$$

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 that assume that $\operatorname{spt}(u) \subseteq \operatorname{spt}(U)$ with U also defined in the proof Theorem 2.36. □

Remark 2.50. Inequalities (2.26) and (2.27) are commonly referblack to as *Poincaré inequalities.* They are 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 $u \in H_0^1(\Omega)$. In particular, there exists constants C_1, C_2 such that

$$C_1\|Du\|_{L^2(\Omega)} \leqslant \|u\|_{H^1(\Omega)} \leqslant C_2\|Du\|_{L^2(\Omega)}\,.$$

A more general form is given as follows:

Lemma 2.51 (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*

$$\|u - \bar{u}\|_{L^2(\Omega)} \leqslant C\|Du\|_{L^2(\Omega)} \quad \forall\, u \in H^1(\Omega)\,, \tag{2.28}$$

where $\bar{u} := \fint_\Omega u(y)dy$ denotes the average value of u over Ω.

Proof. Suppose for the sake of contradiction that (2.28) does not hold. Then there is a sequence $\{u_j\}_{j=1}^\infty \subseteq H^1(\Omega)$ satisfying

$$\|u_j - \bar{u}_j\|_{L^2(\Omega)} > j\|Du_j\|_{L^2(\Omega)}\,, \tag{2.29}$$

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 (2.29), $\|Dw_j\|_{L^2(\Omega)} < j^{-1}$, so that $\|w_j\|^2_{H^1(\Omega)} < 1 + j^{-2} < \infty$. Strong compactness, given by Theorem 2.65 (see also Theorem 2.149) 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 $\|w\|_{L^2(\Omega)} = 1$. □

Corollary 2.52. *Whenever $\bar{u} = \fint_\Omega u(y)dy = 0$, $\|Du\|_{L^2(\Omega)}$ 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 2.53 (Poincaré inequality for $H_0^1(\Omega)$). *Let $\Omega \subseteq \mathbb{R}^n$ denote an open, bounded, connected, and smooth domain. Then*

$$\|u\|_{L^2(\Omega)} \leqslant C\|Du\|_{L^2(\Omega)} \qquad \forall\, u \in H_0^1(\Omega)\,, \tag{2.30}$$

where the constant C depends on Ω.

Lemma 2.54 (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*

$$\|u\|_{L^2(\Omega)} \leqslant C\big(\|\sqrt{k}u\|_{L^2(\partial\Omega)} + \|Du\|_{L^2(\Omega)}\big) \qquad \forall\, u \in H^1(\Omega)\,, \tag{2.31}$$

where the constant C depends on Ω.

Integration by parts for functions in $W_0^{1,p}(\Omega)$

Having established Theorem 2.45, using the density argument we can conclude the following

Theorem 2.55. *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.

2.1.12 Weak Solutions to Dirichlet's 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*:

$$-\Delta u = f \qquad \text{in} \quad \Omega\,, \tag{2.32a}$$

$$u = 0 \qquad \text{on} \quad \partial\Omega\,, \tag{2.32b}$$

where $\Delta = \sum_{i=1}^{\mathrm{n}} \frac{\partial^2}{\partial x_i^2}$ denotes the Laplace operator or *Laplacian.* As written, (2.32) 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 (2.32) 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 2.56. 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 2.57 (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 $v \in H_0^1(\Omega)$,*

$$\langle f, v \rangle = \int_\Omega \Big[f_0(x) v(x) + \sum_{i=1}^n f_i(x) \frac{\partial v}{\partial x_i}(x) \Big] dx \,, \tag{2.33}$$

and

$$\|f\|_{H^{-1}(\Omega)} = \inf \Big\{ \Big(\int_\Omega \sum_{j=0}^n |f_j(x)|^2 \, dx \Big)^{\frac{1}{2}} \,\Big|\, f \text{ satisfying } (2.33) \Big\} . \tag{2.34}$$

Proof. By the Riesz Representation Theorem, for every $f \in H^{-1}(\Omega)$ there exists $u \in H_0^1(\Omega)$ satisfying

$$(u, v)_{L^2(\Omega)} + (Du, Dv)_{L^2(\Omega)} = \langle f, v \rangle \qquad \forall\, v \in H_0^1(\Omega) . \tag{2.35}$$

Letting $f_0 = u$ and $f_i = \partial u / \partial x_i$ for $i = 1, ..., n$ gives the relation (2.33).

Then for $f \in H^{-1}(\Omega)$, we may write

$$\langle f, v \rangle = \int_\Omega \Big[g_0(x) v(x) + \sum_{i=1}^n g_i(x) \frac{\partial v}{\partial x_i}(x) \Big] dx \,, \tag{2.36}$$

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 (2.35) yields

$$\|u\|^2_{H^1(\Omega)} \leqslant \int_\Omega \sum_{j=0}^n |g_0(x)|^2 \, dx \, .$$

Hence, since $f_0 = u$ and $f_i = \partial u/\partial x_i$, we see that

$$\int_\Omega \sum_{j=0}^n |f_j(x)|^2\, dx \leqslant \int_\Omega \sum_{j=0}^n |g_j(x)|^2\, dx\,. \tag{2.37}$$

From (2.33), we infer that

$$\|f\|_{H^{-1}(\Omega)} \leqslant \Big(\int_\Omega \sum_{j=0}^n |f_j(x)|^2\, dx\Big)^{\frac{1}{2}} \text{ if } \|v\|_{H^1_0(\Omega)} \leqslant 1\,.$$

Thus, with $v = u\|u\|^{-1}_{H^1_0(\Omega)}$ in (2.35), we have that

$$\|f\|^2_{H^{-1}(\Omega)} = \int_\Omega \sum_{j=0}^n |f_j(x)|^2\, dx\,. \tag{2.38}$$

Then, (2.34) follows from (2.36)-(2.38). ▫

Definition 2.58. A function $u \in H^1_0(\Omega)$ is a weak solution of (2.32) if

$$\int_\Omega Du \cdot Dv\, dx = \langle f, v\rangle \qquad \forall\, v \in H^1_0(\Omega)\,.$$

Remark 2.59. 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 2.60. The motivation for Definition 2.58 is as follows. Since $\mathscr{C}^\infty_c(\Omega)$ is dense in $H^1_0(\Omega)$, multiply equation (2.32a) by $\varphi \in \mathscr{C}^\infty_c(\Omega)$, 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 2.61 (Existence and uniqueness of weak solutions)**.** *For any $f \in H^{-1}(\Omega)$, there exists a unique weak solution to* (2.32).

Proof. Using the Poincaré inequality, $\|Du\|_{L^2(\Omega)}$ is an H^1-equivalent norm for all $u \in H^1_0(\Omega)$, and $(Du, Dv)_{L^2(\Omega)}$ defines the inner-product on $H^1_0(\Omega)$. As such, according to the definition of weak solutions to (2.32), we are seeking $u \in H^1_0(\Omega)$ such that

$$(u, v)_{H^1_0(\Omega)} = \langle f, v\rangle \ \ \forall\, v \in H^1_0(\Omega)\,. \tag{2.39}$$

The existence of a unique $u \in H^1_0(\Omega)$ satisfying (2.39) is provided by the Riesz representation theorem for Hilbert spaces. ▫

Remark 2.62. Note that the Riesz representation theorem shows that there exists a distribution, denoted by $-\Delta u \in H^{-1}(\Omega)$ such that

$$\langle -\Delta u, v\rangle = \langle f, v\rangle \qquad \forall\, v \in H^1_0(\Omega)\,.$$

The operator $-\Delta : H^1_0(\Omega) \to H^{-1}(\Omega)$ is thus an isomorphism.

A fundamental question in the theory of linear partial differential equations is commonly referblack 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 2.6, where it will be shown that for integers $k \geqslant 2$, $-\Delta : H^k(\Omega) \cap H^1_0(\Omega) \to H^{k-2}(\Omega)$ is also an isomorphism. An important consequence of this result is that $(-\Delta)^{-1} : H^{k-2}(\Omega) \to H^k(\Omega) \cap H^1_0(\Omega)$ 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.

2.1.13 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 2.63. 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 1.55) then implies the following theorem.

THEOREM 2.64 (Weak compactness in $W^{1,p}(\Omega)$)**.** *Let $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ 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 2.65 (Rellich's theorem on a bounded domain Ω)**.** *Let $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ be an open, bounded domain with $\mathscr{C}^1$-boundary, and $1 \leqslant p < \mathrm{n}$. Then $W^{1,p}(\Omega)$ is compactly embedded in $L^q(\Omega)$ for all $1 \leqslant q < \frac{\mathrm{n}p}{\mathrm{n}-p}$; 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 $\mathrm{n} = 2$ 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 2.66 (Arzelà-Ascoli Theorem)**.** *Let $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$ 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 $\{u_{j_k}\}_{k=1}^{\infty}$ 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 $\mathbb{R}^{\mathrm{n}}$. 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 $1 \leqslant q < \frac{\mathrm{n}p}{\mathrm{n}-p}$. 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 $\operatorname{spt}(Eu_j) \subseteq \widetilde{\Omega}$, 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

$$\bar{u}_j^\epsilon = \int_{B(0,\epsilon)} \frac{1}{\epsilon^{\mathrm{n}}}\eta\big(\frac{y}{\epsilon}\big)\bar{u}_j(x-y)dy = \int_{B(0,1)} \eta(z)\bar{u}_j(x-\epsilon z)dz\,. \tag{2.40}$$

Applying the fundamental theorem of calculus to $\bar{u}_j$, we see that

$$\bar{u}_j(x-\epsilon z) - \bar{u}_j(x) = \int_0^1 \frac{d}{dt}\bar{u}_j(x-\epsilon tz)dt = -\epsilon\int_0^1 D\bar{u}_j(x-\epsilon tz)\cdot z\,dt\,. \tag{2.41}$$

Substitution of (2.41) into (2.40) shows that

$$\left|\bar{u}_j^\epsilon(x) - \bar{u}_j(x)\right| = \epsilon\int_{B(0,1)} \eta(z)\int_0^1 |D\bar{u}_j(x-\epsilon tz)|\,dzdt\,,$$

so that

$$\begin{aligned}\int_{\tilde{\Omega}} |\bar{u}_j^\epsilon(x) - \bar{u}_j(x)|\,dx &= \epsilon\int_{B(0,1)} \eta(z)\int_0^1\int_{\tilde{\Omega}} |D\bar{u}_j(x-\epsilon tz)|\,dxdzdt \\ &\leqslant \epsilon\|D\bar{u}_j\|_{L^1(\tilde{\Omega})} \leqslant \epsilon\|D\bar{u}_j\|_{L^p(\tilde{\Omega})} < \epsilon CM\,,\end{aligned}$$

where we have used the mean value theorem for integrals and Young's inequality for convolution for the first inequality above. Using the L^p-interpolation Lemma 1.19, for any $1 < q < \mathrm{n}p/(\mathrm{n}-p)$,

$$\begin{aligned}\|\bar{u}_j^\epsilon - \bar{u}_j\|_{L^q(\tilde{\Omega})} &\leqslant \|\bar{u}_j^\epsilon - \bar{u}_j\|^a_{L^1(\tilde{\Omega})}\|\bar{u}_j^\epsilon - \bar{u}_j\|^{1-a}_{L^{\frac{\mathrm{n}p}{\mathrm{n}-p}}(\tilde{\Omega})} \\ &\leqslant \epsilon^a CM^a\,\|D\bar{u}_j^\epsilon - D\bar{u}_j\|^{1-a}_{L^p(\tilde{\Omega})} \\ &\leqslant \epsilon^a CM\,.\end{aligned} \tag{2.42}$$

The inequality (2.42) shows that $\bar{u}^\epsilon_j$ 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 $\bar{u}^\epsilon_j$ to construct a convergent subsequence $\bar{u}^\epsilon_{j_k}$.

Step 3. Extracting a convergent subsequence. Our goal is to employ the Arzela-Ascoli Theorem, so we show that for $\epsilon > 0$ fixed,

$$\sup_j \|\bar{u}^\epsilon_j\|_{\mathscr{C}^0(\bar{\tilde{\Omega}})} \leqslant \widetilde{M} < \infty \quad \text{and} \quad \{\bar{u}^\epsilon_j\}_{j=1}^\infty \text{ is equi-continous.}$$

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

$$\begin{aligned}\sup_j \|\bar{u}^\epsilon_j\|_{\mathscr{C}^0(\bar{\tilde{\Omega}})} &\leqslant \sup_j \sup_{x \in \bar{\tilde{\Omega}}} \int_{B(x,\epsilon)} \eta_\epsilon(x-y)|\bar{u}_j(y)|dy \\ &\leqslant \|\eta_\epsilon\|_{L^\infty(\mathbb{R}^{\mathrm{n}})} \sup_j \|\bar{u}_j\|_{L^1(\tilde{\Omega})} \leqslant C\epsilon^{-\mathrm{n}} < \infty\,,\end{aligned}$$

and similarly

$$\sup_j \|\bar{D}u^\epsilon_j\|_{\mathscr{C}^0(\bar{\tilde{\Omega}})} \leqslant \|D\eta_\epsilon\|_{L^\infty(\mathbb{R}^{\mathrm{n}})} \sup_j \|\bar{u}_j\|_{L^1(\tilde{\Omega})} \leqslant C\epsilon^{-\mathrm{n}-1} < \infty\,.$$

The latter inequality proves equicontinuity of the sequence $\{\bar{u}^\epsilon_j\}_{j=1}^\infty$, and hence there exists a subsequence $\{u_{j_k}\}_{k=1}^\infty$ which converges uniformly on $\widetilde{\Omega}$, so that

$$\limsup_{k,\ell\to\infty} \|\bar{u}^\epsilon_{j_k} - \bar{u}^\epsilon_{j_\ell}\|_{L^q(\tilde{\Omega})} = 0\,.$$

Step 4. Diagonal argument. Now, fix $\delta > 0$ and choose ϵ sufficiently small in (2.42) 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 $\delta = \frac{1}{2}, \frac{1}{3}$, 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. $\square$

Weak convergence in $W^{1,p}(\Omega)$ for $1 < p < \infty$

If we know a priori that $W^{1,p}(\Omega)$ is reflexive, then the Alaoglu theorem (Theorem 1.55) can be used to study the weak-* convergence of bounded sequence in $W^{1,p}(\Omega)$ which in turn is equivalent to the weak convergence of 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 2.67 (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 $\boldsymbol{v} = (v_0, v_1, \cdots, v_{\rm n}) \in L^{p'}(\Omega)^{\rm n+1}$ such that*

$$\langle f, u\rangle = \int_\Omega u v_0 \, dx + \sum_{\ell=1}^{\rm n} \int_\Omega \frac{\partial u}{\partial x_\ell} v_\ell \, dx \,,$$

where $\langle \cdot, \cdot\rangle$ denotes the duality pairing between distributions in $W^{1,p}(\Omega)'$ and functions in $W^{1,p}(\Omega)$. Moreover,

$$\|f\|_{W^{1,p}(\Omega)'} = \sum_{\ell=0}^{\rm n} \|v_\ell\|_{L^p(\Omega)} \,.$$

Proof. Define a bounded linear map $\mathrm{P} : W^{1,p}(\Omega) \to L^p(\Omega)^{\rm n+1}$ by

$$\mathrm{P}u = \Big(u, \frac{\partial u}{\partial x_1}, \cdots, \frac{\partial u}{\partial x_{\rm n}}\Big) \,,$$

and let W be the range of P. Since P is an isometry, W is a closed subspace of $L^p(\Omega)^{\rm n+1}$. 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 $\widetilde{L} : L^p(\Omega)^{\rm n+1} \to \mathbb{R}$ satisfying

1. $\widetilde{L}(\boldsymbol{w}) = L\boldsymbol{w}$ for all $\boldsymbol{w} \in W$; 2. $\|\widetilde{L}\|_{L^p(\Omega)^{\rm n+1}{}'} = \|L\|_{W'}$.

By the Risez representation theorem (Theorem 1.49), there exists a unique $\boldsymbol{v} = (v_0, v_1, \cdots, v_{\rm n}) \in L^{p'}(\Omega)^{\rm n+1}$ such that

$$\widetilde{L}(\boldsymbol{w}) = \sum_{\ell=0}^{\rm n} \int_\Omega w_\ell v_\ell \, dx \qquad \forall\, \boldsymbol{w} = (w_0, w_1, \cdots, w_{\rm n}) \in L^p(\Omega)^{\rm n+1} \,,$$

and $\|\widetilde{L}\|_{L^p(\Omega)^{n+1\prime}} = \sum_{\ell=0}^{n} \|v_\ell\|_{L^p(\Omega)}$. In particular, we have

$$\langle f, u\rangle = L(\mathrm{P}u) = \widetilde{L}(\mathrm{P}u) = \int_\Omega uv_0\,dx + \sum_{\ell=1}^{n}\int_\Omega \frac{\partial u}{\partial x_\ell} v_\ell\,dx$$

which concludes the theorem. ▫

Remark 2.68. For the case $p = 2$, the existence of such a $\boldsymbol{v}$ in Theorem 2.67 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 1.57 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 2.69. *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 2.69 is in fact the weak convergence in $W^{1,p}(\Omega)$, we make use of Theorem 2.67. Let $\{u_k\}_{k=1}^\infty$ be a bounded sequence in $W^{1,p}(\Omega)$ and $f \in W^{1,p}(\Omega)'$. Then Theorem 2.67 provides a unique $\boldsymbol{v} = (v_0, v_1, \cdots, v_n) \in L^{p'}(\Omega)^{n+1}$ such that

$$\langle f, u\rangle = \int_\Omega uv_0\,dx + \sum_{\ell=1}^{n}\int_\Omega \frac{\partial u}{\partial x_\ell} v_\ell\,dx\,.$$

Therefore, the subsequence $\{u_{k_j}\}_{j=1}^\infty$ provides by Theorem 2.69 satisfies

$$\begin{aligned}\lim_{j\to\infty}\langle f, u_{k_j}\rangle &= \lim_{j\to\infty}\Big[\int_\Omega u_{k_j}v_0\,dx + \sum_{\ell=1}^{n}\int_\Omega \frac{\partial u_{k_j}}{\partial x_\ell} v_\ell\,dx\Big] = \int_\Omega uv_0\,dx + \sum_{\ell=1}^{n}\int_\Omega \frac{\partial u}{\partial x_\ell} v_\ell\,dx\\ &= \langle f, u\rangle\,.\end{aligned}$$

The argument above establishes the following

Theorem 2.70. *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 $W^{1,p}(\Omega)$.*

2.1.14 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 consideblack, and the goal is to show the weak convergence of the product of two weakly convergent sequence under certain additional constraints.

Theorem 2.71 (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 $\operatorname{div}\boldsymbol{u}_k$ and $\operatorname{curl}\boldsymbol{v}_k$ 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 $\operatorname{div}\boldsymbol{u}_k$ and $\operatorname{curl}\boldsymbol{v}_k$ in $H^{-1}(\mathbb{R}^{\rm n})$, there exists a subsequence $\{k_j\}_{j=1}^\infty$ such that

$$\begin{aligned} \operatorname{div}\boldsymbol{u}_{k_j} \to \operatorname{div}\boldsymbol{u} & \qquad \text{in} \quad H^{-1}(\mathbb{R}^{\rm n})\,, \\ \operatorname{curl}\boldsymbol{v}_{k_j} \to \operatorname{curl}\boldsymbol{v} & \qquad \text{in} \quad H^{-1}(\mathbb{R}^{\rm n})\,. \end{aligned}$$

Let $\varphi \in \mathcal{D}(\mathbb{R}^{\rm n})$ be given. Then

$$-\Delta(\varphi \operatorname{curl}\boldsymbol{w}_k) = -\operatorname{curl}\boldsymbol{w}_k(\varphi + \Delta\varphi) - 2\nabla\varphi \cdot \nabla \operatorname{curl}\boldsymbol{w}_k + \varphi \operatorname{curl}\boldsymbol{v}_k\,. \tag{2.43}$$

Since $\varphi + \Delta\varphi$ is compactly supported,

$$\operatorname{curl}\boldsymbol{w}_k(\varphi + \Delta\varphi) \to \operatorname{curl}\boldsymbol{w}(\varphi + \Delta\varphi) \qquad \text{in} \quad L^2(\mathbb{R}^{\rm n})\,.$$

For the second term on the right-hand side of (2.43), by the definition of the dual space norm,

$$
\begin{aligned}
\|\nabla\varphi\cdot\nabla\mathrm{curl}(\boldsymbol{w}_{k_j}-\boldsymbol{w})\|_{H^{-1}(\mathbb{R}^n)} &= \sup_{\substack{\psi\in H^1_0(\mathbb{R}^n)\\ \|\psi\|_{H^1(\mathbb{R}^n)}=1}} \langle\nabla\varphi\cdot\nabla\mathrm{curl}(\boldsymbol{w}_{k_j}-\boldsymbol{w}),\psi\rangle \\
&= \sup_{\substack{\psi\in H^1_0(\mathbb{R}^n)\\ \|\psi\|_{H^1(\mathbb{R}^n)}=1}} \Big[\langle\Delta\varphi\,\mathrm{curl}(\boldsymbol{w}_{k_j}-\boldsymbol{w}),\psi\rangle + \langle\nabla\varphi\otimes\mathrm{curl}(\boldsymbol{w}_{k_j}-\boldsymbol{w}),\nabla\psi\rangle\Big] \\
&\leqslant 2\|\nabla\varphi\|_{W^{1,\infty}(\mathbb{R}^n)}\|\mathrm{curl}(\boldsymbol{w}_{k_j}-\boldsymbol{w})\|_{L^2(\mathrm{spt}(\varphi))}\to 0 \quad\text{as}\quad j\to\infty\,.
\end{aligned}
$$

As a consequence, the right-hand side of (2.43) 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

$$
\|\varphi\,\mathrm{curl}(\boldsymbol{w}_{k_j}-\boldsymbol{w})\|_{H^1(\mathbb{R}^n)}\to 0 \quad\text{as}\quad j\to\infty\,. \tag{2.44}
$$

Finally, observing that

$$
\begin{aligned}
&\int_{\mathbb{R}^n}\boldsymbol{u}_{k_j}\cdot\boldsymbol{v}_{k_j}\varphi\,dx \\
&\quad= \int_{\mathbb{R}^n}\boldsymbol{u}_{k_j}\cdot\mathrm{curlcurl}\boldsymbol{w}_{k_j}\varphi\,dx - \int_{\mathbb{R}^n}\boldsymbol{u}_{k_j}\cdot\nabla\mathrm{div}\boldsymbol{w}_{k_j}\varphi\,dx + \int_{\mathbb{R}^n}\boldsymbol{u}_{k_j}\cdot\boldsymbol{w}_{k_j}\varphi\,dx \\
&\quad= \int_{\mathbb{R}^n}\boldsymbol{u}_{k_j}\cdot\mathrm{curl}(\varphi\,\mathrm{curl}\boldsymbol{w}_{k_j})\,dx - \int_{\mathbb{R}^n}\boldsymbol{u}_{k_j}\cdot(\nabla\varphi\times\mathrm{curl}\boldsymbol{w}_{k_j})\,dx \\
&\qquad+ \int_{\mathbb{R}^n}\mathrm{div}\boldsymbol{u}_{k_j}\mathrm{div}\boldsymbol{w}_{k_j}\varphi\,dx + \int_{\mathbb{R}^n}\boldsymbol{u}_{k_j}\cdot\nabla\varphi\,\mathrm{div}\boldsymbol{w}_{k_j}\,dx + \int_{\mathbb{R}^n}\boldsymbol{u}_{k_j}\cdot\boldsymbol{w}_{k_j}\varphi\,dx\,,
\end{aligned}
$$

we conclude that the right-hand side converges to corresponding terms without k_j; thus it is clear that

$$
\lim_{j\to\infty}\int_{\mathbb{R}^n}\boldsymbol{u}_{k_j}\cdot\boldsymbol{v}_{k_j}\varphi\,dx = \int_{\mathbb{R}^n}\boldsymbol{u}\cdot\boldsymbol{v}\varphi\,dx\,. \qquad \square
$$

2.1.15 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. 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 $\dfrac{1}{r} = \dfrac{1}{p_1} + \dfrac{1}{p_2}$ 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 $\sum_{j=1}^{m} \dfrac{1}{p_j} = 1$, then

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

Problem 2.3. 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

$$\lim_{h \to 0} \frac{1}{h} \int_{\mathbb{R}} \left[\varphi(x+h) - \varphi(x)\right] g(x) \, dx = - \lim_{h \to 0} \frac{1}{h} \int_{\mathbb{R}} \int_{x}^{x+h} f(y) \varphi(x) \, dy \, dx \, .)$$

Problem 2.4. Show that $W^{n,1}(\mathbb{R}^n) \subseteq C(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$.
(**Hint**$u(x) = \displaystyle\int_{-\infty}^{0} \cdots \int_{-\infty}^{0} \frac{\partial^n}{\partial x_1 \cdots \partial x_n} u(x+y) dy_1 \cdots dy_n$.)

Problem 2.5. 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 $\|\eta_\epsilon * u - u\|_{W^{1,p}(\mathbb{R}^n)} < \delta$. 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 2.6. In case n $= 1$, deduce from Problems 2.3 and 2.5 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 2.7. 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:

$$\|Du\|^2_{L^2(\Omega)} \leqslant C\|u\|_{L^2(\Omega)}\|D^2 u\|_{L^2(\Omega)} \qquad \forall\, u \in H^2(\Omega) \cap H^1_0(\Omega)\,,$$

where $D^2 u$ 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

$$u(x) = \left[-\log|x|\right]^\alpha.$$

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},

$$f_{n_\ell}\, Df_{n_\ell} \rightharpoonup f\, Df \quad \text{weakly in} \;\; L^p(\Omega)\,?$$

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 $u'(x) = \frac{du}{dx}(x)$ 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 $\bar{u} = \int_\Omega u(x)\, dx$, 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* $|u(x)-u(0)| \leqslant \int_0^x |u'(y)|dy$ *and* $|u(x)-u(1)| \leqslant \int_x^1 |u'(y)|dy$.)

(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 $\frac{u(x)}{x} \in L^p(\Omega)$ 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

$$v(x) := \begin{cases} \dfrac{u}{x} & \text{if} \quad x \in (0,1] \\ 0 & \text{if} \quad\quad x = 0\,. \end{cases}$$

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).)

2.2 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 tempeblack distributions.

2.2.1 Fourier Transform on $L^1(\mathbb{R}^n)$ and the Space $\mathscr{S}(\mathbb{R}^n)$

Definition 2.72. 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 2.73. The Schwartz space is the collection of smooth functions of rapid decay denoted by

$$\mathscr{S}(\mathbb{R}^n) = \left\{u \in \mathscr{C}^\infty(\mathbb{R}^n) \,\middle|\, x^\beta D^\alpha u \in L^\infty(\mathbb{R}^n) \quad \forall\, \alpha, \beta \in \mathbb{N}^n\right\}.$$

Elements in $\mathscr{S}(\mathbb{R}^n)$ are called Schwartz functions.

It is not difficult to show (as it follows from the definition) that

$$\mathscr{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)\,,$$

and that

$$\xi^\alpha D_\xi^\beta \widehat{f} = (-i)^{|\alpha|}(-1)^{|\beta|}\mathscr{F}(D_x^\alpha x^\beta f)\,.$$

The Schwartz space $\mathscr{S}(\mathbb{R}^n)$ is also known as the space of rapidly decreasing functions; thus, after multiplying by any polynomial functions $\mathcal{P}(x)$,

$$\mathcal{P}(x) D^\alpha u(x) \to 0 \text{ as } x \to \infty \text{ for all } \alpha \in \mathbb{N}^n.$$

The classical space of test functions $\mathscr{D}(\mathbb{R}^n) := \mathscr{C}_c^\infty(\mathbb{R}^n) \subseteq \mathscr{S}(\mathbb{R}^n)$. The prototype element of $\mathscr{S}(\mathbb{R}^n)$ 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 $\mathscr{S}(\mathbb{R}^n)$:

1. $\mathscr{S}(\mathbb{R}^n)$ is a vector space.
2. $\mathscr{S}(\mathbb{R}^n)$ is an algebra under the pointwise product of functions.
3. $\mathcal{P}u \in \mathscr{S}(\mathbb{R}^n)$ for all $u \in \mathscr{S}(\mathbb{R}^n)$ and all polynomial functions $\mathcal{P}$.
4. $\mathscr{S}(\mathbb{R}^n)$ is closed under differentiation.
5. $\mathscr{S}(\mathbb{R}^n)$ is closed under translations and multiplication by complex exponentials $e^{ix\cdot\xi}$.
6. $\mathscr{S}(\mathbb{R}^n) \subseteq L^1(\mathbb{R}^n)$ (since if $u \in \mathscr{S}(\mathbb{R}^n)$, $|u(x)| \leqslant C(1+|x|)^{-(n+1)}$ for some $C > 0$, and $(1+|x|)^{-(n+1)}dx$ decays like $|x|^{-2}$ as $|x| \to \infty$).

Definition 2.74. For all $f \in L^1(\mathbb{R}^n)$, we define operator $\mathscr{F}^*$ by

$$(\mathscr{F}^* f)(x) = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} f(\xi) e^{ix\cdot\xi} d\xi\,.$$

The function $\mathscr{F}^* f$ sometimes is also denoted by $\check{f}$.

Lemma 2.75. *$(\mathscr{F}u, v)_{L^2(\mathbb{R}^n)} = (u, \mathscr{F}^* v)_{L^2(\mathbb{R}^n)}$ for all $u, v \in \mathscr{S}(\mathbb{R}^n)$.*

Recall that the $L^2(\mathbb{R}^n)$ inner-product for complex-valued functions is given by $(u, v)_{L^2(\mathbb{R}^n)} = \int_{\mathbb{R}^n} u(x)\overline{v(x)}\, dx$.

Proof. Since $u, v \in \mathscr{S}(\mathbb{R}^n)$, by Fubini's Theorem,

$$\begin{aligned}(\mathscr{F}u, v)_{L^2(\mathbb{R}^n)} &= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} u(x) e^{-ix\cdot\xi}\, dx\, \overline{v(\xi)}\, d\xi \\ &= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} u(x) \overline{e^{ix\cdot\xi} v(\xi)}\, d\xi\, dx \\ &= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} u(x) \int_{\mathbb{R}^n} \overline{e^{ix\cdot\xi} v(\xi)}\, d\xi\, dx = (u, \mathscr{F}^* v)_{L^2(\mathbb{R}^n)} \,. \end{aligned}$$ □

Theorem 2.76. *$\mathscr{F}^*\mathscr{F} = \mathrm{Id} = \mathscr{F}\mathscr{F}^*$ 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,

$$(\mathscr{F}^*\mathscr{F} f)(x) = \lim_{\epsilon\to 0} (2\pi)^{-n} \int_{\mathbb{R}^n} f(y) \Big(\int_{\mathbb{R}^n} e^{-\epsilon|\xi|^2} e^{i(y-x)\cdot\xi}\, d\xi \Big) dy \,.$$

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

$$p_\epsilon(x) = \frac{1}{(4\pi\epsilon)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4\epsilon}} \quad \text{and that} \int_{\mathbb{R}^n} p(x)\, dx = 1\,. \tag{2.45}$$

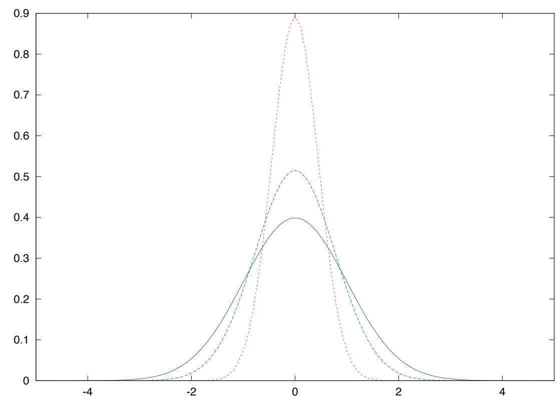


Figure 2.1: As $\epsilon \to 0$, the sequence of functions p_ϵ becomes more localized about the origin.

Given (2.45), then for all $f \in \mathscr{S}(\mathbb{R}^n)$, $p_\epsilon * f \to f$ 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 1.40, 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 (2.45) is true, then even though p_ϵ does not have compact support, $\int_{B(0,\delta)^\complement} p_\epsilon(x)\, dx \to 0$ as $\epsilon \to 0$ for all $\delta > 0$.)

Thus, it remains to prove (2.45). It suffices to consider the case $\epsilon = \frac{1}{2}$; then by definition

$$p_{\frac{1}{2}}(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{ix\cdot\xi} e^{-\frac{|\xi|^2}{2}} d\xi = \mathscr{F}\Big((2\pi)^{-\frac{n}{2}} e^{-\frac{|\cdot|^2}{2}}\Big)(x)\,.$$

In order to prove that $p_{\frac{1}{2}}(x) = (2\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{2}}$, we must show that with the Gaussian function $G(x) = (2\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{2}}$,

$$G(\xi) = \mathscr{F}(G)(\xi)\,.$$

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 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

$$\int_{\mathbb{R}} e^{\frac{x^2}{2}}\,dx = (2\pi)^{\frac{1}{2}}\,. \tag{2.46}$$

To prove (2.46), one can again rely on the multiplication property of the exponential to observe that

$$\int_{\mathbb{R}} e^{x_1^2/2}\,dx_1 \int_{\mathbb{R}} e^{x_2^2/2}\,dx_2 = \int_{\mathbb{R}^2} e^{(x_1^2+x_2^2)/2}\,dx = \int_0^{2\pi}\int_0^{\infty} e^{-\frac{r^2}{2}}rdrd\theta = 2\pi\,. \qquad \square$$

It follows from Lemma 2.75 that for all $u, v \in \mathscr{S}(\mathbb{R}^n)$,

$$(\mathscr{F}u, \mathscr{F}v)_{L^2(\mathbb{R}^n)} = (u, \mathscr{F}^*\mathscr{F}v)_{L^2(\mathbb{R}^n)} = (u, v)_{L^2(\mathbb{R}^n)}\,.$$

Thus, we have established the *Plancheral theorem* on $\mathscr{S}(\mathbb{R}^n)$.

Theorem 2.77 (Plancheral's theorem)**.** *$\mathscr{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is an isomorphism with inverse $\mathscr{F}^*$ preserving the $L^2(\mathbb{R}^n)$ inner-product.*

2.2.2 The Topology on $\mathscr{S}(\mathbb{R}^n)$ and Tempeblack Distributions

An alternative to Definition 2.73 can be stated as follows:

Definition 2.78 (The space $\mathscr{S}(\mathbb{R}^n)$)**.** Setting $\langle x\rangle = \sqrt{1+|x|^2}$,

$$\mathscr{S}(\mathbb{R}^n) = \left\{ u \in \mathscr{C}^\infty(\mathbb{R}^n) \,\middle|\, \langle x\rangle^k |D^\alpha u| \leqslant C_{k,\alpha} \quad \forall\, k \in \mathbb{N} \right\}.$$

The space $\mathscr{S}(\mathbb{R}^n)$ has a Fréchet topology determined by semi-norms.

Definition 2.79 (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 2.80 (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 2.81 (Tempeblack Distributions)**.** A linear map $T : \mathscr{S}(\mathbb{R}^n) \to \mathbb{C}$ is continuous if for each $k \in \mathbb{N}$, 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 tempeblack distributions.

Definition 2.82 (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 2.83. 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 2.84. 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 2.85 (Distributional derivative of Dirac measure)**.**

$$\left\langle \frac{d\delta}{dx}, u \right\rangle = -\frac{du}{dx}(0) \ \ \forall\, u \in \mathscr{S}(\mathbb{R}^n)\,.$$

2.2.3 Fourier Transform on $\mathscr{S}'(\mathbb{R}^n)$

Definition 2.86. 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 2.87. *$\mathscr{F}^*\mathscr{F} = \mathrm{Id} = \mathscr{F}\mathscr{F}^*$ on $\mathscr{S}'(\mathbb{R}^n)$.*

Proof. By Definition 2.86, 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 2.76. $\square$

Example 2.88 (Fourier transform of δ)**.** We claim that $\mathscr{F}\delta = (2\pi)^{-\frac{\mathrm{n}}{2}}$. According to Definition 2.86, for all $u \in \mathscr{S}(\mathbb{R}^{\mathrm{n}})$,

$$\langle \mathscr{F}\delta, u\rangle = \langle \delta, \mathscr{F}u\rangle = \mathscr{F}u(0) = \int_{\mathbb{R}^{\mathrm{n}}} (2\pi)^{-\frac{\mathrm{n}}{2}}\, u(x)\, dx\,,$$

so that $\mathscr{F}\delta = (2\pi)^{-\frac{\mathrm{n}}{2}}$.

Example 2.89. The same argument shows that $\mathscr{F}^*(\delta) = (2\pi)^{-\frac{\mathrm{n}}{2}}$ so that $\mathscr{F}^*[(2\pi)^{\frac{\mathrm{n}}{2}}\delta] = 1$. Using Theorem 2.87, we see that $\mathscr{F}(1) = (2\pi)^{\frac{\mathrm{n}}{2}}\delta$. 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 2.90. A function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^{\mathrm{n}})$ generates a distribution $f \in \mathscr{S}'(\mathbb{R}^{\mathrm{n}})$. We now show that the Fourier transform given by Definition 2.86 agrees with the Fourier transform of a function.

For $\varphi \in \mathscr{S}(\mathbb{R}^{\mathrm{n}})$,

$$\begin{aligned}\langle \widehat{f}, \varphi\rangle \equiv \langle f, \widehat{\varphi}\rangle &= \int_{\mathbb{R}^{\mathrm{n}}} f(\xi)\widehat{\varphi}(\xi)\, d\xi = \frac{1}{(2\pi)^{\frac{\mathrm{n}}{2}}}\int_{\mathbb{R}^{\mathrm{n}}}\int_{\mathbb{R}^{\mathrm{n}}} f(\xi)\varphi(x)e^{-ix\cdot\xi}\, dx\, d\xi \\ &= \lim_{m\to\infty}\frac{1}{(2\pi)^{\frac{\mathrm{n}}{2}}}\int_{\Omega_m}\int_{\mathbb{R}^{\mathrm{n}}} f(\xi)\varphi(x)e^{-ix\cdot\xi}\, dx\, d\xi\,,\end{aligned}$$

where Ω_m is an increasing sequence of bounded sets such that $\bigcup_{m=1}^{\infty}\Omega_m = \mathbb{R}^{\mathrm{n}}$. Letting $f_m = \mathbf{1}_{\Omega_m} f$ or $\widehat{f_m}(x) = \frac{1}{(2\pi)^{\frac{\mathrm{n}}{2}}}\int_{\Omega_m} f(\xi)e^{-ix\cdot\xi}d\xi$, we find that

$$\langle \lim_{m\to\infty}\widehat{f_m}, \varphi\rangle = \frac{1}{(2\pi)^{\frac{\mathrm{n}}{2}}}\int_{\mathbb{R}^{\mathrm{n}}}\lim_{m\to\infty}\int_{\Omega_m} f(\xi)e^{-ix\cdot\xi}\varphi(x)\, d\xi\, dx\,.$$

Therefore, if we define $\widehat{f} = \lim_{m\to\infty}\widehat{f_m}$ whenever the limit makes sense, then we have the following identity

$$\int_{\mathbb{R}^{\mathrm{n}}}\int_{\mathbb{R}^{\mathrm{n}}} f(\xi)\varphi(x)e^{-ix\cdot\xi}d\xi\, dx = \int_{\mathbb{R}^{\mathrm{n}}}\int_{\mathbb{R}^{\mathrm{n}}} f(\xi)\varphi(x)e^{-ix\cdot\xi}\, dx\, d\xi\,, \tag{2.47}$$

and $\widehat{f}$ agrees with the Fourier transform of a function. Note that (2.47) shows that we can interchange the order of integration even though $f(\xi)\varphi(x)e^{-ix\cdot\xi}$ does not belong to $L^1(\mathbb{R}^{2\mathrm{n}})$.

2.2.4 The Fourier Transform on $L^2(\mathbb{R}^n)$

In Theorem 1.40, 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 2.77,

$$\|\widehat{u}_j - \widehat{u}_k\|_{L^2(\mathbb{R}^n)} = \|u_j - u_k\|_{L^2(\mathbb{R}^n)} < \epsilon .$$

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 2.91 (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 2.92. $(u,v)_{L^2(\mathbb{R}^n)} = (\mathscr{F}u, \mathscr{F}v)_{L^2(\mathbb{R}^n)}$ *for all* $u, v \in L^2(\mathbb{R}^n)$.

2.2.5 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 **??**) (by interpolating p between 1 and 2) yields the following result.

[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.

Theorem 2.93 (Hausdorff-Young inequality)**.** *If $u \in L^p(\mathbb{R}^n)$ for $1 \leqslant p \leqslant 2$, then for $q = \frac{p-1}{p}$, there exists a constant C such that*

$$\|\widehat{u}\|_{L^q(\mathbb{R}^n)} \leqslant C\|u\|_{L^p(\mathbb{R}^n)} .$$

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 2.94 (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

$$\left|\widehat{f}(\xi) - \int_{B_M} e^{-ix\cdot\xi} f(x)\, dx\right| < \epsilon .$$

Using Lemma 1.34, 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,

$$\left|\widehat{f}(\xi) - \int_{B_M} \varphi_j(x) e^{-ix\cdot\xi}\, dx\right| < 2\epsilon .$$

Writing $\varphi_j(x) = \sum_{\ell=1}^N C_\ell \mathbf{1}_{E_\ell}(x)$, we have that

$$\left|\widehat{f}(\xi) - \sum_{\ell=1}^N C_\ell \int_{E_\ell} \varphi_j(x) e^{-ix\cdot\xi}\, dx\right| < 2\epsilon .$$

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, $O_\ell = \bigcup_{\alpha \in A_\ell} \mathcal{V}_\alpha^\ell$ for some open rectangle $\mathcal{V}_\alpha^\ell$ and index set A_ℓ. By the compactness of K_ℓ, $K_\ell \subseteq \bigcup_{j=1}^{N_\ell} \mathcal{V}_{\alpha_j}^\ell$ 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

$$\Big| \int_{E_\ell} e^{-ix\cdot\xi}\, dx - \int_{\bigcup_{j=1}^{N_\ell} \mathcal{V}^\ell_{\alpha_j}} e^{-ix\cdot\xi}\, dx \Big| < \epsilon\,.$$

On the other hand, for each rectangle $\mathcal{V}^\ell_{\alpha_j}$, $\Big| \int_{\mathcal{V}^\ell_{\alpha_j}} e^{-ix\cdot\xi}\, dx \Big| \leqslant \dfrac{C}{\xi_1\cdots\xi_n}$, 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. ▫

2.2.6 Convolution and the Fourier Transform

Theorem 2.95. *If $u, v \in L^1(\mathbb{R}^n)$, then $u * v \in L^1(\mathbb{R}^n)$ and*

$$\mathscr{F}(u*v) = (2\pi)^{\frac{n}{2}}(\mathscr{F}u)(\mathscr{F}v)\,, \quad \mathscr{F}^*(u*v) = (2\pi)^{\frac{n}{2}}(\mathscr{F}^*u)(\mathscr{F}^*v)\,.$$

Proof. Young's inequality (Theorem 1.45) shows that $u * v \in L^1(\mathbb{R}^n)$ so that the Fourier transform is well-defined. The assertion then follows from a direct computation:

$$\begin{aligned}
\mathscr{F}(u*v) &= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi}(u*v)(x)\, dx \\
&= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} u(x-y)v(y)dy\, e^{-ix\cdot\xi}\, dx \\
&= (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} u(x-y)e^{-i(x-y)\cdot\xi}\, dx\, v(x)\, e^{-iy\cdot\xi}\, dy \\
&= (2\pi)^{\frac{n}{2}}\widehat{u}\widehat{v} \quad \text{(by Fubini's theorem)}\,.
\end{aligned}$$

That $\mathscr{F}^*(u*v) = (2\pi)^{\frac{n}{2}}(\mathscr{F}^*u)(\mathscr{F}^*v)$ can be proved in a similar way. ▫

By using Young's inequality (Theorem 1.46) together with the Hausdorff-Young inequality, we can generalize the convolution result to the following

Theorem 2.96. *Suppose that $u \in L^p(\mathbb{R}^n)$ and $v \in L^q(\mathbb{R}^n)$, and let r satisfy $\dfrac{1}{r} = \dfrac{1}{p} + \dfrac{1}{q} - 1$ for $1 \leqslant p, q, r \leqslant 2$. Then $\mathscr{F}(u*v), \mathscr{F}^*(u*v) \in L^{\frac{r}{r-1}}(\mathbb{R}^n)$ and*

$$\mathscr{F}(u*v) = (2\pi)^{\frac{n}{2}}(\mathscr{F}u)(\mathscr{F}v)\,, \quad \mathscr{F}^*(u*v) = (2\pi)^{\frac{n}{2}}(\mathscr{F}^*u)(\mathscr{F}^*v)\,.$$

Let $\star$ denote the convolution operator defined by $f \star g = (2\pi)^{-\frac{n}{2}}(f * g)$. Then the theorem above implies that if $u \in L^p(\mathbb{R}^n)$ and $v \in L^q(\mathbb{R}^n)$,

$$\mathscr{F}(u \star v) = (\mathscr{F}u)(\mathscr{F}v)\,, \quad \mathscr{F}^*(u \star v) = (\mathscr{F}^*u)(\mathscr{F}^*v)\,.$$

Using this notation, we find that for $f, g \in \mathscr{S}(\mathbb{R}^n)$,

$$\langle f \star g, \varphi\rangle = \Big(\frac{1}{\sqrt{2\pi}}\Big)^n \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(x-y)g(y)\varphi(x)\,dy\,dx = \langle g, \widetilde{f} \star \varphi\rangle\,,$$

where $\widetilde{\ }: \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is the reflection operator defined by $\widetilde{f}(x) = f(-x)$. This observation motivates the following

Definition 2.97. Let $f \in \mathscr{S}(\mathbb{R}^n)$ and $T \in \mathscr{S}'(\mathbb{R}^n)$. The convolution $f \star T$ is the tempeblack distribution defined by

$$\langle f \star T, \varphi\rangle = \langle T, \widetilde{f} \star \varphi\rangle \qquad \forall\, \varphi \in \mathscr{S}(\mathbb{R}^n)\,,$$

Theorem 2.98. *Let $f \in \mathscr{S}(\mathbb{R}^n)$ and $T \in \mathscr{S}'(\mathbb{R}^n)$. 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,

$$\langle \mathscr{F}(f \star T), \varphi\rangle = \langle f \star T, \widehat{\varphi}\rangle = \langle T, \widetilde{f} \star \widehat{\varphi}\rangle = \langle \widehat{T}, \check{\widetilde{f}} \cdot \varphi\rangle = \langle \widehat{T}, \widehat{f} \cdot \varphi\rangle = \langle \widehat{f}\widehat{T}, \varphi\rangle\,. \qquad \square$$

2.2.7 An Explicit Computation with the Fourier Transform

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) $f(x) = |x|^\alpha$, $-n < \alpha < 0$.

Case (1) In this case, f_1 is rapidly decreasing but not in the Schwartz class $\mathcal{S}(\mathbb{R}^n)$. We begin with $n = 1$. 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 $e^{-t|x|} = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{t}{t^2+\xi^2}\, e^{ix\xi}d\xi$. Next, when $n > 1$ we begin with the observation

$$\int_0^\infty e^{-st^2}e^{-s\xi^2}ds = \frac{e^{-s(t^2+\xi^2)}}{-(t^2+\xi^2)}\Big|_0^\infty = \frac{1}{t^2+\xi^2}\,;$$

thus

$$e^{-t|x|} = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{t}{t^2+\xi^2}\; e^{i|x|\xi} d\xi = \frac{1}{\pi}\int_{-\infty}^{\infty} t\Big(\int_0^\infty e^{-st^2} e^{-s\xi^2} ds\Big) e^{i\lambda\xi} d\xi$$
$$= \frac{1}{\pi}\int_0^\infty t\Big(\int_{-\infty}^{\infty} e^{-s\xi^2} e^{i|x|\xi} d\xi\Big) e^{-st^2} ds = \int_0^\infty \frac{t}{\sqrt{s\pi}} e^{-st^2} e^{-\frac{|x|^2}{4s}} ds\,.$$

As a consequence,

$$\mathscr{F}(e^{-t|\cdot|})(\xi) = \int_0^\infty g(t,s)\mathcal{F}(e^{-\frac{|x|^2}{4s}})ds = \int_0^\infty \frac{t}{\sqrt{s\pi}}(2s)^{\frac{\mathrm{n}}{2}} e^{-s(t^2+|\xi|^2)} ds$$
$$= \frac{t}{(t^2+|\xi|^2)^{\frac{\mathrm{n}+1}{2}}}\int_0^\infty \frac{1}{\sqrt{\pi s}}(2s)^{\frac{\mathrm{n}}{2}} e^{-s} ds = \frac{C(\mathrm{n})t}{(t^2+|\xi|^2)^{\frac{\mathrm{n}+1}{2}}}\,,$$

where the constant $C(\mathrm{n}) = \displaystyle\int_0^\infty \frac{1}{\sqrt{\pi s}}(2s)^{\frac{\mathrm{n}}{2}} ds = \sqrt{\frac{2^{\mathrm{n}}}{\pi}}\Gamma\big(\frac{\mathrm{n}+1}{2}\big)$, and Γ is the so-called *Gamma*-function. It follows that

$$\mathscr{F}^{-1}(e^{-t|\cdot|})(x) = \mathcal{F}(e^{-t|\cdot|})(-x) = \sqrt{\frac{2^{\mathrm{n}}}{\pi}}\Gamma(\frac{\mathrm{n}+1}{2})\frac{t}{(t^2+|x|^2)^{\frac{\mathrm{n}+1}{2}}}\,. \tag{2.48}$$

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,

$$\mathscr{F}(|\cdot|^\alpha)(\xi) = \frac{1}{\Gamma(-\frac{\alpha}{2})}\int_0^\infty s^{-\frac{\alpha}{2}-1}\mathscr{F}(e^{-s|\cdot|^2})ds = \frac{1}{2^{\frac{\mathrm{n}}{2}}\Gamma(-\frac{\alpha}{2})}\int_0^\infty s^{-\frac{\alpha}{2}-\frac{\mathrm{n}}{2}-1} e^{-\frac{|\xi|^2}{4s}} ds$$
$$= \frac{1}{2^{\frac{\mathrm{n}}{2}}\Gamma(-\frac{\alpha}{2})}\int_0^\infty \Big(\frac{|\xi|^2}{4s}\Big)^{-\frac{\alpha}{2}-\frac{\mathrm{n}}{2}-1} e^{-s}\frac{|\xi|^2}{4s^2} ds = \frac{2^{\alpha+\frac{\mathrm{n}}{2}}\Gamma(\frac{\alpha+\mathrm{n}}{2})}{\Gamma(-\frac{\alpha}{2})}|\xi|^{-\alpha-\mathrm{n}}\,,$$

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

$$\mathscr{F}^{-1}(|\cdot|^{-2})(x) = \sqrt{\frac{\pi}{2}}\frac{1}{|x|}\,. \tag{2.49}$$

2.2.8 Applications to the Poisson, Heat, and Wave Equations

The Poisson equation on $\mathbb{R}^3$

In Theorem 2.61, 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

$$-|\xi|^2 \widehat{u}(\xi) = \widehat{f}(\xi). \tag{2.50}$$

Distributional solutions to (2.50) 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 (2.49) we see that $\widehat{u}(\xi) = \frac{\widehat{f}(\xi)}{|\xi|^2}$; 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 $\Phi(x) = \frac{1}{4\pi|x|}$. 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

$$\begin{aligned} \langle \Delta(\Phi * f), \widehat{\varphi} \rangle &= (2\pi)^{\frac{\mathrm{n}}{2}} \langle -|\xi|^2 \widehat{\Phi} \widehat{f}, \varphi \rangle = (2\pi)^{\frac{\mathrm{n}}{2}} \langle \mathscr{F}(\Delta\Phi), \widehat{f}\varphi \rangle \\ &= (2\pi)^{\frac{\mathrm{n}}{2}} \langle \Delta\Phi, \mathscr{F}(\widehat{f}\varphi) \rangle = \langle \Delta\Phi, \widetilde{f} * \widehat{\varphi} \rangle = \langle \Delta\Phi * f, \widehat{\varphi} \rangle. \end{aligned}$$

EXAMPLE 2.99. On $\mathbb{R}^2$, $\Delta(e^{x_1}\cos x_2) = 0$. The function $e^{x_1}\cos x_2$ is not a tempeblack 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 ∞.

The Poisson integral formula on the half-space

Let $\Omega = \mathbb{R}^{\mathrm{n}} \times \mathbb{R}_+$, and consider the Dirichlet problem

$$\begin{aligned} \Big(\frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_{\mathrm{n}}^2}\Big)u = \Big(\frac{\partial^2}{\partial t^2} + \Delta\Big)u &= 0 \quad \text{ in } \ \Omega \times (0,\infty)\,, \\ u(\cdot,0) &= f(\cdot) \qquad \text{ on } \ \Omega \times \{t=0\} \end{aligned}$$

for some $f \in \mathscr{S}(\mathbb{R}^{\mathrm{n}})$. 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 $\mathscr{S}'(\mathbb{R}^{\mathrm{n}})$ and hence we may use the Fourier transform. Applying the Fourier transform (in the x variable) $\mathscr{F}_x$, we see that

$$\begin{aligned} \frac{\partial^2}{\partial t^2}(\mathscr{F}_x u)(\xi,t) &= |\xi|^2(\mathscr{F}_x u)(\xi,t) \qquad \forall\,(\xi,t) \in \mathbb{R}^{\mathrm{n}} \times \mathbb{R}_+\,, \\ (\mathscr{F}_x u)(\xi,0) &= \widehat{f}(\xi) \qquad\qquad\qquad \forall\,\xi \in \mathbb{R}^{\mathrm{n}}\,. \end{aligned}$$

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 (2.48),

$$\begin{aligned} u(x,t) &= \mathscr{F}_x^*(\widehat{f}(\cdot)e^{-t|\cdot|})(x) = \big[\mathscr{F}_x^*(e^{-t|\cdot|}) * f\big](x) \\ &= \frac{\Gamma(\frac{\mathrm{n}+1}{2})}{\pi^{\frac{\mathrm{n}+1}{2}}} \int_{\mathbb{R}^{\mathrm{n}}} \frac{tf(y)}{(t^2+|x-y|^2)^{\frac{\mathrm{n}+1}{2}}}\,dy\,. \end{aligned}$$

This is the *Poisson integral formula on the half-space.*

If f is bounded; that is, $f \in L^\infty(\mathbb{R}^{\mathrm{n}})$, then the integral converges and $u \in L^\infty(\mathbb{R}^{\mathrm{n}} \times \mathbb{R}_+)$. Therefore, $u \in \mathscr{C}^\infty(\mathbb{R}^{\mathrm{n}} \times \mathbb{R}_+) \cap L^\infty(\mathbb{R}^{\mathrm{n}} \times \mathbb{R}_+)$.

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*

$$u_t(x,t) = \Delta u(x,t) \qquad \forall\, (x,t) \in \mathbb{R}^{\mathrm{n}} \times (0,\infty)\,, \tag{2.51a}$$
$$u(x,0) = g(x) \qquad \forall\, x \in \mathbb{R}^{\mathrm{n}}. \tag{2.51b}$$

Taking the Fourier transform of (2.51), we find that

$$\partial_t \widehat{u}(\xi,t) = -|\xi|^2 \widehat{u}(\xi,t)\,,$$
$$\widehat{u}(\xi,0) = \widehat{g}(\xi)\,.$$

Therefore, $\widehat{u}(\xi,t) = \widehat{g}(\xi) e^{-|\xi|^2 t}$ and hence

$$\begin{aligned} u(x,t) &= \mathscr{F}^*\big(\widehat{g}(\cdot)e^{-|\cdot|^2 t}\big)(x) = \big[\mathscr{F}^*(e^{-|\cdot|^2 t}) \star g\big](x) \\ &= \frac{1}{(4\pi t)^{\mathrm{n}/2}} \int_{\mathbb{R}^{\mathrm{n}}} e^{-\frac{|x-y|^2}{4t}} g(y)\, dy \quad \big(\equiv (\mathscr{H}(\cdot,t) * g)(x)\big)\,, \end{aligned} \tag{2.52}$$

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 2.100. *If $g \in L^\infty(\mathbb{R}^{\mathrm{n}})$, then the solution u to (2.51) 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 2.101. The representation formula (2.52) shows that whenever g is bounded, continuous, and positive, the solution $u(x,t)$ to (2.51) is positive everywhere for $t > 0$.

The representation formula (2.52) can also be used to prove the following

Theorem 2.102. *Assume that $g \in \mathscr{C}(\mathbb{R}^{\mathrm{n}}) \cap L^\infty(\mathbb{R}^{\mathrm{n}})$. Then u defined by (2.52) 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*

$$u_t(x,t) - \Delta u(x,t) = f(x,t) \qquad \forall\, (x,t) \in \mathbb{R}^{\mathrm{n}} \times (0,\infty)\,, \tag{2.53a}$$
$$u(x,0) = 0 \qquad \forall\, x \in \mathbb{R}^{\mathrm{n}}\,. \tag{2.53b}$$

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 (2.52),

$$U(x,t,s) = \int_{\mathbb{R}^n} \mathscr{H}(x-y,t-s) f(y,s) dy\,.$$

We next invoke *Duhamel's principle* to find a solution $u(x,t)$ to (2.53):

$$u(x,t) = \int_0^t U(x,t,s)ds = \int_0^t \int_{\mathbb{R}^n} \mathscr{H}(x-y,t-s) f(y,s) dyds\,. \tag{2.54}$$

The principle of linear superposition then shows that the solution of the equations

$$\begin{aligned} u_t(x,t) - \Delta u(x,t) &= f(x,t) && \forall\,(x,t) \in \mathbb{R}^n \times (0,\infty)\,, \\ u(x,0) &= g(x) && \forall\, x \in \mathbb{R}^n\,, \end{aligned}$$

is the sum of (2.52) and (2.54):

$$\begin{aligned} u(x,t) &= \int_0^t \int_{\mathbb{R}^n} \mathscr{H}(x-y,t-s) f(y,s) dyds + \int_{\mathbb{R}^n} \mathscr{H}(x-y,t) g(y) dy \\ &= \big[\mathscr{H}(\cdot,t) * g\big](x) + \int_0^t \big[\mathscr{H}(\cdot,t-s) * f(\cdot,s)\big](x)\, ds\,. \end{aligned} \tag{2.55}$$

The Wave equation

For wave speed $c > 0$, and for $x \in \mathbb{R}^n$, $t \in \mathbb{R}$, consider the following second-order linear hyperbolic equation:

$$\begin{aligned} u_{tt}(x,t) &= c^2 \Delta u(x,t) && \forall\,(x,t) \in \mathbb{R}^n \times (0,\infty)\,, \\ u(x,0) &= f(x) && \forall\, x \in \mathbb{R}^n, \\ u_t(x,0) &= g(x) && \forall\, x \in \mathbb{R}^n. \end{aligned}$$

Taking the Fourier transform of (2.56), we find that

$$\begin{aligned} \widehat{u}_{tt}(\xi,t) &= -c^2|\xi|^2 \widehat{u}(\xi,t) && \forall\,(\xi,t) \in \mathbb{R}^n \times (0,\infty)\,, \\ \widehat{u}(\xi,0) &= \widehat{f}(\xi) && \forall\, \xi \in \mathbb{R}^n, \\ \widehat{u}_t(\xi,0) &= \widehat{g}(\xi) && \forall\, \xi \in \mathbb{R}^n. \end{aligned}$$

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,

$$\begin{aligned} u(x,t) &= \Big[\mathscr{F}^*(\cos c|\cdot|t) \star f + \mathscr{F}^*\Big(\frac{\sin c|\cdot|t}{c|\cdot|}\Big) \star g\Big](x) \\ &= \frac{1}{c}\Big[\frac{d}{dt}\mathscr{F}^*\Big(\frac{\sin c|\cdot|t}{|\cdot|}\Big) \star f + \mathscr{F}^*\Big(\frac{\sin c|\cdot|t}{|\cdot|}\Big) \star g\Big](x)\,. \end{aligned}$$

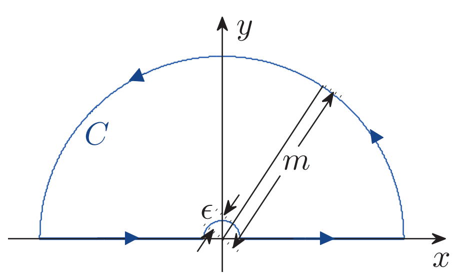
The 1-dimensional case. For the case that n = 1,

$$\int_{-m}^{m} \frac{\sin ct\lambda}{\lambda}\, e^{ix\lambda} d\lambda = \int_{-m}^{m} \frac{e^{i(x+ct)\lambda} - e^{i(x-ct)\lambda}}{2i\lambda}\, d\lambda\,.$$

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 $\lim_{m\to\infty}\int_0^{\pi} e^{-m\sin\theta} d\theta = 0$ (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 2.103. $\mathscr{F}^*\Big(\dfrac{\sin c|\cdot|t}{|\cdot|}\Big)(x) = \sqrt{\dfrac{\pi}{2}}\, \chi_{\{|x|<ct\}}(x)$ *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

$$\Big|\int_{-m}^{m} \frac{\sin c|\xi|t}{|\xi|} e^{ix\xi} d\xi\Big| = \Big|\int_{-m}^{m} \frac{e^{i(x+ct)\xi} - e^{i(x-ct)\xi}}{2i\xi} d\xi\Big| \leqslant 2\pi \quad \forall\, m > 0\,. \tag{2.57}$$

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 (2.57) 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 an 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 the spherical coordinate, 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 $t>0$,

$$\begin{aligned}&\Big[\mathscr{F}^{-1}\Big(\frac{\sin c|\cdot|t}{|\cdot|}\Big)\star\varphi\Big](x)\\ &= \frac{1}{4\pi}\int_{\partial B(0,1)}\int_0^\infty\int_{-\infty}^{\infty}\frac{1}{r}\big(e^{i\rho(r-ct)}-e^{i\rho(|x|+ct)}\big)\varphi(x-r\omega)r^2d\mu_1(\rho)\,d\mu_1(r)\,dS_\omega\\ &= \frac{1}{4\pi}\int_{\partial B(0,1)}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\big(e^{i\rho(r-ct)}-e^{i\rho(|x|+ct)}\big)\varphi(x-r\omega)r\chi_{\{r>0\}}(r)d\mu_1(\rho)\,d\mu_1(r)\,dS_\omega\\ &= \frac{ct}{4\pi}\int_{\partial B(0,1)}\varphi(x-ctw)\,dS_\omega = \frac{1}{4\pi ct}\int_{\partial B(x,ct)}\varphi(y)\,dS_y\,.\end{aligned}$$

Therefore,

$$u(x,t) = \frac{1}{4\pi c^2 t}\int_{\partial B(x,ct)} g(y)\, dS_y + \frac{\partial}{\partial t}\Big[\frac{1}{4\pi c^2 t}\int_{\partial B(x,ct)} f(y)\, dS_y\Big]. \tag{2.58}$$

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.

2.2.9 Exercises

Problem 2.1. (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 2.2. 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 2.3. 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 $\psi(0) = 1$. Use the fact that $w_j \rightharpoonup w$ in $\mathscr{S}'(\mathbb{R}^{\mathrm{n}})$.

Problem 2.4. 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 2.5. Let $T \in \mathscr{S}'(\mathbb{R}^{\mathrm{n}})$ and $f \in S(\mathbb{R}^{\mathrm{n}})$. Show that the Leibniz rule for distributional derivatives holds; that is, show that $\frac{\partial}{\partial x_i}(fT) = f\frac{\partial T}{\partial x_i} + \frac{\partial f}{\partial x_i}T$ in the sense of distribution.

Problem 2.6. 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 2.7. Let d_r denote the map given by $d_r f(x) = f(rx)$. Show that

$$\mathscr{F}(d_r f) = r^{-\mathrm{n}} d_{1/r} \mathscr{F}(f)\,.$$

Problem 2.8. Show that a function $f \in L^2(\mathbb{R}^{\mathrm{n}})$ is real if and only if $\widehat{f}(-\xi) = \overline{\widehat{f}(\xi)}$.

Problem 2.9. Find the Fourier transform of the function $f(x) = xe^{tx^2}$ for $t < 0$.

Problem 2.10. Find the Fourier transform of $\mathbf{1}_{(-a,a)}$, the characteristic (indicator) function of the set $(-a, a)$.

Problem 2.11. 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 2.12. 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 $-\mathrm{n} < \alpha < 0$,

$$\mathscr{F}(|x|^\alpha)(\xi) = \frac{\Gamma(\frac{\mathrm{n}+\alpha}{2})}{\Gamma(-\frac{\alpha}{2})} 2^{\alpha+\frac{\mathrm{n}}{2}} |\xi|^{-(\alpha+\mathrm{n})}$$

and $f(x) = \dfrac{1}{\alpha+2}\dfrac{\partial}{\partial x_1}|x|^{\alpha+2}$.

PROBLEM 2.13. 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 2.14. 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 $\varphi \in L^1_{\text{loc}}(\mathbb{R})$.

PROBLEM 2.15. 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 2.16. Using Definition 2.86, compute the Fourier transform of the function/distribution

$$\mathrm{R}(x) = \begin{cases} x & \text{if } x \geqslant 0\,, \\ 0 & \text{otherwise}\,, \end{cases}$$

by completing the following:

(1) Let H be the Heaviside function. Show that $\widehat{H}(\xi) = \text{p.v.}\dfrac{1}{\sqrt{2\pi} i\xi} + C\delta(\xi)$ for some constant C, where $\text{p.v.}\dfrac{1}{\xi}$ is defined as

$$\Big\langle \text{p.v.}\frac{1}{\xi}, \varphi \Big\rangle = \lim_{\epsilon \to 0^+} \int_{\mathbb{R}\setminus[-\epsilon,\epsilon]} \frac{\varphi(\xi)}{\xi} d\xi = \lim_{\epsilon \to 0^+} \Big(\int_{-\infty}^{-\epsilon} + \int_{\epsilon}^{\infty} \Big) \frac{\varphi(\xi)}{\xi} d\xi\,.$$

Note that the integral above always exists as long as $\varphi \in \mathscr{S}(\mathbb{R})$.

(2) Let $S(x) = H(x) - \dfrac{1}{2}$. 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 2.17. By (2.58), the solution of the 3-dimensional wave equation

$$u_{tt}(x,t) = c^2 \Delta u(x,t) \qquad \text{in} \quad \mathbb{R}^3 \times (0,\infty)\,, \tag{2.59a}$$

$$u(x,0) = f(x) \qquad \text{on} \quad \mathbb{R}^3 \times \{t=0\}\,, \tag{2.59b}$$

$$u_t(x,0) = g(x) \qquad \text{on} \quad \mathbb{R}^3 \times \{t=0\}\,, \tag{2.59c}$$

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 (2.58) as

$$u(x,t) = \frac{1}{4\pi^2 c^2 t^2} \int_{|y-x|=ct} \Big[t g(y) + f(y) + \sum_{i=1}^{3} f_{y_i}(y)(y_i - x_i) \Big] dS_y$$

and convert the integral into an integral over $B(x,ct)$.

Problem 2.18. Let us consider the BBM equation

$$u_t + u_x + u u_x - u_{xxt} = 0 \qquad \forall\, x \in \mathbb{R}\,, t \in (0,T]\,, \tag{2.60a}$$

$$u(x,0) = g(x) \qquad \forall\, x \in \mathbb{R}\,. \tag{2.60b}$$

1. Use the Fourier transform to show that a bounded solution to (2.60) satisfies

$$u(x,t) = g(x) + \int_0^t \int_{-\infty}^{\infty} K(x-y)\Big[u(y,s) + \frac{1}{2}u^2(y,s)\Big]dyds\,, \tag{2.61}$$

where K is defined by

$$K(x) = \frac{1}{2}\,\mathrm{sgn}(x)e^{-|x|}\,.$$

2. Write (2.61) as $u = F(u)$; that is, treat the right-hand side of (2.61) 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.)

PROBLEM 2.19. 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 $f(nT)$.

2.3 The Sobolev Spaces $H^s(\mathbb{R}^n)$, $s \in \mathbb{R}$

2.3.1 $H^s(\mathbb{R}^n)$ 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 2.104. 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 2.105. 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 2.106. *For $s \in \mathbb{R}$, $(H^s(\mathbb{R}^n), \| \cdot \|_{H^s(\mathbb{R}^n)})$ is a Hilbert space.*

Example 2.107 ($H^1(\mathbb{R}^n)$)**.** The $H^1(\mathbb{R}^n)$ in Fourier representation is exactly the same as the that given by Definition 2.15:

$$\begin{aligned}
\|u\|^2_{H^1(\mathbb{R}^n)} &= \int_{\mathbb{R}^n} \langle \xi \rangle^2 |\hat{u}(\xi)|^2 d\xi = \int_{\mathbb{R}^n} (1 + |\xi|^2) |\hat{u}(\xi)|^2 d\xi \\
&= \int_{\mathbb{R}^n} (|u(x)|^2 + |Du(x)|^2)\, dx\,,
\end{aligned}$$

the last equality following from the Plancheral theorem.

Example 2.108 ($H^{\frac{1}{2}}(\mathbb{R}^{\mathrm{n}})$)**.** The $H^{\frac{1}{2}}(\mathbb{R}^{\mathrm{n}})$ can be viewed as interpolating between decay requiblack for $\widehat{u} \in L^2(\mathbb{R}^{\mathrm{n}})$ and $\widehat{u} \in H^1(\mathbb{R}^{\mathrm{n}})$:

$$H^{\frac{1}{2}}(\mathbb{R}^{\mathrm{n}}) = \left\{ u \in L^2(\mathbb{R}^{\mathrm{n}}) \,\Big|\, \int_{\mathbb{R}^{\mathrm{n}}} \sqrt{1+|\xi|^2}|\widehat{u}(\xi)|^2 \, d\xi < \infty \right\}.$$

Example 2.109 ($H^{-1}(\mathbb{R}^{\mathrm{n}})$)**.** The space $H^{-1}(\mathbb{R}^{\mathrm{n}})$ can be heuristically described as those distributions whose anti-derivative is in $L^2(\mathbb{R}^{\mathrm{n}})$; in terms of the Fourier representation, elements of $H^{-1}(\mathbb{R}^{\mathrm{n}})$ possess a transforms that can grow linearly at infinity:

$$H^{-1}(\mathbb{R}^{\mathrm{n}}) = \left\{ u \in \mathscr{S}'(\mathbb{R}^{\mathrm{n}}) \,\Big|\, \int_{\mathbb{R}^{\mathrm{n}}} \frac{|\widehat{u}(\xi)|^2}{1+|\xi|^2} \, d\xi < \infty \right\}.$$

For $T \in H^{-s}(\mathbb{R}^{\mathrm{n}})$ and $u \in H^s(\mathbb{R}^{\mathrm{n}})$, the duality pairing is given by

$$\langle T, u \rangle = (\Lambda^{-s}T, \Lambda^s u)_{L^2(\mathbb{R}^{\mathrm{n}})} ,$$

from which the following result follows.

Proposition 2.110. *For all $s \in \mathbb{R}$, $H^s(\mathbb{R}^{\mathrm{n}})' = H^{-s}(\mathbb{R}^{\mathrm{n}})$.*

The ability to define fractional-order Sobolev spaces $H^s(\mathbb{R}^{\mathrm{n}})$ allows us to refine the estimates of the trace of a function which we previously stated in Theorem 2.42. That result, based on the Gauss-Green theorem, stated that the trace operator was continuous from $H^1(\mathbb{R}^{\mathrm{n}}_+)$ into $L^2(\mathbb{R}^{\mathrm{n}-1})$. In fact, the trace operator is continuous from $H^1(\mathbb{R}^{\mathrm{n}}_+)$ into $H^{\frac{1}{2}}(\mathbb{R}^{\mathrm{n}-1})$.

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 2.111. *$\tau : H^1(\mathbb{R}^2) \to H^{\frac{1}{2}}(\mathbb{R})$, 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 2.112. *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 2.111. Suppose that $u \in \mathscr{S}(\mathbb{R}^2)$ and set $f(x_2) = u(0, x_1)$. According to Lemma 2.112,

$$\begin{aligned} \widehat{f}(\xi_2) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{u}(\xi_1, \xi_2) d\xi_1 = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{u}(\xi_1, \xi_2) \langle\xi\rangle \langle\xi\rangle^{-1} d\xi_1 \\ &\leqslant \frac{1}{\sqrt{2\pi}} \Big(\int_{\mathbb{R}} |\widehat{u}(\xi_1, \xi_2)|^2 \langle\xi\rangle^2 d\xi_1 \Big)^{\frac{1}{2}} \Big(\int_{\mathbb{R}} \langle\xi\rangle^{-2} d\xi_1 \Big)^{\frac{1}{2}} \,, \end{aligned}$$

and hence

$$|f(\xi_2)|^2 \leqslant C \int_{\mathbb{R}} |\widehat{u}(\xi_1, \xi_2)|^2 \langle\xi\rangle^2 d\xi_1 \int_{\mathbb{R}} \langle\xi\rangle^{-2} d\xi_1 \,.$$

The key to this trace estimate is the explicit evaluation of the integral $\displaystyle\int_{\mathbb{R}} \langle\xi\rangle^{-2} d\xi_1$:

$$\int_{\mathbb{R}} \frac{1}{1 + \xi_1^2 + \xi_2^2} d\xi_1 = \frac{\tan^{-1}\big(\frac{\xi_1}{\sqrt{1+\xi_2^2}}\big)}{\sqrt{1+\xi_2^2}} \Bigg|_{\xi_1=-\infty}^{\xi_1=+\infty} \leqslant \pi(1 + \xi_2^2)^{-\frac{1}{2}} \,. \tag{2.62}$$

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 $H^s(\mathbb{R}^n)$ spaces replacing $H^1(\mathbb{R}^n)$, where $n \geqslant 3$ and $s > \dfrac{1}{2}$, proceeds in a very similar fashion; the only difference is that the integral $\displaystyle\int_{\mathbb{R}} \langle\xi\rangle^{-2} d\xi_1$ is replaced by $\displaystyle\int_{\mathbb{R}^{n-1}} \langle\xi\rangle^{-2s} d\xi_1 \cdots d\xi_{n-1}$, and instead of obtaining an explicit anti-derivative of this integral, an upper bound is instead found. The result is the following general trace theorem.

Theorem 2.113 (The trace theorem for $H^s(\mathbb{R}^{\mathrm{n}})$). *For $s > \frac{1}{2}$, the trace operator $\tau : H^s(\mathbb{R}^{\mathrm{n}}) \to H^{s-\frac{1}{2}}(\mathbb{R}^{\mathrm{n}-1})$ is continuous.*

We can extend this result to open, bounded, $\mathscr{C}^\infty$-domains $\Omega \subseteq \mathbb{R}^{\mathrm{n}}$.

Definition 2.114. 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

$$\vartheta_\ell : B(0, r_\ell) \subseteq \mathbb{R}^{\mathrm{n}-1} \to \omega_\ell \text{ is a } \mathscr{C}^\infty\text{-diffeomorphism}\,.$$

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 $u \in H^s(\partial\Omega)$,

$$\|u\|^2_{H^s(\partial\Omega)} = \sum_{\ell=1}^K \|(\varphi_\ell u) \circ \vartheta_\ell\|^2_{H^s(\mathbb{R}^{\mathrm{n}-1})}\,.$$

The space $(H^s(\partial\Omega), \|\cdot\|_{H^s(\partial\Omega)})$ is a Hilbert space by virtue of the completeness of $H^s(\mathbb{R}^{\mathrm{n}-1})$; furthermore, any system of charts for $\partial\Omega$ with subordinate partition of unity will produce an equivalent norm.

Theorem 2.115 (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 (2.21) but with each chart being a $\mathscr{C}^\infty$-map, such that ϑ_ℓ is equal to the restriction of ϑ_ℓ to the $(\mathrm{n}-1)$-dimensional ball $B(0, r_\ell) \subseteq \mathbb{R}^{\mathrm{n}-1}$. 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 2.113, for $s > \frac{1}{2}$,

$$\|u\|^2_{H^{s-\frac{1}{2}}(\partial\Omega)} = \sum_{\ell=1}^K \|(\varphi_\ell u) \circ \vartheta_\ell\|^2_{H^{s-\frac{1}{2}}(\mathbb{R}^{\mathrm{n}-1})} \leqslant C \sum_{\ell=1}^K \|(\varphi_\ell u) \circ \vartheta_\ell\|^2_{H^s(\mathbb{R}^{\mathrm{n}})} \leqslant C\|u\|^2_{H^s(\Omega)}\,. \quad \square$$

One may then ask if the trace operator τ is onto; namely, given $f \in H^{s-\frac{1}{2}}(\mathbb{R}^{\rm n-1})$ for $s > \frac{1}{2}$, does there exist a $u \in H^s(\mathbb{R}^{\rm n})$ such that $f = \tau u$? By essentially reversing the order of the proof of Theorem 2.111, it is possible to answer this question in the affirmative. We first consider the case that $\rm n = 2$ and $s = 1$.

Theorem 2.116. *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 (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 (2.62) for the inequality above.

It remains to prove that $u(x_1, 0) = f(x_1)$, but by Lemma 2.112, 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. □

A similar construction yields the general result.

Theorem 2.117. *For $s > \frac{1}{2}$, the trace operator $\tau : H^s(\mathbb{R}^{\rm n}) \to H^{s-\frac{1}{2}}(\mathbb{R}^{\rm n-1})$ is a surjection.*

By using the system of charts employed for the proof of Theorem 2.115, we also have the surjectivity of the trace map on bounded domains.

Theorem 2.118. *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 2.119. *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 2.77, $u = \mathscr{F}^*\widehat{u}$; thus according to Hölder's inequality and the Riemann-Lebesgue lemma (Theorem 2.94), it suffices to show that

$$\|\widehat{u}\|_{L^1(\mathbb{R}^n)} \leqslant C\|u\|_{H^s(\mathbb{R}^n)}\,. \tag{2.63}$$

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 2.120 (Interpolation inequality)**.** *Let $0 < r < t < \infty$, and $s = \alpha r + (1-\alpha)t$ for some $\alpha \in (0,1)$. Then*

$$\|u\|_{H^s(\mathbb{R}^n)} \leqslant C\|u\|^{\alpha}_{H^r(\mathbb{R}^n)}\|u\|^{1-\alpha}_{H^t(\mathbb{R}^n)}\,. \tag{2.64}$$

Example 2.121 (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

$$\frac{1}{2}\frac{d}{dt}\int_{\mathbb{T}^2}|u(x,t)|^2\,dx+\underbrace{\int_{\mathbb{T}^2}u^i,_j\,u^ju^i\,dx}_{\mathcal{I}_1}+\underbrace{\int_{\mathbb{T}^2}p,_i\,u^idx}_{\mathcal{I}_2}=0\,.$$

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 $\frac{d}{dt}\|u(\cdot,t)\|^2_{L^2(\mathbb{T}^2)}$ 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

$$\frac{d}{dt}\|u(\cdot,t)\|^2_{H^3(\mathbb{T}^2)}\leqslant C\|u(\cdot,t)\|^3_{H^3(\mathbb{T}^2)} \tag{2.65}$$

from which it follows that $\|u(\cdot,t)\|^2_{H^3(\mathbb{T}^2)}\leqslant M$ for some constant $M<\infty$.

To prove (2.65), 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 $D^\alpha\mathrm{div}u=0$; 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

$$\int_{\mathbb{T}^2}D^3(u^i,_j\,u^j)D^3u^i\,dx=\underbrace{\int_{\mathbb{T}^2}D^3u^i,_j\;u^j\,D^3u^i\,dx}_{\mathcal{K}_1}$$
$$+\underbrace{\int_{\mathbb{T}^2}u^i,_j\;D^3u^j\,D^3u^i\,dx}_{\mathcal{K}_2}+\int_{\mathbb{T}^2}\text{l. o. t. }dx\,.$$

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 (2.65) with the $H^2(\mathbb{T}^2)$ norm replacing the $H^3(\mathbb{T}^2)$ norm.

2.3.2 Fractional-Order Sobolev Spaces via Difference Quotient Norms

The case that $s > 0$

Lemma 2.122. *For* $0 < s < 1$, $u \in H^s(\mathbb{R}^n)$ *is equivalent to*

$$u \in L^2(\mathbb{R}^n) , \qquad \iint_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|u(x) - u(y)|^2}{|x-y|^{n+2s}}\, dxdy < \infty .$$

Proof. The Fourier transform shows that for $h \in \mathbb{R}^n$,

$$\int_{\mathbb{R}^n} |u(x+h) - u(x)|^2\, dx = \int_{\mathbb{R}^n} |e^{ih\cdot\xi} - 1|^2 |\widehat{u}(\xi)|^2 d\xi = \int_{\mathbb{R}^n} \sin^2 \frac{h\cdot\xi}{2} |\widehat{u}(\xi)|^2 d\xi .$$

It follows that

$$\begin{aligned}
\iint_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|u(x) - u(y)|^2}{|x-y|^{n+2s}}\, dxdy &= \iint_{\mathbb{R}^n\times\mathbb{R}^n} \frac{\sin^2 \frac{h\cdot\xi}{2}}{|h|^{n+2s}} |\widehat{u}(\xi)|^2 d\xi dh \\
&= \int_{\mathbb{R}^n} |\widehat{u}(\xi)|^2 \Big[\int_{\mathbb{R}^n} \frac{\sin^2 \frac{h\cdot\xi}{2}}{|h|^{n+2s}} dh \Big] d\xi \\
\text{(letting } h = 2|\xi|^{-1}z\text{)} &= 2^{-2s} \int_{\mathbb{R}^n} |\xi|^{2s} |\widehat{u}(\xi)|^2 \Big[\int_{\mathbb{R}^n} \frac{\sin^2 (z\cdot\frac{\xi}{|\xi|})}{|z|^{n+2s}} dz \Big] d\xi .
\end{aligned}$$

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

$$\iint\limits_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|u(x)-u(y)|^2}{|x-y|^{n+2s}}\,dxdy = C\int_{\mathbb{R}^n} |\xi|^{2s}|\widehat{u}(\xi)|^2 d\xi\,,$$

where $C = \displaystyle\int_{\mathbb{R}^n} \frac{\sin^2 z_1}{|z|^{n+2s}} dz < \infty$ since $0 < s < 1$. □

Corollary 2.123. *For* $0 < s < 1$*,*

$$\|u\|_{H^s(\mathbb{R}^n)} = \Big[\|u\|^2_{L^2(\mathbb{R}^n)} + \iint\limits_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|u(x)-u(y)|^2}{|x-y|^{n+2s}}\,dxdy\Big]^{\frac{1}{2}}$$

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

$$\iint\limits_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|D^\alpha u(x)-D^\alpha u(y)|^2}{|x-y|^{n+2(s-[s])}}\,dxdy < \infty$$

for all $|\alpha| = [s]$. Moreover, an equivalent norm on $H^s(\mathbb{R}^n)$ is given by

$$\begin{aligned}\|u\|_{H^s(\mathbb{R}^n)} &= \Big[\sum_{|\alpha|\leqslant[s]} \|D^\alpha u\|^2_{L^2(\Omega(\mathbb{R}^n)} + \iint\limits_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|D^\alpha u(x)-D^\alpha u(y)|^2}{|x-y|^{n+2(s-[s])}}\,dxdy\Big]^{\frac{1}{2}} \\ &= \Big[\|u\|^2_{H^{[s]}(\mathbb{R}^n)} + \iint\limits_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|D^\alpha u(x)-D^\alpha u(y)|^2}{|x-y|^{n+2(s-[s])}}\,dxdy\Big]^{\frac{1}{2}}. \end{aligned} \tag{2.66}$$

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 2.124. *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 2.122, 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 $|(\varphi u)(x) - (\varphi u)(y)| \leqslant |\varphi(x) - \varphi(y)||u(x)| + |u(x) - u(y)||\varphi(y)|$,

$$\begin{aligned}
&\iint_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|(\varphi u)(x) - (\varphi u)(y)|^2}{|x-y|^{n+2s}}\,dxdy \\
&\qquad \leqslant 2 \iint_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|\varphi(x) - \varphi(y)|^2|u(x)|^2 + |u(x)-u(y)|^2|\varphi(y)|^2}{|x-y|^{n+2s}}\,dxdy \\
&\qquad \leqslant \underbrace{2 \iint_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|\varphi(x) - \varphi(y)|^2|u(x)|^2}{|x-y|^{n+2s}}\,dxdy}_{\mathcal{I}_1} + \underbrace{2\|\varphi\|_{L^\infty(\mathbb{R}^n)} \iint_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|u(x)-u(y)|^2}{|x-y|^{n+2s}}\,dxdy}_{\mathcal{I}_2}.
\end{aligned}$$

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,

$$\begin{aligned}
&\int_{\mathbb{R}^n} \int_{|x-y|\leqslant 1} \frac{|\varphi(x) - \varphi(y)|^2|u(x)|^2}{|x-y|^{n+2s}}\,dxdy \\
&\qquad\qquad \leqslant C \int_{\mathbb{R}^n} \int_{|x-y|\leqslant 1} |x-y|^{2-n-2s}|u(x)|^2\,dxdy \\
&\qquad\qquad \leqslant C \int_{|z|\leqslant 1} |z|^{2-n-2s}dz \int_{\mathbb{R}^n} |u(x)|^2\,dx < \infty \qquad \text{if } s < 1\,.
\end{aligned}$$

For the remaining integral,

$$\begin{aligned}
&\int_{\mathbb{R}^n} \int_{|x-y|\geqslant 1} \frac{|\varphi(x) - \varphi(y)|^2|u(x)|^2}{|x-y|^{n+2s}}\,dxdy \\
&\qquad\qquad \leqslant 4\|\varphi\|^2_{L^\infty(\mathbb{R}^n)} \int_{\mathbb{R}^n} \int_{|x-y|\leqslant 1} |x-y|^{-n-2s}|u(x)|^2\,dxdy \\
&\qquad\qquad \leqslant 4\|\varphi\|^2_{L^\infty(\mathbb{R}^n)} \int_{|z|\geqslant 1} |z|^{-n-2s}dz \int_{\mathbb{R}}^{n} |u(x)|^2 < \infty \qquad \text{if } s > 0\,.
\end{aligned}$$

The general case of $s \geqslant 0$ can be proved in a similar fashion, and we leave the details to the reader. ▫

The following theorem shows that $H^s(\mathbb{R}^n)$ is a *multiplicative-algebra*; that is, $fg \in H^s(\mathbb{R}^n)$ if $f, g \in H^s(\mathbb{R}^n)$, provided that $s > \dfrac{n}{2}$.

Theorem 2.125. *Let $s > \frac{\mathrm{n}}{2}$ be a real number. Then there exists a generic constant $\mathrm{C}_s > 0$ such that*

$$\|uv\|_{H^s(\mathbb{R}^\mathrm{n})} \leqslant \mathrm{C}_s \|u\|_{H^s(\mathbb{R}^\mathrm{n})} \|v\|_{H^s(\mathbb{R}^\mathrm{n})} \qquad \forall\, u, v \in H^s(\mathbb{R}^\mathrm{n})\,. \tag{2.67}$$

Proof. Assume that $u, v \in H^s(\mathbb{R}^\mathrm{n})$. 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

$$\begin{aligned} \langle\xi\rangle^s(\widehat{u} \star \widehat{v})(\xi) &= \int_{\mathbb{R}^\mathrm{n}} \langle\xi\rangle^s \big|\widehat{u}(\xi-\eta)\widehat{v}(\eta)\big| d\eta \\ &\leqslant C_s \int_{\mathbb{R}^\mathrm{n}} \big[\langle\xi-\eta\rangle^s + \langle\eta\rangle^s\big]\big|\widehat{u}(\xi-\eta)\widehat{v}(\eta)\big| d\eta \\ &= C_s\Big[\int_{\mathbb{R}^\mathrm{n}} \big|\langle\xi-\eta\rangle^s\widehat{u}(\xi-\eta)\big|\big|\widehat{v}(\eta)\big| d\eta + \int_{\mathbb{R}^\mathrm{n}} \big|\widehat{u}(\xi-\eta)\big|\big|\langle\eta\rangle^s\widehat{v}(\eta)\big| d\eta\Big] \\ &= C_s\Big[\big(|\widehat{u_s}| \star |\widehat{v}|\big)(\xi) + \big(|\widehat{u}| \star |\widehat{v_s}|\big)(\xi)\Big]\,, \end{aligned}$$

where $w_s(x) \equiv \int_{\mathbb{R}^\mathrm{n}} \langle\xi\rangle^s \widehat{w}(\xi) e^{ix\cdot\xi} d\mu_\mathrm{n}(\xi)$ is the inverse Fourier transform of $\langle\cdot\rangle^s\widehat{w}(\cdot)$. As a consequence,

$$\begin{aligned} \|uv\|_{H^s(\mathbb{R}^\mathrm{n})} &= \Big[\int_{\mathbb{R}^\mathrm{n}} \langle\xi\rangle^{2s} |\widehat{uv}(\xi)|^2 d\xi\Big]^{\frac{1}{2}} = \Big[\int_{\mathbb{R}^\mathrm{n}} \langle\xi\rangle^{2s} |(\widehat{u} \star \widehat{v})(\xi)|^2 d\xi\Big]^{\frac{1}{2}} \\ &= \frac{1}{\sqrt{2\pi}^{\mathrm{n}}}\Big[\int_{\mathbb{R}^\mathrm{n}} \langle\xi\rangle^{2s} |(\widehat{u} * \widehat{v})(\xi)|^2 d\xi\Big]^{\frac{1}{2}} \\ &\leqslant \frac{C_s}{\sqrt{2\pi}^{\mathrm{n}}} \big\| |\widehat{u_s}| * |\widehat{v}| + |\widehat{u}| * \|\widehat{v_s}| \big\|_{L^2(\mathbb{R}^\mathrm{n})}\,, \end{aligned}$$

while Plancherel's formula and Young's inequality further imply that

$$\begin{aligned} \big\| |\widehat{u_s}| * |\widehat{v}| + |\widehat{u}| * \|\widehat{v_s}| \big\|_{L^2(\mathbb{R}^\mathrm{n})} &\leqslant \big\| |\widehat{u_s}| * |\widehat{v}| \big\|_{L^2(\mathbb{R}^\mathrm{n})} + \big\| |\widehat{u}| * \|\widehat{v_s}| \big\|_{L^2(\mathbb{R}^\mathrm{n})} \\ &\leqslant \|\widehat{u_s}\|_{L^2(\mathbb{R}^\mathrm{n})} \|\widehat{v}\|_{L^1(\mathbb{R}^\mathrm{n})} + \|\widehat{u}\|_{L^1(\mathbb{R}^\mathrm{n})} \|\widehat{v_s}\|_{L^2(\mathbb{R}^\mathrm{n})} \\ &= \|u_s\|_{L^2(\mathbb{R}^\mathrm{n})} \|\widehat{v}\|_{L^1(\mathbb{R}^\mathrm{n})} + \|\widehat{u}\|_{L^1(\mathbb{R}^\mathrm{n})} \|v_s\|_{L^2(\mathbb{R}^\mathrm{n})} \\ &\leqslant 2\|\langle\cdot\rangle^{-s}\|_{L^2(\mathbb{R}^\mathrm{n})} \|u\|_{H^s(\mathbb{R}^\mathrm{n})} \|v\|_{H^s(\mathbb{R}^\mathrm{n})}\,, \end{aligned}$$

where (2.63) is used to conclude the last inequality. Estimate (2.67) is then established by choosing $\mathrm{C}_s = \dfrac{2C_s\|\langle\cdot\rangle^{-s}\|_{L^2(\mathbb{R}^\mathrm{n})}}{\sqrt{2\pi}^{\mathrm{n}}}$. ◻

The case that $s < 0$

For $s < 0$, we define the space $H^s(\mathbb{R}^{\rm n})$ to be the dual space of $H^{-s}(\mathbb{R}^{\rm n})$ with the corresponding dual space norm (or operator norm) defined by

$$\|u\|_{H^s(\mathbb{R}^{\rm n})} = \sup_{v\in H^{-s}(\mathbb{R}^{\rm n})} \frac{\langle u,v\rangle}{\|v\|_{H^{-s}(\mathbb{R}^{\rm n})}} = \sup_{\|v\|_{H^{-s}(\mathbb{R}^{\rm n})}=1} \langle u,v\rangle\,. \tag{2.68}$$

If $u \in H^s(\mathbb{R}^{\rm n})$ for some $s < 0$, the Riesz representation theorem suggests the existence of $w \in H^{-s}(\mathbb{R}^{\rm n})$ satisfying

$$\langle u,v\rangle = (w,v)_{H^{-s}(\mathbb{R}^{\rm n})} \qquad \forall\, v \in H^{-s}(\mathbb{R}^{\rm n})\,, \tag{2.69}$$

and the operator norm of u is the same as the $H^{-s}(\mathbb{R}^{\rm n})$-norm of w; that is,

$$\|u\|_{H^s(\mathbb{R}^{\rm n})} = \|w\|_{H^{-s}(\mathbb{R}^{\rm n})}\,. \tag{2.70}$$

On the other hand, by the definition of the Fourier transform of a tempeblack distribution,

$$\begin{aligned}\langle \widehat{u},\check{\varphi}\rangle = \langle u,\varphi\rangle = (w,\varphi)_{H^{-s}(\mathbb{R}^{\rm n})} &= \int_{\mathbb{R}^{\rm n}} \langle\xi\rangle^{-2s}\widehat{w}(\xi)\overline{\widehat{\varphi}(\xi)}\,d\xi \\ &= \int_{\mathbb{R}^{\rm n}} \langle\xi\rangle^{-2s}\widehat{w}(\xi)\check{\varphi}(\xi)\,d\xi \qquad \forall\,\varphi \in \mathscr{S}(\mathbb{R}^{\rm n})\,.\end{aligned}$$

The equality above implies that $\widehat{u}(\xi) = \langle\xi\rangle^{-2s}\widehat{w}(\xi)$; thus

$$\langle\xi\rangle^{s}\widehat{u}(\xi) = \langle\xi\rangle^{-s}\widehat{w}(\xi) \in L^2(\mathbb{R}^{\rm n})\,.$$

Therefore, (2.73) implies that

$$\|u\|_{H^s(\mathbb{R}^{\rm n})} = \Big[\int_{\mathbb{R}^{\rm n}} \langle\xi\rangle^{2s}|\widehat{u}(\xi)|^2 d\xi\Big]^{\frac{1}{2}}\,. \tag{2.71}$$

In other words, for all real s, the space $H^s(\mathbb{R}^{\rm n})$ can be defined as the collection of tempeblack distributions u such that $\langle\cdot\rangle^s\widehat{u}(\cdot) \in L^2(\mathbb{R}^{\rm n})$, and the $H^s(\mathbb{R}^{\rm n})$-norm of u is given by (2.74).

The dual space of $H^s(\mathbb{R}^{\rm n})$ 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

$$\langle u, v\rangle = (w, v)_{H^s(\mathbb{R}^n)} \qquad \forall\, v \in H^s(\mathbb{R}^n)\,, \tag{2.72}$$

and the operator norm of u is the same as the $H^s(\mathbb{R}^n)$-norm of w; that is,

$$\|u\|_{H^s(\mathbb{R}^n)'} = \|w\|_{H^s(\mathbb{R}^n)}\,. \tag{2.73}$$

Moreover, since $\mathscr{S}(\mathbb{R}^n) \subseteq H^s(\mathbb{R}^n)$, $u \in H^s(\mathbb{R}^n)'$ is a tempeblack distribution; thus the definition of the Fourier transform of a tempeblack 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 $\tilde{}$ 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 (2.73) that

$$\|u\|_{H^s(\mathbb{R}^n)'} = \Big[\int_{\mathbb{R}^n} \langle\xi\rangle^{-2s}|\widehat{u}(\xi)|^2 d\xi\Big]^{\frac{1}{2}} = \|u\|_{H^{-s}(\mathbb{R}^n)}\,. \tag{2.74}$$

In other words, for $s > 0$ the space $H^{-s}(\mathbb{R}^n) = H^s(\mathbb{R}^n)'$.

2.3.3 The Interpolation Spaces[2]

Given $t > 0$, let

$$K(t, u; r, s) \equiv \inf_{\substack{u=u_r+u_s \\ u_r\in H^r(\mathbb{R}^n), u_s\in H^s(\mathbb{R}^n)}} \Big[\|u_r\|^2_{H^r(\mathbb{R}^n)} + t^2\|u_s\|^2_{H^s(\mathbb{R}^n)}\Big]^{1/2}.$$

For $\alpha \in (0,1)$, define the interpolation space $\big(H^r(\mathbb{R}^n), H^s(\mathbb{R}^n)\big)_\alpha$ by

$$\big(H^r(\mathbb{R}^n), H^s(\mathbb{R}^n)\big)_\alpha \equiv \Big\{u \,\Big|\, \int_0^\infty t^{-1-2\alpha} K(t, u; r, s)^2 dt < \infty\Big\} \tag{2.75}$$

equipped with norm

$$\|u\|_{(H^r(\mathbb{R}^n), H^s(\mathbb{R}^n))_\alpha} \equiv \Big[\int_0^\infty t^{-1-2\alpha} K(t, u; r, s)^2 dt\Big]^{1/2}.$$

Our first goal in this section is to show that the space $\big(H^r(\mathbb{R}^n), H^s(\mathbb{R}^n)\big)_\alpha$ is the same as $H^{\alpha s+(1-\alpha)r}(\mathbb{R}^n)$. To be more precise, we shall prove the following

Proposition 2.126. *Let $0 < r < s < \infty$, and $\alpha \in (0,1)$. Then*

$$\int_0^\infty t^{-1-2\alpha} K(t, u; r, s)^2 dt = C_\alpha \|u\|^2_{H^{\alpha r+(1-\alpha)s}(\mathbb{R}^n)}, \tag{2.76}$$

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 $H^r(\mathbb{R}^n)$,

$$K(t, u; r, s) = \inf_{\substack{u=u_r+u_s \\ u_r\in H^r(\mathbb{R}^n), u_s\in H^s(\mathbb{R}^n)}} \Big[\int_{\mathbb{R}^n} \big(|\widehat{u_r}(\xi)|^2\langle\xi\rangle^{2r} + t^2|\widehat{u_s}(\xi)|^2\langle\xi\rangle^{2s}\big) d\xi\Big]^{1/2},$$

where we recall that $\langle\xi\rangle = (1+|\xi|^2)^{1/2}$. For each $\xi \in \mathbb{R}^n$, 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

$$r(\xi) = \frac{t^2\langle\xi\rangle^{2s}}{\langle\xi\rangle^{2r} + t^2\langle\xi\rangle^{2s}}\mathrm{Re}\big(\widehat{u}(\xi)\, e^{i\theta(\xi)}\big) \quad \text{and} \quad \theta(\xi) = \arg\big(\widehat{u}(\xi)\big).$$

[2]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

$$\begin{aligned} \int_0^\infty t^{-1-2\alpha}K(t,u;r,s)^2 dt &= \int_{\mathbb{R}^n}\int_0^\infty \frac{t^{1-2\alpha}\langle\xi\rangle^{2r}}{t^2+\langle\xi\rangle^{2(r-s)}}|\widehat{u}(\xi)|^2 dt d\xi \\ {\scriptstyle (t\,=\,\langle\xi\rangle^{r-s}t')} &= C_\alpha\int_{\mathbb{R}^n}\langle\xi\rangle^{2\alpha r+2(1-\alpha)s}|\widehat{u}(\xi)|^2 d\xi \\ &= C_\alpha\|u\|^2_{H^{\alpha s+(1-\alpha)r}(\mathbb{R}^n)}\,, \end{aligned}$$

where the constant C_α is given by $\displaystyle\int_0^\infty \frac{t^{1-2\alpha}}{t^2+1}dt$. □

Theorem 2.127. *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

$$\|Au\|_{H^{\alpha s_2+(1-\alpha)r_2}(\mathbb{R}^n)} \leqslant \sqrt{2}\,\mathrm{M}_0^{1-\alpha}\mathrm{M}_1^{\alpha}\|u\|_{H^{\alpha s_1+(1-\alpha)r_1}(\mathbb{R}^n)}\,. \tag{2.77}$$

Proof. Let $u \in H^{\alpha s_1+(1-\alpha)r_1}(\mathbb{R}^n)$. By Proposition 2.126

$$\begin{aligned} G_1(t,u) \equiv & \inf_{\substack{u=u_s+u_r \\ u_r\in H^{r_1}(\mathbb{R}^n),u_s\in H^{s_1}(\mathbb{R}^n)}}\Big[\|u_r\|_{H^{r_1}(\mathbb{R}^n)} + t\|u_s\|_{H^{s_1}(\mathbb{R}^n)}\Big] \\ & \leqslant \sqrt{2}\,K(t,u;r_1,s_1) < \infty \end{aligned}$$

for almost all $t \in (0,\infty)$. For each decomposition $u = u_s + u_r$ with $u_s \in H^{s_1}(\mathbb{R}^{\mathrm{n}})$ and $u_r \in H^{r_1}(\mathbb{R}^{\mathrm{n}})$, we have $Au = Au_r + Au_s$; thus

$$\begin{aligned}
G_2(t, Au) &\equiv \inf_{\substack{Au=v_r+v_s \\ v_r\in H^{r_2}(\mathbb{R}^{\mathrm{n}}), v_s\in H^{s_2}(\mathbb{R}^{\mathrm{n}})}} \Big[\|v_r\|_{H^{r_2}(\mathbb{R}^{\mathrm{n}})} + t\|v_s\|_{H^{s_2}(\mathbb{R}^{\mathrm{n}})}\Big] \\
&\leqslant \|Au_r\|_{H^{r_2}(\mathbb{R}^{\mathrm{n}})} + t\|Au_s\|_{H^{s_2}(\mathbb{R}^{\mathrm{n}})} \leqslant \mathrm{M}_0\|u_r\|_{H^{r_1}(\mathbb{R}^{\mathrm{n}})} + t\mathrm{M}_1\|u_s\|_{H^{s_1}(\mathbb{R}^{\mathrm{n}})} \\
&= \mathrm{M}_0\Big(\|u_s\|_{H^{s_1}(\mathbb{R}^{\mathrm{n}})} + \frac{t\mathrm{M}_1}{\mathrm{M}_0}\|u_r\|_{H^{r_1}(\mathbb{R}^{\mathrm{n}})}\Big)\,.
\end{aligned}$$

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, (2.76) suggests that

$$\begin{aligned}
C_\alpha\|Au\|^2_{H^{\alpha s_2+(1-\alpha)r_2}(\mathbb{R}^{\mathrm{n}})} &= \int_0^\infty t^{-1-2\alpha}K(t, Au; r_2, s_2)^2 dt \\
&\leqslant 2\int_0^\infty t^{-1-2\alpha}\mathrm{M}_0^2 K\Big(\frac{t\mathrm{M}_1}{\mathrm{M}_0}, u; r_1, s_1\Big)^2 dt \\
&= 2\int_0^\infty \mathrm{M}_1^{2\alpha}\mathrm{M}_0^{2-2\alpha}\tilde{t}^{-1-2\alpha}K(\tilde{t}, u; r_1, s_1)^2 d\tilde{t} \\
&= 2C_\alpha \mathrm{M}_0^{2-2\alpha}\mathrm{M}_1^{2\alpha}\|u\|^2_{H^{\alpha s_1+(1-\alpha)r_1}(\mathbb{R}^{\mathrm{n}})}\,.
\end{aligned}$$

□

Corollary 2.128. *Let $s > \frac{\mathrm{n}}{2}$ be a real number. Then there exists a generic constant $\mathrm{C}_s > 0$ such that for all $0 \leqslant r \leqslant s$,*

$$\|uv\|_{H^r(\mathbb{R}^{\mathrm{n}})} \leqslant C_s\|u\|_{H^s(\mathbb{R}^{\mathrm{n}})}\|v\|_{H^r(\mathbb{R}^{\mathrm{n}})} \qquad \forall\, u \in H^s(\mathbb{R}^{\mathrm{n}}) \text{ and } v \in H^r(\mathbb{R}^{\mathrm{n}})\,. \tag{2.78}$$

Proof. Let $u \in H^s(\mathbb{R}^{\mathrm{n}})$ be given, and define a linear map A by $Av = uv$. Then Theorem 2.125 implies that $A \in \mathscr{B}(H^s(\mathbb{R}^{\mathrm{n}}), H^s(\mathbb{R}^{\mathrm{n}}))$ with estimate

$$\|Av\|_{H^s(\mathbb{R}^{\mathrm{n}})} \leqslant \mathrm{C}_s\|u\|_{H^s(\mathbb{R}^{\mathrm{n}})}\|v\|_{H^s(\mathbb{R}^{\mathrm{n}})} \qquad \forall\, v \in H^s(\mathbb{R}^{\mathrm{n}})\,.$$

Moreover, by the Sobolev embedding (Theorem 2.119),

$$\|Av\|_{L^2(\mathbb{R}^{\mathrm{n}})} \leqslant \|u\|_{L^\infty(\mathbb{R}^{\mathrm{n}})}\|v\|_{L^2(\mathbb{R}^{\mathrm{n}})} \leqslant C\|u\|_{H^s(\mathbb{R}^{\mathrm{n}})}\|v\|_{L^2(\mathbb{R}^{\mathrm{n}})} \qquad \forall\, v \in L^2(\mathbb{R}^{\mathrm{n}})$$

which implies that $A \in \mathscr{B}(L^2(\mathbb{R}^n), L^2(\mathbb{R}^n))$. Therefore, Theorem 2.127 implies that $A \in \mathscr{B}(H^r(\mathbb{R}^n), H^r(\mathbb{R}^n))$, and for $v \in H^r(\mathbb{R}^n)$,

$$\begin{aligned}\|Av\|_{H^r(\mathbb{R}^n)} &\leqslant \sqrt{2}\big(\mathrm{C}_s\|u\|_{H^s(\mathbb{R}^n)}\big)^{\frac{s-r}{s}}\big(C\|u\|_{H^s(\mathbb{R}^n)}\big)^{\frac{r}{s}}\|v\|_{H^r(\mathbb{R}^n)}\\ &= C_s\|u\|_{H^s(\mathbb{R}^n)}\|v\|_{H^r(\mathbb{R}^n)}\,,\end{aligned}$$

where $C_s = \max\limits_{r\in[0,s]} \sqrt{2}C^{r/s}\mathrm{C}_s^{(s-r)/s} < \infty$. □

2.4 Fractional-Order Sobolev Spaces on Domains with Boundary

2.4.1 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

$$\|u\|_{H^k(\mathbb{R}^n_+)} = \Big[\sum_{|\alpha|\leqslant k} \|D^\alpha u\|^2_{L^2(\mathbb{R}^n_+)}\Big]^{\frac{1}{2}}\,. \tag{2.79}$$

Note that we are not able to directly use the Fourier transform to define the $H^k(\mathbb{R}^n_+)$.

Definition 2.129 (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

$$(Eu)(x) = \begin{cases} u(x) & \text{if} \quad x_n \geqslant 0\,,\\ \sum\limits_{j=1}^N a_j u(x', -2^{1-j}x_n) & \text{if} \quad x_n < 0\,.\end{cases} \tag{2.80}$$

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 2.130 (Sobolev extension theorem)**.** *For any $N \in \mathbb{N}$, the operator E defined in* (2.80) *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 2.131. The extension operator E given by (2.80) 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 2.132. *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 2.130, 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 (2.80), 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

$$\|u\|_{H^s(\mathbb{R}^n_+)} := \|Eu\|_{H^s(\mathbb{R}^n)}\,. \tag{2.81}$$

When $s = k \in \mathbb{N}$, it may not be immediately clear that the $H^s(\mathbb{R}^n_+)$-norm defined by (2.81) is equivalent to the $H^k(\mathbb{R}^n_+)$-norm defined by (2.79). Let $\|\cdot\|_1$ be the norm defined by (2.79) and $\|\cdot\|_2$ be the norm defined by (2.81). 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 2.122 (or (2.66)), we in fact have the following

Theorem 2.133. *For $s > 0$ and $s \notin \mathbb{N}$, then $\|\cdot\|_{H^s(\mathbb{R}^n_+)}$ is equivalent to the norm*

$$|||u|||_{H^s(\mathbb{R}^n_+)} := \Big[\|u\|^2_{H^{[s]}(\mathbb{R}^n_+)} + \sum_{|\alpha|=[s]} \iint\limits_{\mathbb{R}^n_+\times\mathbb{R}^n_+} \frac{|D^\alpha u(x) - D^\alpha u(y)|^2}{|x-y|^{n+2(s-[s])}}\, dxdy\Big]^{\frac{1}{2}}.$$

Proof. Recall that the norm $|||w|||_{H^s(\mathbb{R}^n)}$ defined by

$$|||w|||_{H^s(\mathbb{R}^n)} := \Big[\|w\|^2_{H^{[s]}(\mathbb{R}^n)} + \sum_{|\alpha|=[s]} \iint\limits_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|D^\alpha w(x) - D^\alpha w(y)|^2}{|x-y|^{n+2(s-[s])}}\, dxdy\Big]^{\frac{1}{2}}$$

is equivalent to the norm $\|w\|_{H^s(\mathbb{R}^n)}$, so it is clear that $|||u|||_{H^s(\mathbb{R}^n_+)} \leqslant C_1\|Eu\|_{H^s(\mathbb{R}^n)}$ for some constant $C_1 > 0$ since $Eu = u$ on $\mathbb{R}^n_+$.

For the reversed inequality, since

$$\begin{aligned}
&\iint\limits_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|D^\alpha(Eu)(x) - D^\alpha(Eu)(y)|^2}{|x-y|^{n+2(s-[s])}}\, dxdy \\
&\qquad = \Big(\iint\limits_{\mathbb{R}^n_+\times\mathbb{R}^n_+} + \iint\limits_{\mathbb{R}^n_+\times\mathbb{R}^n_-} + \iint\limits_{\mathbb{R}^n_-\times\mathbb{R}^n_+} + \iint\limits_{\mathbb{R}^n_-\times\mathbb{R}^n_-} \Big) \frac{|D^\alpha(Eu)(x) - D^\alpha(Eu)(y)|^2}{|x-y|^{n+2(s-[s])}}\, dxdy\,,
\end{aligned}$$

by the boundedness of the extension operator we find that

$$\begin{aligned}
&|||Eu|||^2_{H^s(\mathbb{R}^n)} \\
&\quad = \|Eu\|^2_{H^{[s]}(\mathbb{R}^n)} + \sum_{|\alpha|=[s]} \iint\limits_{\mathbb{R}^n\times\mathbb{R}^n} \frac{|D^\alpha(Eu)(x) - D^\alpha(Eu)(y)|^2}{|x-y|^{n+2(s-[s])}}\, dxdy \\
&\quad \leqslant C\Big[|||u|||^2_{H^s(\mathbb{R}^n_+)} + \sum_{j\leqslant|\alpha|=[s]} \iint\limits_{\mathbb{R}^n_-\times\mathbb{R}^n_-} \frac{|(D^\alpha u)(x', -jx_n) - (D^\alpha u)(y', -jy_n)|^2}{|x-y|^{n+2(s-[s])}}\, dxdy \\
&\qquad\qquad + \sum_{j\leqslant|\alpha|=[s]} \iint\limits_{\mathbb{R}^n_-\times\mathbb{R}^n_+} \frac{|(D^\alpha u)(x', -jx_n) - (D^\alpha u)(y', y_n)|^2}{|(x', x_n) - (y', y_n)|^{n+2(s-[s])}}\, dxdy\Big] \\
&\quad \leqslant C\Big[|||u|||^2_{H^s(\mathbb{R}^n_+)} + \sum_{|\alpha|=[s]} \iint\limits_{\mathbb{R}^n_+\times\mathbb{R}^n_+} \frac{|D^\alpha u(x) - D^\alpha u(y)|^2}{|x-y|^{n+2(s-[s])}}\, dxdy\Big]
\end{aligned}$$

which implies $\|Eu\|_{H^s(\mathbb{R}^n)} \leqslant C|\!|\!|Eu|\!|\!|_{H^s(\mathbb{R}^n)} \leqslant C_2|\!|\!|u|\!|\!|_{H^s(\mathbb{R}^n_+)}$ for some constant $C_2 > 0$; thus the equivalence of these two norms is established. □

2.4.2 The Sobolev Space $H^s(\Omega)$

We can now define the Sobolev spaces $H^s(\Omega)$ for any open and bounded domain $\Omega \subseteq \mathbb{R}^n$ with smooth boundary $\partial\Omega$.

Definition 2.134 (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 $\gamma : \mathbb{R}^{n-1} \to \mathbb{R}$ such that - upon relabeling and reorienting the coordinates axes if necessary - we have

$$\Omega \cap B(x_0, r) = \left\{x \in B(x_0, r) \,\middle|\, x_n > \gamma(x_1, \cdots, x_{n-1})\right\}.$$

$\partial\Omega$ 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 2.135 (Partition of unity)**.** Let X be a topological space. A partition of unity is a collection of continuous functions $\left\{\chi_j : X \to [0,1]\right\}$ 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 2.136. *Let $\Omega \subseteq \mathbb{R}^n$ 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 $\left\{\sqrt{\zeta_m}\right\}_{m=1}^K \subseteq \mathscr{C}_c^\infty(\mathbb{R}^n)$.*

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 $\left\{\mathcal{U}_m^{(\delta)}\right\}_{m=1}^\infty$ 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 $\left\{x_{k_j}\right\}_{j=1}^\infty$ 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 $\bar{\Omega} \subseteq \bigcup_{m=1}^{K} \mathcal{U}_m^{(\delta)}$ 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 $0 \leqslant \zeta_m \leqslant 1$, $\text{spt}(\zeta_m) \subseteq \mathcal{U}_m$, and $\sqrt{\zeta_m} \in \mathscr{C}_c^\infty(\mathbb{R}^{\rm n})$ for all $1 \leqslant m \leqslant K$, and $\sum_{m=1}^{K} \xi_m = \sum_{m=1}^{K} \zeta_m = 1$. In other words, $\{\zeta_m\}_{m=1}^K$ is a partition of unity subordinate to $\{\mathcal{U}_m\}_{m=1}^K$ satisfying that $\sqrt{\zeta_m} \in \mathscr{C}_c^\infty(\mathbb{R}^{\rm n})$ for all $1 \leqslant m \leqslant K$. □

For a domain Ω with smooth boundary, we may assume that there exist $x_1, \cdots x_N \in \partial\Omega$, $r_1, \cdots r_N > 0$, $\gamma_j \in \mathscr{C}^\infty$ such that, upon relabeling and reorienting the coordinates axes if necessary,

$$\Omega \cap \mathcal{U}_j = \big\{x \in \mathcal{U}_j \,\big|\, x_{\rm n} > \gamma_j(x_1, \cdots, x_{{\rm n}-1})\big\} \text{ where } \mathcal{U}_j = B(x_j, r_j)\,,$$

and $\Omega \subseteq \bigcup_{j=0}^{N} \mathcal{U}_j$, and $\{\chi_j\}_{j=0}^N$ is a partition of unity subordinate to the open cover $\{\mathcal{U}_j\}_{j=0}^N$ such that $\chi_j \in \mathscr{C}_c^\infty(\mathcal{U}_j)$, and the function ψ_j defined by

$$\psi_j(x) = (x_1, \cdots, x_{{\rm n}-1}, \gamma_j(x_1, \cdots, x_{{\rm n}-1}) + x_{\rm n})\,.$$

is a diffeomorphism between $\mathcal{U}_j$ and a small neighborhood $\mathcal{V}_j$ of $\mathbb{R}^{\rm n}$.

Let $v_0 = \chi_0 u$ and $v_j = (\chi_j u) \circ \psi_j$. Then v_0 can be consideblack as a function defined on $\mathbb{R}^{\rm n}$, and v_j can be consideblack as a function defined on $\mathbb{R}^{\rm n}_+$. We then have the following definition.

Definition 2.137. The space $H^s(\Omega)$ for $s > 0$ is the collection of all measurable functions u such that $\chi_0 u \in H^s(\mathbb{R}^{\rm n})$ and $(\chi_j u) \circ \psi_j \in H^s(\mathbb{R}^{\rm n}_+)$. The $H^s(\Omega)$-norm is defined by

$$\|u\|_{H^s(\Omega)} = \Big[\|\chi_0 u\|^2_{H^s(\mathbb{R}^{\rm n})} + \sum_{j=1}^N \|(\chi_j u) \circ \psi_j\|^2_{H^s(\mathbb{R}^{\rm n}_+)}\Big]^{1/2}\,.$$

Theorem 2.138 (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 2.136. 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 (2.80) for some $k \geqslant s$. Then E satisfies properties (i)-(iii), and the proofs of these three properties are left to the readers. □

Theorem 2.139 (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 $\lim_{j\to\infty} \|u_{j_k} - u\|_{H^s(\Omega)} = 0$.

Let E be the extension operator defined in Theorem 2.138, 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,

$$\begin{aligned}\|v_j^\epsilon - v_j\|_{H^s(\mathbb{R}^n)}^2 &= \int_{\mathbb{R}^n} \big|\sqrt{2\pi}^{\,n}\widehat{\eta}(\epsilon\xi) - 1\big|^2 \langle\xi\rangle^{2s} \big|\widehat{v}_j(\xi)\big|^2 d\xi \\ &\leqslant \Big[\frac{4\epsilon^\delta}{(1+\epsilon)^\delta} + 4\sin^2\frac{\sqrt{\epsilon}}{2}\Big]\|v_j\|_{H^{s+\delta}(\mathbb{R}^n)}^2\end{aligned}$$

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

$$\|v_j^\epsilon - v_j\|_{H^s(\mathbb{R}^n)} < \frac{\epsilon'}{3} \qquad \forall\, j \in \mathbb{N}\,. \tag{2.82}$$

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 (2.64) implies that

$$\limsup_{k,\ell\to\infty} \|v_{j_k}^\epsilon - v_{j_\ell}^\epsilon\|_{H^s(\mathbb{R}^n)} \leqslant C \limsup_{k,\ell\to\infty} \|v_{j_k}^\epsilon - v_{j_\ell}^\epsilon\|_{L^2(\mathbb{R}^n)}^{\frac{\delta}{s+\delta}} = 0\,.$$

As a consequence, there exists $N > 0$ such that

$$\|v_{j_k}^\epsilon - v_{j_\ell}^\epsilon\|_{H^s(\mathbb{R}^n)} < \frac{\epsilon'}{3} \quad \text{whenever} \quad k,\ell \geqslant N\,. \tag{2.83}$$

The triangle inequality together with (2.82) and (2.83) 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 2.120, Theorem 2.127 and Corollary 2.128. The proofs for these two theorems are left as an exercise.

THEOREM 2.140 (Interpolation inequality). *Let $\Omega \subseteq \mathbb{R}^{\rm n}$ be a bounded smooth domain, $0 < r < t < \infty$, and $s = \alpha r + (1-\alpha)t$ for some $\alpha \in (0,1)$. Then*

$$\|u\|_{H^s(\Omega)} \leqslant C\|u\|^{\alpha}_{H^r(\Omega)}\|u\|^{1-\alpha}_{H^t(\Omega)} . \tag{2.84}$$

THEOREM 2.141. *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

$$\|Au\|_{H^{\alpha s_2+(1-\alpha)r_2}(\Omega)} \leqslant C\mathrm{M}_0^{1-\alpha}\mathrm{M}_1^{\alpha}\|u\|_{H^{\alpha s_1+(1-\alpha)r_1}(\Omega)} \tag{2.85}$$

for some generic constant $C > 0$ (independent of u).

THEOREM 2.142. *Let $\Omega \subseteq \mathbb{R}^{\rm n}$ be a bounded smooth domain, and $s > \frac{\rm n}{2}$ be a real number. Then there exists a generic constant $\mathrm{C}_s > 0$ such that for all $0 \leqslant r \leqslant s$,*

$$\|uv\|_{H^r(\Omega)} \leqslant C_s\|u\|_{H^s(\Omega)}\|v\|_{H^r(\Omega)} \qquad \forall\, u \in H^s(\Omega) \text{ and } v \in H^r(\Omega) . \tag{2.86}$$

2.5 The Sobolev Spaces $H^s(\mathbb{T}^{\rm n})$, $s \in \mathbb{R}$

2.5.1 The Fourier Series: Revisited

DEFINITION 2.143. For $u \in L^1(\mathbb{T}^{\rm n})$, define

$$(\mathscr{F}u)(k) = \widehat{u}_k = (2\pi)^{-\frac{\rm n}{2}} \int_{\mathbb{T}^{\rm n}} e^{-ik\cdot x}u(x)\,dx ,$$

and for $\widehat{u} \in \ell^1(\mathbb{Z}^{\rm n})$, define

$$(\mathscr{F}^*\widehat{u})(x) = (2\pi)^{-\frac{\rm n}{2}} \sum_{k\in\mathbb{Z}^{\rm n}} \widehat{u}_k e^{ik\cdot x} .$$

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 2.144. 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 2.145. 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

$$\mathscr{F} : \mathscr{C}^\infty(\mathbb{T}^{\mathrm{n}}) \to \mathfrak{s}\,, \quad \mathscr{F}^* : \mathfrak{s} \to \mathscr{C}^\infty(\mathbb{T}^{\mathrm{n}})\,,$$

and $\mathscr{F}^*\mathscr{F} = \mathrm{Id}$ on $\mathscr{C}^\infty(\mathbb{T}^{\mathrm{n}})$ and $\mathscr{F}\mathscr{F}^* = \mathrm{Id}$ on $\mathfrak{s}$. These properties smoothly extend to the Hilbert space setting:

$$\begin{array}{ll} \mathscr{F} : L^2(\mathbb{T}^{\mathrm{n}}) \to \ell^2(\mathbb{Z}^{\mathrm{n}}) & \mathscr{F}^* : \ell^2(\mathbb{Z}^{\mathrm{n}}) \to L^2(\mathbb{T}^{\mathrm{n}}) \\ \mathscr{F}^*\mathscr{F} = \mathrm{Id} \text{ on } L^2(\mathbb{T}^{\mathrm{n}}) & \mathscr{F}\mathscr{F}^* = \mathrm{Id} \text{ on } \ell^2(\mathbb{Z}^{\mathrm{n}})\,. \end{array}$$

Definition 2.146. The inner-products on $L^2(\mathbb{T}^{\mathrm{n}})$ and $\ell^2(\mathbb{Z}^{\mathrm{n}})$ are

$$(u, v)_{L^2(\mathbb{T}^{\mathrm{n}})} = \int_{\mathbb{T}^{\mathrm{n}}} u(x)\overline{v(x)}\, dx$$

and

$$(\widehat{u}, \widehat{v})_{\ell^2(\mathbb{Z}^{\mathrm{n}})} = \sum_{k \in \mathbb{Z}^{\mathrm{n}}} \widehat{u}_k \bar{\widehat{v}}_k\,,$$

respectively.

Parseval's identity shows that $\|u\|_{L^2(\mathbb{T}^{\mathrm{n}})} = \|\widehat{u}\|_{\ell^2(\mathbb{Z}^{\mathrm{n}})}$.

DEFINITION 2.147. 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:

$$\begin{array}{ll} \mathscr{F} : \mathscr{D}'(\mathbb{T}^n) \to \mathfrak{s}' & \mathscr{F}^* : \mathfrak{s}' \to \mathscr{D}'(\mathbb{T}^n) \\ \mathscr{F}^*\mathscr{F} = \mathrm{Id} \text{ on } \mathscr{D}'(\mathbb{T}^n) & \mathscr{F}\mathscr{F}^* = \mathrm{Id} \text{ on } \mathfrak{s}'. \end{array}$$

DEFINITION 2.148 (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 $u \in \mathscr{D}'(\mathbb{T}^n)$,

$$\Lambda^s u(x) = \sum_{k \in \mathbb{Z}^n} |\widehat{u}_k|^2 \langle k \rangle^s e^{ik\cdot x}.$$

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 2.149 (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 $H^s(\mathbb{T}^n)$.

2.5.2 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:

$$\begin{aligned} \Delta\, \mathrm{PI}(f) &= 0 \ \text{ in } \ \mathrm{D} \\ \mathrm{PI}(f) &= f \ \text{ on } \ \partial \mathrm{D} = \mathbb{S}^1 \,. \end{aligned}$$

$\mathrm{PI}(f)$ has an explicit representation via the Fourier series

$$\mathrm{PI}(f)(r,\theta) = \sum_{k\in\mathbb{Z}} \widehat{f}_k r^{|k|} e^{ik\theta} \quad \forall\, r < 1, 0 \leqslant \theta < 2\pi\,, \tag{2.87}$$

as well as the integral representation

$$\mathrm{PI}(f)(r,\theta) = \frac{1-r^2}{2\pi} \int_{\mathbb{S}^1} \frac{f(\varphi)}{r^2 - 2r\cos(\theta-\varphi)+1} d\varphi \ \ \forall\, r < 1, 0 \leqslant \theta < 2\pi\,. \tag{2.88}$$

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 2.150. 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 $\{e^{i\ell\theta} \,|\, \ell \in \mathbb{Z}\}$ 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 $\frac{1}{r}u_\theta \in L^2(\mathrm{D})$. 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,

$$\begin{aligned}
\|\frac{1}{r}u_\theta\|^2_{L^2(\mathrm{D})} &= \int_0^1 \Big(\int_0^{2\pi} \Big|\sum_{\ell\neq 0} i\ell\widehat{f}_\ell r^{|\ell|-1} e^{i\ell\theta}\Big|^2 d\theta\Big) r dr \\
&= 2\pi \sum_{\ell\neq 0} |\ell|^2 |\widehat{f}_\ell|^2 \int_0^1 r^{2|\ell|-1} dr \leqslant \pi \|f\|^2_{H^{\frac{1}{2}}(\mathbb{S}^1)}\,.
\end{aligned}$$

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 (D.1) and the fact that $\frac{1}{|\ell| - k + 1} \leqslant \frac{k+1}{|\ell|}$ for all $|\ell| \geqslant k$,

$$\begin{aligned}
\|\partial_r^k u\|^2_{L^2(\mathrm{D})} &= \int_0^1 \Big(\int_0^{2\pi} \Big|\sum_{|\ell|\geqslant k} |\ell|(|\ell|-1)\cdots(|\ell|-k+1)\widehat{f}_\ell r^{|\ell|-k} e^{i\ell\theta}\Big|^2 d\theta\Big) r dr \\
&= \int_0^1 \sum_{|\ell|\geqslant k} |\ell|^2(|\ell|-1)^2\cdots(|\ell|-k+1)^2 |\widehat{f}_\ell|^2 r^{2|\ell|-2k+1} dr \\
&\leqslant 2\pi \sum_{|\ell|\geqslant k} |\ell|^{2k} |\widehat{f}_\ell|^2 \int_0^1 r^{2|\ell|-2k+1} dr \\
&= \sum_{|\ell|\geqslant k} \frac{|\ell|^{2k}}{|\ell|-k+1} |\widehat{f}_\ell|^2 \\
&\leqslant (k+1)\pi \sum_{|\ell|\geqslant k} |\ell|^{2k-1} |\widehat{f}_\ell|^2 \leqslant (k+1)\pi\|f\|^2_{H^{k-\frac{1}{2}}(\mathbb{S}^1)}\,. \qquad (2.89)
\end{aligned}$$

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,

$$\begin{aligned}\Big\|\frac{1}{r^k}\partial_\theta^k u\Big\|^2_{L^2(\mathrm{D})} &= 2\pi\sum_{|\ell|\geqslant 2}|\ell|^{2(k-1)}(|\ell|-1)^2|\widehat{f}_\ell|^2\int_0^1 r^{2|\ell|-2k+1}dr\\ &\quad + 2\pi\sum_{\ell\neq 0}|\ell|^{2(k-1)}(2|\ell|-1)|\widehat{f}_\ell|^2\int_0^1 r^{2|\ell|-2k+1}dr\\ &\leqslant 2\pi\sum_{|\ell|\geqslant 2}|\ell|^{2(k-1)}(|\ell|-1)^2|\widehat{f}_\ell|^2\int_0^1 r^{2|\ell|-2k+1}dr + 2\pi\|f\|^2_{H^{k-\frac12}(\mathbb{S}^1)}\\ &\leqslant 2\pi\sum_{|\ell|\geqslant 3}|\ell|^{2(k-2)}(|\ell|-1)^2(|\ell|-2)^2|\widehat{f}_\ell|^2\int_0^1 r^{2|\ell|-2k+1}dr\\ &\quad + 2\pi(1+2)\|f\|^2_{H^{k-\frac12}(\mathbb{S}^1)}\\ &\leqslant \cdots\cdots\\ &\leqslant 2\pi\sum_{|\ell|\geqslant k}|\ell|^2(|\ell|-1)^2\cdots(|\ell|-k+1)^2|\widehat{f}_\ell|^2\int_0^1 r^{2|\ell|-2k+1}dr\\ &\quad + 2\pi(1+2+\cdots+k)\|f\|^2_{H^{k-\frac12}(\mathbb{S}^1)}\end{aligned}$$

which, with the help of (2.89), implies that

$$\Big\|\frac{1}{r^k}\partial_\theta^k u\Big\|^2_{L^2(\mathrm{D})} \leqslant \pi(k+1)^2\|f\|^2_{H^{k-\frac12}(\mathbb{S}^1)}\,.$$

As a consequence, we conclude that $\|u\|_{H^k(\mathrm{D})} \leqslant C_k\|f\|_{H^{k-\frac12}(\mathbb{S}^1)}$ 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 2.150, 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:

$f \in H^{3/2}(\mathbb{S}^1)$ implies that $f \in H^{1/2+\alpha}(\mathbb{S}^1)$ for $\alpha \in (0,1)$
which further implies that $f \in \mathscr{C}^{0,\alpha}(\mathbb{S}^1)$,

where the last assertion means that $|f(x+y) - f(x)| \leqslant C|y|^\alpha$.

We start with the identity

$$\begin{aligned} |f(x+y) - f(y)| &= \Big| \sum_{k\in\mathbb{Z}} \widehat{f}_k e^{ikx}(e^{iky} - 1)\Big| = \Big| \sum_{k\neq 0} \widehat{f}_k e^{ikx}(e^{iky} - 1)\Big| \\ &\leqslant \Big(\sum_{k\neq 0} |\widehat{f}_k|^2 \langle k\rangle^{1+2\alpha}\Big)^{\frac{1}{2}} \Big(\sum_{k\neq 0} |e^{iky} - 1|^2 \langle k\rangle^{-1-2\alpha}\Big)^{\frac{1}{2}} \\ &= \|f\|_{H^{1/2+\alpha}(\mathbb{S}^1)} \Big(\sum_{k\neq 0} |e^{iky} - 1|^2 \langle k\rangle^{-1-2\alpha}\Big)^{\frac{1}{2}} . \end{aligned}$$

We consider $|y| \leqslant \frac{1}{2}$ 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 2.150, 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 $\alpha \in (0,1)$.

2.5.3 Exercises

Problem 2.20. 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 $\dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} p_r(\theta)d\theta = 1$.

PROBLEM 2.21. 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 2.22. Let $\mathrm{D} := B(0,1) \subseteq \mathbb{R}^2$ and let u satisfy the Neumann problem

$$\Delta u = 0 \qquad \text{in} \quad \mathrm{D}\,, \tag{2.90a}$$

$$\frac{\partial u}{\partial r} = g \qquad \text{on} \quad \partial \mathrm{D} := \mathbb{S}^1 \,. \tag{2.90b}$$

If $u = \mathrm{PI}(f) := \sum_{k\in\mathbb{Z}} \widehat{f}_k r^{|k|} e^{ik\theta}$, show that for $f \in H^{3/2}(\mathbb{S}^1)$,

$$g = Nf, \tag{2.91}$$

which is the same as

$$\widehat{g}_k = |k| \widehat{f}_k \,.$$

N denotes the Dirichlet to Neumann map given by $Nf(\theta) = \sum_{k\in\mathbb{Z}} \widehat{f}_k |k| e^{ik\theta}$ or $Nf = -i\frac{\partial}{\partial\theta} Hf = -iH\frac{\partial f}{\partial \theta}$, 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 2.23. 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 (2.91) is given by $f(\theta) = (2\pi)^{-1} \int_{\mathbb{S}^1} K(\theta - \varphi) g(\varphi) d\varphi$.

PROBLEM 2.24. Consider the solution to the Neumann problem (2.90a) and (2.90b). Show that $g \in H^{1/2}(\mathbb{S}^1)$ implies that $u \in H^2(\mathrm{D})$ and that

$$\|u\|^2_{H^2(\mathrm{D})} \leqslant C\big(\|g\|^2_{H^{1/2}(\mathbb{S}^1)} + \|u\|^2_{L^2(\mathrm{D})}\big) \,.$$

2.6 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

$$\Delta u = 0 \qquad \text{in} \quad \Omega\,, \tag{2.92a}$$

$$u = f \qquad \text{on} \quad \partial\Omega \,. \tag{2.92b}$$

Theorem 2.151. *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 (2.92) 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 (2.92) to a problem with homogeneous boundary conditions. Using the surjectivity of the trace operator provided by Theorem 2.118, 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 2.45 that $U \in H^1_0(\Omega)$ and satisfies $-\Delta U = \Delta F$ in $H^1_0(\Omega)$; that is

$$\langle -\Delta U, v\rangle = \langle \Delta F, v\rangle \qquad \forall\, v \in H^1_0(\Omega).$$

According to Remark 2.62, $-\Delta : H^1_0(\Omega) \to H^{-1}(\Omega)$ is an isomorphism, so that $\Delta F \in H^{-1}(\Omega)$; therefore, by Theorem 2.61, there exists a unique weak solution $U \in H^1_0(\Omega)$, satisfying

$$\int_\Omega DU \cdot Dv\, dx = \langle \Delta F, v\rangle \qquad \forall\, v \in H^1_0(\Omega),$$

with

$$\|U\|_{H^1(\Omega)} \leqslant C\|\Delta F\|_{H^{-1}(\Omega)}, \tag{2.93}$$

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 2.118, 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,

$$\int_\Omega DU \cdot Dv\, dx = \int_\Omega \Delta F\, v\, dx \qquad \forall\, v \in H^1_0(\Omega). \tag{2.94}$$

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 (2.94); thus,

$$-\int_\Omega U,_i \; \eta_\epsilon * (\zeta^2 U^\epsilon,_j),_{ji} \; dx = -\int_\Omega U,_i \; \eta_\epsilon * [\zeta^2 U^\epsilon,_{ij} + 2\zeta\zeta,_i U^\epsilon,_j],_j \; dx$$
$$= \int_{\Omega_2} \zeta^2 U^\epsilon,_{ij} U^\epsilon,_{ij} \; dx - 2\int_\Omega \eta_\epsilon * [\zeta\zeta,_i U^\epsilon,_j],_j \; U,_i \; dx \,,$$

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 1.45),

$$\|\eta_\epsilon * [\zeta^2 U^\epsilon,_{jj} + 2\zeta\zeta,_j U^\epsilon,_j]\|_{L^2(\Omega_2)} \leqslant \|\zeta^2 U^\epsilon,_{jj} + 2\zeta\zeta,_j U^\epsilon,_j\|_{L^2(\Omega_2)};$$

hence, by the Cauchy-Young inequality with δ, Lemma 1.44, 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

$$\|D^2 U^\epsilon\|^2_{L^2(\Omega_1)} \leqslant \|\zeta D^2 U^\epsilon\|^2_{L^2(\Omega_2)} \leqslant C_\delta \big[\|DU^\epsilon\|^2_{L^2(\Omega_2)} + \|\Delta F\|^2_{L^2(\Omega)}\big] \leqslant C_\delta \|\Delta F\|^2_{L^2(\Omega)} \,,$$

the last inequality following from (2.93), and Young's inequality.

Since the right-hand side does not depend on $\epsilon > 0$, there exists a subsequence

$$D^2 U^{\epsilon'} \rightharpoonup \mathcal{W} \quad \text{in} \quad L^2(\Omega_1) \,.$$

By Theorem 2.21, $U^\epsilon \to U$ in $H^1(\Omega_1)$, so that $\mathcal{W} = D^2 U$ on Ω_1. As weak convergence is lower semi-continuous, $\|D^2 U\|_{L^2(\Omega_1)} \leqslant C_\epsilon \|\Delta F\|_{L^2(\Omega)}$. As Ω_1 and Ω_2 are arbitrary, we have established that $U \in H^2_{\rm loc}(\Omega)$ and that

$$\|U\|_{H^2_{\rm loc}(\Omega)} \leqslant C \|\Delta F\|_{L^2(\Omega)} \,.$$

For any $w \in H^1_0(\Omega)$, set $v = \zeta w$ in (2.94). Since $u \in H^2_{\rm loc}(\Omega)$, 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 $\operatorname{spt}(\zeta)$ can be chosen arbitrarily close to $\partial\Omega$, it follows that for all x in the interior of Ω, we have that

$$-\Delta U(x) = \Delta F(x) \quad \text{for almost every } x \in \Omega\,. \tag{2.95}$$

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

$$\begin{aligned}
&\vartheta_\ell : B(0,r_\ell) \to \mathcal{U}_\ell \text{ is a } \mathscr{C}^\infty \text{ diffeomorphism }\,,\\
&\det(D\vartheta_\ell) = 1\,,\\
&\vartheta_\ell(B(0,r_\ell) \cap \{x_{\mathrm{n}} = 0\}) \to \mathcal{U}_\ell \cap \partial\Omega\,,\\
&\vartheta_\ell(B(0,r_\ell) \cap \{x_{\mathrm{n}} > 0\}) \to \mathcal{U}_\ell \cap \Omega\,.
\end{aligned}$$

Let $0 \leqslant \zeta_\ell \leqslant 1$ in $\mathscr{C}_c^\infty(\mathcal{U}_\ell)$ denote a partition of unity subordinate to the open covering $\mathcal{U}_\ell$, and define the horizontal convolution operator, smoothing functions defined on $\mathbb{R}^{\mathrm{n}}$ in the first $1, ..., \mathrm{n}-1$ directions, as follows:

$$\rho_\epsilon *_h F(x_h, x_{\mathrm{n}}) = \int_{\mathbb{R}^{\mathrm{n}-1}} \rho_\epsilon(x_h - y_h) F(y_h, x_{\mathrm{n}}) dy_h\,,$$

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

$$v = -\big(\rho_\epsilon *_h [(\zeta_\ell \circ \vartheta_\ell)^2 \rho_\epsilon *_h (U \circ \vartheta_\ell),_\alpha],_\alpha\big) \circ \vartheta_\ell^{-1} \in H_0^1(\Omega)$$

into (2.94), and use the change of variables formula to obtain the identity

$$\int_{B_+(0,r_\ell)} A_i^k (U \circ \vartheta_\ell),_k A_i^j (v \circ \vartheta_\ell),_j \, dx = \int_{B_+(0,r_\ell)} (\Delta F) \circ \vartheta_\ell \, v \circ \vartheta_\ell \, dx\,, \tag{2.96}$$

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 (2.96) 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

$$\begin{aligned} \mathcal{I}_1 &= -\int_{B_+(0,r_\ell)} A_i^j A_i^k U^\ell,_k \; H_\epsilon(\xi_\ell^2 H_\epsilon U^\ell,_{j\alpha}),_\alpha \, dx \,, \\ \mathcal{I}_2 &= -2\int_{B_+(0,r_\ell)} A_i^j A_i^k U^\ell,_k \; H_\epsilon(\xi_\ell \xi_\ell,_j H_\epsilon U^\ell,_\alpha),_\alpha \, dx \,. \end{aligned}$$

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

$$\begin{aligned} \mathcal{I}_{1a} &= \int_{B_+(0,r_\ell)} (A_i^j A_i^k H_\epsilon U^\ell,_k),_\alpha \; \xi_\ell^2 H_\epsilon U^\ell,_{j\alpha} \, dx \,, \\ \mathcal{I}_{1b} &= \int_{B_+(0,r_\ell)} \big(\llbracket H_\epsilon, A_i^j A_i^k \rrbracket U^\ell,_k\big),_\alpha \; \xi_\ell^2 H_\epsilon U^\ell,_{j\alpha} \, dx \,, \end{aligned}$$

and where

$$\llbracket H_\epsilon, A_i^j A_i^k \rrbracket U^\ell,_k = H_\epsilon(A_i^j A_i^k U^\ell,_k) - A_i^j A_i^k \, H_\epsilon U^\ell,_k \tag{2.97}$$

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 (2.96), we see that

$$\begin{aligned} \int_{B_+(0,r_\ell)} \xi_\ell^2 A_i^j A_i^k H_\epsilon U^\ell,_{k\alpha} \; H_\epsilon U^\ell,_{j\alpha} \, dx &\leqslant \Big| \int_{B_+(0,r_\ell)} (A_i^j A_i^k),_\alpha H_\epsilon U^\ell,_k \; \xi_\ell^2 H_\epsilon U^\ell,_{j\alpha} \,] \, dx \Big| \\ &\quad + |\mathcal{I}_{1b}| + |\mathcal{I}_2| + \Big| \int_{B_+(0,r_\ell)} (\Delta F) \circ \vartheta_\ell \, v \circ \vartheta_\ell \, dx \Big| \,. \end{aligned}$$

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

$$\lambda |Y|^2 \leqslant A^j_i A^k_i Y_j Y_k \quad \forall\, Y \in \mathbb{R}^{\mathrm{n}}\,.$$

It follows that

$$\lambda \int_{B_+(0,r_\ell)} \xi_\ell^2 |\bar\partial D H_\epsilon U^\ell|^2\, dx \leqslant \Big| \int_{B_+(0,r_\ell)} (A^j_i A^k_i),_\alpha\, H_\epsilon U^\ell,_k\ \xi_\ell^2 H_\epsilon U^\ell,_{j\alpha}\,]\, dx \Big| \\ + |\mathcal{I}_{1b}| + |\mathcal{I}_2| + \Big| \int_{B_+(0,r_\ell)} (\Delta F)\circ\vartheta_\ell\, v\circ\vartheta_\ell\, dx \Big|\,,$$

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

$$\Big| \int_{B_+(0,r_\ell)} (A^j_i A^k_i),_\alpha\, H_\epsilon U^\ell,_k\ \xi_\ell^2 H_\epsilon U^\ell,_{j\alpha}\,]\, dx \Big| + |\mathcal{I}_2| + \Big| \int_{B_+(0,r_\ell)} (\Delta F)\circ\vartheta_\ell\, v\circ\vartheta_\ell\, dx \Big| \\ \leqslant \delta \int_{B_+(0,r_\ell)} \xi_\ell^2 |\bar\partial D H_\epsilon U^\ell|^2\, dx + C_\delta \|\Delta F\|^2_{L^2(\Omega)}\,.$$

It remains to establish such an upper bound for $|\mathcal{I}_{1b}|$.

To do so, we first establish a pointwise bound for (2.97): for $\mathcal{A}^{jk} = A^j_i A^k_i$,

$$[\![H_\epsilon, A^j_i A^k_i]\!] U^\ell,_k(x) \\ = \int_{B(x_h,\epsilon)} \rho_\epsilon(x_h - y_h) \big[\mathcal{A}^{jk}(y_h, x_{\mathrm{n}}) - \mathcal{A}^{jk}(x_h, x_{\mathrm{n}})\big] U^\ell,_k(y_h, x_{\mathrm{n}})\, dy_h\,.$$

By Morrey's inequality, $\big|[\mathcal{A}^{jk}(y_h, x_{\mathrm{n}}) - \mathcal{A}^{jk}(x_h, x_{\mathrm{n}})]\big| \leqslant C\epsilon \|\mathcal{A}\|_{W^{1,\infty}(B_+(0,r_\ell))}$. 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

$$\big| \partial_{x_\alpha} \big([\![H_\epsilon, A^j_i A^k_i]\!] U^\ell,_k \big)(x) \big| \leqslant C \int_{B(x_h,\epsilon)} \frac{1}{\epsilon} \rho' \Big(\frac{x - h - y_h}{\epsilon} \Big) \big| U^\ell,_k(y_h, x_{\mathrm{n}}) \big|\, dy_h$$

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

$$\int_{B_+(0,r_\ell)} \xi_\ell^2 |\bar\partial D U^\ell|^2 \, dx \leqslant C_\delta \|\Delta F\|^2_{L^2(\Omega)} \,. \tag{2.98}$$

From (2.95), 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,

$$-\mathcal{A}^{\mathrm{nn}} U^\ell,_{\mathrm{nn}} = \mathcal{A}^{j\alpha} U^\ell,_{\alpha j} + \mathcal{A}^{\beta k} U^\ell,_{k\beta} + \mathcal{A}^{jk},_j U^\ell,_k + \Delta F \circ \vartheta_\ell \,. \tag{2.99}$$

Since $\mathcal{A}^{\mathrm{nn}} > 0$, it follows from (2.98) that

$$\int_{B_+(0,r_\ell)} \xi_\ell^2 |D^2 U^\ell|^2 \, dx \leqslant C_\delta \|\Delta F\|^2_{L^2(\Omega)} \,. \tag{2.100}$$

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 (2.99). $\square$

Appendix A

Elementary Analysis

A.1 The Inverse Function Theorem

The main goal of this section is to introduce the global inverse function theorem which will be used to determine when a local diffeomorphism is a global diffeomorphism.

Theorem A.1 (Global Existence of 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

$$\begin{aligned} &f : \mathcal{U}_1 \cap \mathcal{O} \cap K \to \mathcal{V}_1 \cap \mathcal{V}_2 \cap f(K) \text{ is one-to-one and onto}\,,\\ &f : \mathcal{U}_2 \cap \mathcal{O} \cap K \to \mathcal{V}_1 \cap \mathcal{V}_2 \cap f(K) \text{ is one-to-one and onto}\,. \end{aligned}$$

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, we must have $E = K$ 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 $f(\partial K) \subseteq f\big(\text{int}(K)\big)$.

Since the compactness of K implies that $f(K)$ is compact, there is $b \in \mathbb{R}^n$ such that $b \notin f(K)$. Consider the function $\varphi : K \to \mathbb{R}$ defined by

$$\varphi(x) = \frac{1}{2}\|f(x) - b\|^2_{\mathbb{R}^n} = \frac{1}{2}\sum_{j=1}^n |f_i(x) - b_i|^2\,.$$

Then φ is a continuous function on K; thus φ attains its maximum at $x_0 \in K$. Since $f(\partial K) \subseteq f\big(\text{int}(K)\big)$, we can assume that $x_0 \in \text{int}(K)$; 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 conclude that $(Df)(x_0)$ is not invertible, a contradiction. $\square$

Appendix B

Preliminaries

B.1 Linear Algebra

B.1.1 Vector spaces

Definition B.1 (Vector spaces)**.** A vector space $\mathcal{V}$ over a scalar field $\mathbb{F}$ is a set of elements called *vectors*, together with two operations $+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ and $\cdot : \mathbb{F} \times \mathcal{V} \to \mathcal{V}$, called the *vector addition* and *scalar multiplication* respectively, such that

1. $\boldsymbol{v} + \boldsymbol{w} = \boldsymbol{w} + \boldsymbol{v}$ for all $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$.
2. $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$ for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$.
3. There is a zero vector $\mathbf{0}$ such that $\boldsymbol{v} + \mathbf{0} = \boldsymbol{v}$ for all $\boldsymbol{v} \in \mathcal{V}$.
4. For every $\boldsymbol{v}$ in $\mathcal{V}$, there is a vector $\boldsymbol{w}$ such that $\boldsymbol{v} + \boldsymbol{w} = \mathbf{0}$.
5. $\alpha \cdot (\boldsymbol{v} + \boldsymbol{w}) = \alpha \cdot \boldsymbol{v} + \alpha \cdot \boldsymbol{w}$ for all $\alpha \in \mathbb{F}$ and $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$.
6. $\alpha \cdot (\beta \cdot \boldsymbol{v}) = (\alpha\beta) \cdot \boldsymbol{v}$ for all $\alpha, \beta \in \mathbb{F}$ and $\boldsymbol{v} \in \mathcal{V}$.
7. $(\alpha + \beta) \cdot \boldsymbol{v} = \alpha \cdot \boldsymbol{v} + \beta \cdot \boldsymbol{v}$ for all $\alpha, \beta \in \mathbb{F}$ and $\boldsymbol{v} \in \mathcal{V}$.
8. $1 \cdot \boldsymbol{v} = \boldsymbol{v}$ for all $\boldsymbol{v} \in \mathcal{V}$.

For notational convenience, we often drop the $\cdot$ and write $\alpha\boldsymbol{v}$ instead of $\alpha \cdot \boldsymbol{v}$.

Remark B.2. In property 4 of the definition above, it is easy to see that for each $\boldsymbol{v}$, there is only one vector $\boldsymbol{w}$ such that $\boldsymbol{v} + \boldsymbol{w} = \mathbf{0}$. We often denote this $\boldsymbol{w}$ by

$-\boldsymbol{v}$, and the vector substraction $-:\mathcal{V}\times\mathcal{V}\to\mathcal{V}$ is then defined (or understood) as $\boldsymbol{v}-\boldsymbol{w}=\boldsymbol{v}+(-\boldsymbol{w})$.

EXAMPLE B.3. Let $\mathbb{F}$ be a scalar field. The space $\mathbb{F}^{\rm n}$ is the collection of n-tuple $\boldsymbol{v}=({\rm v}_1,{\rm v}_2,\cdots,{\rm v}_{\rm n})$ with ${\rm v}_i\in\mathbb{F}$ with addition $+$ and scalar multiplication $\cdot$ defined by

$$({\rm v}_1,\cdots,{\rm v}_{\rm n})+({\rm w}_1,\cdots,{\rm w}_{\rm n})\equiv({\rm v}_1+{\rm w}_1,\cdots,{\rm v}_{\rm n}+{\rm w}_{\rm n})\,,$$
$$\alpha({\rm v}_1,\cdots,{\rm v}_{\rm n})\equiv(\alpha{\rm v}_1,\cdots,\alpha{\rm v}_{\rm n})\,.$$

Then $\mathbb{F}^{\rm n}$ is a vector space.

EXAMPLE B.4. Let $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$, and $\mathcal{V}$ be the collection of all $\mathbb{R}$-valued continuous functions on $[0,1]$. The vector addition $+$ and scalar multiplication $\cdot$ is defined by

$$\begin{aligned}(f+g)(x)&=f(x)+g(x) &&\forall\, f,g\in\mathcal{V}\,,\\ (\alpha\cdot f)(x)&=\alpha f(x) &&\forall\, f\in\mathcal{V},\alpha\in\mathbb{F}\,.\end{aligned}$$

Then $\mathcal{V}$ is a vector space, and is denoted by $\mathscr{C}([0,1];\mathbb{F})$. When the scalar field under consideration is clear, we simply use $\mathscr{C}([0,1])$ to denote this vector space.

The linear independence of vectors

DEFINITION B.5. Let $\mathcal{V}$ be a vector space over a scalar field $\mathbb{F}$. k vectors $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\cdots$, $\boldsymbol{v}_k$ in $\mathcal{V}$ is said to be ***linearly dependent*** if there exists $(\alpha_1,\cdots,\alpha_k)\subseteq\mathbb{F}^k$, $(\alpha_1,\cdots,\alpha_k)\neq\mathbf{0}$ such that $\alpha_1\boldsymbol{v}_1+\alpha_2\boldsymbol{v}_2+\cdots+\alpha_k\boldsymbol{v}_k=\mathbf{0}$. k vectors $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\cdots$, $\boldsymbol{v}_k$ in $\mathcal{V}$ is said to be ***linearly independent*** if they are not linearly dependent. In other words, $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_k\}$ are linearly independent if

$$\alpha_1\boldsymbol{v}_1+\alpha_2\boldsymbol{v}_2+\cdots+\alpha_k\boldsymbol{v}_k=\mathbf{0}\quad\Rightarrow\quad\alpha_1=\alpha_2=\cdots=\alpha_k=0\,.$$

EXAMPLE B.6. The k vectors $\{1,x,x^2,\cdots,x^{k-1}\}$ are linearly independent in $\mathscr{C}([0,1])$ for all $k\in\mathbb{N}$.

The dimension of a vector space

DEFINITION B.7. The ***dimension*** of a vector space $\mathcal{V}$ is the number of maximum linearly independent set in $\mathcal{V}$, and in such case $\mathcal{V}$ is called an n-dimensional vector space, where n the the dimension of $\mathcal{V}$. If for every number ${\rm n}\in\mathbb{N}$ there exists n linearly independent vectors in $\mathcal{V}$, the vector space $\mathcal{V}$ is said to be infinitely dimensional.

Example B.8. The space $\mathbb{F}^{\mathrm{n}}$ is n-dimensional, and $\mathscr{C}([0,1])$ is infinitely dimensional (since $1, x, \cdots, x^{\mathrm{n}-1}$ are n linearly independent vectors in $\mathscr{C}([0,1])$).

Bases of a vector space

Definition B.9 (Basis)**.** Let $\mathcal{V}$ be a vector space over $\mathbb{F}$. A set of vectors $\{\boldsymbol{v}_i\}_{i\in\mathcal{I}}$ in $\mathcal{V}$ is called a ***basis*** of $\mathcal{V}$ if for every $\boldsymbol{v} \in \mathcal{V}$, there exists a unique $\{\alpha_i\}_{i\in\mathcal{I}} \subseteq \mathbb{F}$ such that

$$\boldsymbol{v} = \sum_{\alpha\in\mathcal{I}} \alpha_i \boldsymbol{v}_i \,.$$

For a given basis $\mathcal{B} = \{\boldsymbol{v}_i\}_{i\in\mathcal{I}}$, the coefficients $\{\alpha_i\}_{i\in\mathcal{I}}$ given in the above relation is denoted by $[\boldsymbol{v}]_{\mathcal{B}}$.

Example B.10. Let $\mathrm{e}_i = (0, , \cdots, 0, 1, 0, \cdots, 0)$, where 1 locates at the i-th slot. Then the collection $\{\mathrm{e}_i\}_{i=1}^{\mathrm{n}}$ is a basis of $\mathbb{F}^{\mathrm{n}}$ since

$$(\alpha_1, \cdots, \alpha_{\mathrm{n}}) = \sum_{i=1}^{\mathrm{n}} \alpha_i \mathrm{e}_i \qquad \forall\, \alpha_i \in \mathbb{F}.$$

The collection $\{\mathrm{e}_i\}_{i=1}^{\mathrm{n}}$ is called the standard basis of $\mathbb{F}^{\mathrm{n}}$.

Example B.11. Even though $\{1, x, \cdots, x^k, \cdots\}$ is a set of linearly independent vectors, it is not a basis of $\mathscr{C}([0,1])$. However, let $\mathscr{P}([0,1])$ be the collection of polynomials defined on $[0,1]$. Then $\mathscr{P}([0,1])$ is still a vector space, and $\{1, x, \cdots, x^k, \cdots\}$ is a basis of $\mathscr{P}([0,1])$.

B.1.2 Inner products and inner product spaces

Definition B.12 (Inner product space)**.** Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. A vector space $\mathcal{V}$ over a scalar field $\mathbb{F}$ with a bilinear form $(\cdot,\cdot) : \mathcal{V} \times \mathcal{V} \to \mathbb{F}$ is called an inner product space if the bilinear form satisfies

1. $(\boldsymbol{v}, \boldsymbol{v}) \geqslant 0$ for all $\boldsymbol{v} \in \mathcal{V}$.
2. $(\boldsymbol{v}, \boldsymbol{v}) = 0$ if and only if $\boldsymbol{v} = 0$.
3. $(\boldsymbol{v}, \boldsymbol{w}) = \overline{(\boldsymbol{w}, \boldsymbol{v})}$ for all $\boldsymbol{v}, w \in \mathcal{V}$, where the bar over the scalar $(\boldsymbol{w}, \boldsymbol{v})$ is the complex conjugate.

4. $(\boldsymbol{v}+\boldsymbol{w},\boldsymbol{u}) = (\boldsymbol{v},\boldsymbol{u}) + (\boldsymbol{w},\boldsymbol{u})$ for all $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in \mathcal{V}$.

5. $(\alpha\boldsymbol{v},\boldsymbol{w}) = \alpha(\boldsymbol{v},\boldsymbol{w})$ for all $\alpha \in \mathbb{F}$ and $\boldsymbol{v},\boldsymbol{w} \in \mathcal{V}$.

The bilinear form $(\cdot,\cdot)$ is called an *inner product* on $\mathcal{V}$.

Example B.13. Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, and $\mathbb{F}^{\mathrm{n}}$ be the vector space defined in Example B.3. We may define an inner product on $\mathbb{F}^{\mathrm{n}}$ by

$$(\boldsymbol{v},\boldsymbol{w}) \equiv \sum_{i=1}^{\mathrm{n}} \mathrm{v}_i \overline{\mathrm{w}}_i,$$

where v_i and w_i are the i-th component of $\boldsymbol{v}$ and $\boldsymbol{w}$, respectively, and $\overline{\mathrm{w}}_i$ is the complex conjugate of w_i. We sometimes use $\boldsymbol{v}\cdot\boldsymbol{w}$ to denote $(\boldsymbol{v},\boldsymbol{w})$.

Example B.14. Let $\mathcal{V} = \mathscr{C}([0,1];\mathbb{R})$. Define

$$(f,g) = \int_0^1 f(x)g(x)dx\,.$$

Then $\big(\mathscr{C}([0,1];\mathbb{R}),(\cdot,\cdot)\big)$ is an inner product space. The norm induced by this inner product is given by

$$\|f\| = \Big[\int_0^1 |f(x)|^2 dx\Big]^{\frac{1}{2}},$$

and is called the L^2-norm.

Proposition B.15. *Let $\mathcal{V}$ be an inner product space with inner product $(\cdot,\cdot)$. The inner product $(\cdot,\cdot)$ on $\mathcal{V}$ induces a norm defined by*

$$\|\boldsymbol{v}\| \equiv \sqrt{(\boldsymbol{v},\boldsymbol{v})}$$

satisfying

1. $\|\boldsymbol{v}\| \geqslant 0$ *for all* $\boldsymbol{v} \in \mathcal{V}$.

2. $\|\boldsymbol{v}\| = 0$ *if and only if* $\boldsymbol{v} = 0$.

3. $\|\alpha\boldsymbol{v}\| = |\alpha|\|\boldsymbol{v}\|$ *for all* $\alpha \in \mathbb{F}$ *and* $\boldsymbol{v} \in \mathcal{V}$.

4. $\|\boldsymbol{v}+\boldsymbol{w}\| \leqslant \|\boldsymbol{v}\| + \|\boldsymbol{w}\|$ *for all* $\boldsymbol{v},\boldsymbol{w} \in \mathcal{V}$.

5. $|(\boldsymbol{v},\boldsymbol{w})| \leqslant \|\boldsymbol{v}\|\|\boldsymbol{w}\|$ *for all* $\boldsymbol{v},\boldsymbol{w} \in \mathcal{V}$.

Proof. Properties 1 through 3 are obvious. We focus on proving property 5 first, and as we will see, property 4 is a direct consequence of property 5.

Let $\alpha \in \mathbb{F}$ satisfy $\alpha(\boldsymbol{v}, \boldsymbol{w}) = |(\boldsymbol{v}, \boldsymbol{w})|$. Then $|\alpha| = 1$. For all $\lambda \in \mathbb{R}$,

$$\begin{aligned}(\lambda\alpha\boldsymbol{v} + \boldsymbol{w}, \lambda\alpha\boldsymbol{v} + \boldsymbol{w}) &= (\lambda\alpha\boldsymbol{v}, \lambda\alpha\boldsymbol{v}) + (\lambda\alpha\boldsymbol{v}, \boldsymbol{w}) + (\boldsymbol{w}, \lambda\alpha\boldsymbol{v}) + (\boldsymbol{w}, \boldsymbol{w}) \\ &= \lambda^2\|\boldsymbol{v}\|^2 + \lambda\alpha(\boldsymbol{v}, \boldsymbol{w}) + \overline{\lambda\alpha(\boldsymbol{v}, \boldsymbol{w})} + \|\boldsymbol{w}\|^2 \\ &= \lambda^2\|\boldsymbol{v}\|^2 + 2\lambda|(\boldsymbol{v}, \boldsymbol{w})| + \|\boldsymbol{w}\|^2 .\end{aligned}$$

Since the left-hand side of the quantity above is always non-negative for all $\lambda \in \mathbb{R}$, we must have

$$|(\boldsymbol{v}, \boldsymbol{w})|^2 - \|\boldsymbol{v}\|^2\|\boldsymbol{w}\|^2 \leqslant 0$$

which implies property 5. To prove property 4, we note that

$$\begin{aligned}\|\boldsymbol{v} + \boldsymbol{w}\| \leqslant \|\boldsymbol{v}\| + \|\boldsymbol{w}\| &\Leftrightarrow \|\boldsymbol{v} + \boldsymbol{w}\|^2 \leqslant (\|\boldsymbol{v}\| + \|\boldsymbol{w}\|)^2 \\ &\Leftrightarrow (\boldsymbol{v} + \boldsymbol{w}, v + \boldsymbol{w}) \leqslant \|\boldsymbol{v}\|^2 + 2\|\boldsymbol{v}\|\|\boldsymbol{w}\| + \|\boldsymbol{w}\|^2 \\ &\Leftrightarrow \mathrm{Re}(\boldsymbol{v}, \boldsymbol{w}) \leqslant \|\boldsymbol{v}\|\|\boldsymbol{w}\|\end{aligned}$$

while the last inequality is valid because of property 5. ◻

Remark B.16. The inequality in property 5 is called the ***Cauchy-Schwarz inequality***.

B.1.3 Normed vector spaces

The norm introduced in Proposition B.15 is a good way of measure the magnitude of vectors. In general if a real-valued function can be used as a measurement of the magnitude of vectors if certain properties are satisfied.

Definition B.17. Let $\mathcal{V}$ be a vector space over scalar field $\mathbb{F}$. A real-valued function $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ is said to be a norm of $\mathcal{V}$ if

1. $\|\boldsymbol{v}\| \geqslant 0$ for all $\boldsymbol{v} \in \mathcal{V}$.

2. $\|\boldsymbol{v}\| = 0$ if and only if $\boldsymbol{v} = 0$.

3. $\|\alpha\boldsymbol{v}\| = |\alpha|\|\boldsymbol{v}\|$ for all $\boldsymbol{v} \in \mathcal{V}$ and $\alpha \in \mathbb{F}$.

4. $\|\boldsymbol{v} + \boldsymbol{w}\| \leqslant \|\boldsymbol{v}\| + \|\boldsymbol{w}\|$ for all $\boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$.

The pair $(\mathcal{V}, \|\cdot\|)$ is called a normed vector space.

Example B.18. Let $\mathcal{V} = \mathbb{F}^{\mathrm{n}}$, and $\|\cdot\|_p$ be defined by

$$\|x\|_p = \begin{cases} \Big[\sum_{i=1}^{\mathrm{n}} |x_i|^p\Big]^{\frac{1}{p}} & \text{if} \quad 1 \leqslant p < \infty\,, \\ \max_{1 \leqslant i \leqslant \mathrm{n}} |x_i| & \text{if} \quad p = \infty\,, \end{cases}$$

where $x = (x_1, \cdots, x_{\mathrm{n}})$. The function $\|\cdot\|_p$ is a norm of $\mathbb{F}^{\mathrm{n}}$, and is called the p-norm of $\mathbb{F}^{\mathrm{n}}$.

Example B.19. For each $f \in \mathscr{C}([0,1];\mathbb{R})$, we define

$$\|f\|_p = \begin{cases} \Big[\int_0^1 |f(x)|^p dx\Big]^{\frac{1}{p}} & \text{if } 1 \leqslant p < \infty\,, \\ \max_{x \in [0,1]} |f(x)| & \text{if } p = \infty\,. \end{cases}$$

The function $\|\cdot\|_p : \mathscr{C}([0,1];\mathbb{R}) \to \mathbb{R}$ is a norm on $\mathscr{C}([0,1];\mathbb{R})$ (Minkowski's inequality), and is called the L^p-norm.

B.1.4 Bounded linear maps

Definition B.20 (Linear map)**.** Let $\mathcal{V}$ and $\mathcal{W}$ be two vector spaces over a scalar field $\mathbb{F}$. A map $L : \mathcal{V} \to \mathcal{W}$ is called a ***linear map*** from $\mathcal{V}$ into $\mathcal{W}$ if

$$L(\alpha \boldsymbol{v} + \boldsymbol{w}) = \alpha L(\boldsymbol{v}) + L(\boldsymbol{w}) \qquad \forall\, \alpha \in \mathbb{F} \text{ and } \boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}.$$

For notational convenience, we often write $L\boldsymbol{v}$ instead of $L(\boldsymbol{v})$. When $\mathcal{V}$ and $\mathcal{W}$ are finite dimensional, linear maps (from $\mathcal{V}$ into $\mathcal{W}$) are sometimes called ***linear transformations*** (from $\mathcal{V}$ into $\mathcal{W}$).

Let $L_1, L_2 : \mathcal{V} \to \mathcal{W}$ be two linear maps, and $\alpha \in \mathbb{F}$ be a scalar. It is easy to see that $\alpha L_1 + L_2 : \mathcal{V} \to \mathcal{W}$ is also a linear map. This is equivalent to say that the collection of linear maps is a vector space, and this induces the following

Definition B.21. The vector space $\mathscr{L}(\mathcal{V}, \mathcal{W})$ is the collection of linear maps from $\mathcal{V}$ to $\mathcal{W}$.

Definition B.22 (Boundedness of linear maps)**.** Let $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ and $(\mathcal{W}, \|\cdot\|_{\mathcal{W}})$ be two normed vector spaces over a scalar field $\mathbb{F}$. A linear map $L : \mathcal{V} \to \mathcal{W}$ is said to be bounded if the number

$$\|L\|_{\mathscr{B}(\mathcal{V},\mathcal{W})} \equiv \sup_{\|\boldsymbol{v}\|_{\mathcal{V}}=1} \|L\boldsymbol{v}\|_{\mathcal{W}} = \sup_{\boldsymbol{v}\neq 0} \frac{\|L\boldsymbol{v}\|_{\mathcal{W}}}{\|\boldsymbol{v}\|_{\mathcal{V}}} \tag{B.1}$$

is finite. The collection of all bounded linear map from $\mathcal{V}$ to $\mathcal{W}$ is denoted by $\mathscr{B}(\mathcal{V}, \mathcal{W})$.

Remark B.23. When the domain $\mathcal{V}$ and the target $\mathcal{W}$ under consideration are clear, we use $\|\cdot\|$ instead of $\|\cdot\|_{\mathscr{B}(\mathcal{V},\mathcal{W})}$ to simplify the notation of operator norm.

Proposition B.24. *Let $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ and $(\mathcal{W}, \|\cdot\|_{\mathcal{W}})$ be two normed vector spaces over a scalar field $\mathbb{F}$. Then $\big(\mathscr{B}(\mathcal{V}, \mathcal{W}), \|\cdot\|\big)$ with $\|\cdot\|$ defined by* (B.1) *is a normed vector space. (Therefore, $\|\cdot\|$ is also called an operator norm).*

Theorem B.25. *Let $\mathrm{GL}(n)$ be the set of all invertible linear maps on $(\mathbb{R}^n, \|\cdot\|_2)$; that is,*

$$\mathrm{GL}(n) = \big\{L \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^n) \,\big|\, L \textit{ is one-to-one (and onto)}\big\}\,.$$

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. *The mapping $L \mapsto L^{-1}$ is continuous on $\mathrm{GL}(n)$; that is,*

$$\forall\, \varepsilon > 0\,, \exists\, \delta > 0 \ni \|K^{-1} - L^{-1}\| < \varepsilon \quad \textit{whenever} \quad \|K - L\| < \delta\,.$$

Proof. 1. Let $\|L^{-1}\| = \dfrac{1}{\alpha}$ and $\|K - L\| = \beta$. Then $\beta < \alpha$; thus for every $x \in \mathbb{R}^n$,

$$\begin{aligned}\alpha\|x\|_{\mathbb{R}^n} = \alpha\|L^{-1}Lx\|_{\mathbb{R}^n} \leqslant \alpha\|L^{-1}\|\|Lx\|_{\mathbb{R}^n} &= \|Lx\|_{\mathbb{R}^n} \leqslant \|(L-K)x\|_{\mathbb{R}^n} + \|Kx\|_{\mathbb{R}^n} \\ &\leqslant \beta\|x\|_{\mathbb{R}^n} + \|Kx\|_{\mathbb{R}^n}\,.\end{aligned}$$

As a consequence, $(\alpha - \beta)\|x\|_{\mathbb{R}^n} \leqslant \|Kx\|_{\mathbb{R}^n}$ and this implies that $K : \mathbb{R}^n \to \mathbb{R}^n$ is one-to-one hence invertible.

2. Let $L \in \mathrm{GL}(n)$ and $\varepsilon > 0$ be given. Choose $\delta = \min\Big\{\dfrac{1}{2\|L^{-1}\|}, \dfrac{\varepsilon}{2\|L^{-1}\|^2}\Big\}$. If $\|K - L\| < \delta$, then $K \in \mathrm{GL}(n)$. Since $L^{-1} - K^{-1} = K^{-1}(K - L)L^{-1}$, we find that if $\|K - L\| < \delta$,

$$\|K^{-1}\| - \|L^{-1}\| \leqslant \|K^{-1} - L^{-1}\| \leqslant \|K^{-1}\|\|K - L\|\|L^{-1}\| < \frac{1}{2}\|K^{-1}\|$$

which implies that $\|K^{-1}\| < 2\|L^{-1}\|$. Therefore, if $\|K - L\| < \delta$,

$$\|L^{-1} - K^{-1}\| \leqslant \|K^{-1}\|\|K - L\|\|L^{-1}\| < 2\|L^{-1}\|^2\delta < \varepsilon\,. \qquad \square$$

Definition B.26. Let $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ and $(\mathcal{W}, \|\cdot\|_{\mathcal{W}})$ be two normed vector spaces over a scalar field $\mathbb{F}$, and $L \in \mathscr{B}(\mathcal{V}, \mathcal{W})$. The collection of all elements $\boldsymbol{v} \in \mathcal{V}$ such that $L\boldsymbol{v} = \mathbf{0}$ is called the kernel (or the null space) of L and is denoted by $\mathrm{Ker}(L)$ or $\mathrm{Null}(L)$. In other words,

$$\mathrm{Ker}(L) = \left\{ \boldsymbol{v} \in \mathcal{V} \,\middle|\, L\boldsymbol{v} = \mathbf{0} \right\}.$$

Theorem B.27 (Riesz Representation Theorem)**.** *Let $(\mathcal{V}, (\cdot,\cdot)_{\mathcal{V}})$ be an inner product space, and $f : \mathcal{V} \to \mathbb{R}$ be a bounded linear map. Then there exists a unique $\boldsymbol{w} \in \mathcal{V}$ such that $f(\boldsymbol{v}) = (\boldsymbol{v}, \boldsymbol{w})_{\mathcal{V}}$ for all $\boldsymbol{v} \in \mathcal{V}$.*

Proof. The uniqueness for such a vector $\boldsymbol{w}$ is simply due to the fact that there is no non-trivial vector which is orthogonal to itself.

Now we show the existence of $\boldsymbol{w}$. If $f(\boldsymbol{v}) = 0$ for all $\boldsymbol{v} \in \mathcal{V}$, then $\boldsymbol{w} = \mathbf{0}$ does the job. Now suppose that $\mathrm{Ker}(f) \subsetneq \mathcal{V}$. Then there exists $\boldsymbol{w} \in \mathrm{Ker}(f)^{\perp}$ such that $f(\boldsymbol{w}) = 1$. We claim that $f(\boldsymbol{v}) = (\boldsymbol{v}, \boldsymbol{w})_{\mathcal{V}}$. □

Remark B.28. For a normed vector space $(X, \|\cdot\|)$, an element in $\mathscr{B}(X, \mathbb{R})$ is usually called a bounded linear functional (defined on X), and the space $\mathscr{B}(X, \mathbb{R})$ is often denoted by X', called the dual space of X.

By the Riesz representation theorem, we conclude the following

Theorem B.29. *Let $(\mathcal{V}, (\cdot,\cdot)_{\mathcal{V}})$ and $(\mathcal{W}, (\cdot,\cdot)_{\mathcal{W}})$ be two inner product spaces. Then for all $L \in \mathscr{B}(\mathcal{V}, \mathcal{W})$, there exists a unique $L^* \in \mathscr{B}(\mathcal{W}, \mathcal{V})$ such that*

$$(L\boldsymbol{v}, \boldsymbol{w})_{\mathcal{W}} = (\boldsymbol{v}, L^*\boldsymbol{w})_{\mathcal{V}} \qquad \forall\, \boldsymbol{v} \in \mathcal{V}, \boldsymbol{w} \in \mathcal{W}.$$

Definition B.30 (Dual map)**.** Let $\mathcal{V}$ and $\mathcal{W}$ be two inner product spaces, and $L : \mathcal{V} \to \mathcal{W}$ be a bounded linear map. The dual map of L, denoted by L^*, is the unique linear map from $\mathcal{W}$ into $\mathcal{V}$ satisfying

$$(L\boldsymbol{v}, \boldsymbol{w})_{\mathcal{W}} = (\boldsymbol{v}, L^*\boldsymbol{w})_{\mathcal{V}} \qquad \forall\, \boldsymbol{v} \in \mathcal{V}, \boldsymbol{w} \in \mathcal{W},$$

where $(\cdot,\cdot)_{\mathcal{V}}$ and $(\cdot,\cdot)_{\mathcal{W}}$ are inner products on $\mathcal{V}$ and $\mathcal{W}$, respectively.

Remark B.31. If $\mathcal{V}$ is finite dimensional, then $\mathscr{L}(\mathcal{V}, \mathcal{W}) = \mathscr{B}(\mathcal{V}, \mathcal{W})$.

Definition B.32 (Dual space)**.**

B.1.5 Matrices

Definition B.33 (Matrix)**.** Let $\mathbb{F}$ be a scalar field. The space $\mathbb{M}(\boldsymbol{m}, \mathrm{n}; \mathbb{F})$ is the collection of elements, called an m-by-n matrix or $\boldsymbol{m} \times \mathrm{n}$ matrix over $\mathbb{F}$, of the form

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

where $a_{ij} \in \mathbb{F}$ is called the (i, j)-th entry of A, and is denoted by $[A]_{ij}$. We write $A = [a_{ij}]_{1 \leqslant i \leqslant m; 1 \leqslant j \leqslant n}$ or simply $A = [a_{ij}]_{\boldsymbol{m} \times \mathrm{n}}$ to denote that A is an $\boldsymbol{m} \times \mathrm{n}$ matrix whose (i, j)-th entry is a_{ij}. The $1 \times \boldsymbol{m}$ matrix

$$a_{i*} = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix}$$

is called the i-th row of A, and the $m \times 1$ matrix

$$a_{*j} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

is called the j-th column of A.

Definition B.34 (Matrix addition)**.** Let $A = [a_{ij}]_{\boldsymbol{m} \times \mathrm{n}}$ and $B = [b_{ij}]_{\boldsymbol{m} \times \mathrm{n}}$ be two $\boldsymbol{m} \times \mathrm{n}$ matrices over a scalar field $\mathbb{F}$. The sum of A and B, denoted by $A + B$, is another $\boldsymbol{m} \times \mathrm{n}$ matrix defined by $A + B = [a_{ij} + b_{ij}]_{\boldsymbol{m} \times \mathrm{n}}$ or more precisely,

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{bmatrix}.$$

Definition B.35 (Scalar multiplication)**.** Let $A = [a_{ij}]_{\boldsymbol{m} \times \mathrm{n}}$ be an $\boldsymbol{m} \times \mathrm{n}$ matrix over a scalar field $\mathbb{F}$, and $\alpha \in \mathbb{F}$. The scalar multiplication of α and A, denoted by αA, is an $\boldsymbol{m} \times \mathrm{n}$ matrix defined by $\alpha A = [\alpha a_{ij}]_{\boldsymbol{m} \times \mathrm{n}}$ or more precisely,

$$\alpha A = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{bmatrix}.$$

Proposition B.36. *The space $\mathbb{M}(m, n; \mathbb{F})$ is a vector space over $\mathbb{F}$ under the matrix addition and scalar multiplication defined in previous two definitions.*

Definition B.37 (Matrix product)**.** Let $A \in \mathbb{M}(m, n; \mathbb{F})$ and $B \in \mathbb{M}(n, \ell; \mathbb{F})$ be two matrices over a scalar field $\mathbb{F}$. The matrix product of A and B, denoted by AB, is an $m \times \ell$ matrix given by $AB = [c_{ij}]_{\boldsymbol{m} \times \mathrm{n}}$ with $c_{ij} = \sum\limits_{k=1}^{\mathrm{n}} a_{ik} b_{kj}$. In other words, the (i, j)-th entry of the product AB is the inner product of the i-th row of A and the j-th column of B.

Remark B.38. The matrix product AB is only defined if the number of columns of A is the same as the number of rows of B. Therefore, even if AB is defined, BA might not make sense. When A and B are both $n \times \mathrm{n}$ square matrix, AB and BA are both defined; however, in general $AB \neq BA$.

Remark B.39. Let $\boldsymbol{v} \in \mathbb{F}^{\mathrm{n}}$ be a vector such that the k-th component of $\boldsymbol{v}$ is the same as the (i, k)-th entry of $A \in \mathbb{M}(m, n; \mathbb{F})$, and $\boldsymbol{w} \in \mathbb{F}^{\mathrm{n}}$ be a vector such that the k-th component of $\boldsymbol{w}$ is the same as the (k, j)-th entry of $B \in \mathbb{M}(n, \ell; \mathbb{F})$. Then the (i, j)-th entry of AB is simply the inner product of $\boldsymbol{v}$ and $\boldsymbol{w}$ in $\mathbb{F}^{\mathrm{n}}$.

Example B.40.

Definition B.41 (Identity matrix)**.** The

Definition B.42 (Transpose)**.** Let $A = [a_{ij}]_{\boldsymbol{m} \times \mathrm{n}}$ be a $\boldsymbol{m} \times \mathrm{n}$ matrix over scalar field $\mathbb{F}$. The transpose of A, denoted by A^{T}, is the $\mathrm{n} \times \boldsymbol{m}$ matrix given by $[A^{\mathrm{T}}]_{ij} = a_{ji}$.

Definition B.43 (Rank)**.** The ***rank*** of a matrix A is the dimension of the vector space generated (or spanned) by its columns.

Remark B.44. The rank defined above is also referblack to the ***column rank***, and the ***row rank*** of a matrix is the dimension of the vector space spanned by its rows.

Theorem B.45. *The rank of a matrix is the same as the rank of its transpose. In other words, for a given matrix the row rank equals the column rank.*

The proof of this theorem will be postponed and be given in Section B.1.6.

The matrix norm

Each $\boldsymbol{m} \times \mathrm{n}$ matrix $A \in \mathbb{M}(\boldsymbol{m}, \mathrm{n}; \mathbb{F})$ induces a linear map $L : \mathbb{F}^{\mathrm{n}} \to \mathbb{F}^{\boldsymbol{m}}$ in a natural way: let $A = [a_{ij}]_{\boldsymbol{m}\times\mathrm{n}}$ be a $\boldsymbol{m} \times \mathrm{n}$ matrix, $\mathcal{B} = \{\mathrm{e}_j\}_{j=1}^{\mathrm{n}}$ and $\widetilde{\mathcal{B}} = \{\widetilde{\mathrm{e}}_k\}_{k=1}^{\boldsymbol{m}}$ be the standard basis of $\mathbb{F}^{\mathrm{n}}$ and $\mathbb{F}^{\boldsymbol{m}}$, respectively. We define the linear map $L : \mathbb{F}^{\mathrm{n}} \to \mathbb{F}^{\boldsymbol{m}}$ by

$$Lx = \sum_{i=1}^{\boldsymbol{m}} \sum_{j=1}^{\mathrm{n}} a_{ij} x_j \widetilde{\mathrm{e}}_i \in \mathbb{F}^{\boldsymbol{m}}, \quad \text{where} \;\; x = \sum_{j=1}^{\mathrm{n}} x_j \mathrm{e}_j \in \mathbb{F}^{\mathrm{n}} ,$$

or equivalently, $[Lx]_{\widetilde{\mathcal{B}}} = A[x]_{\mathcal{B}}$. The linear map L is called the ***linear map induced by the matrix*** A.

By matrix norms it means the operator norm of the induced linear map. However, as introduced in Section B.1.4, the operator norm of a linear map depends on the norms equipped on the vector spaces. In particular, we have introduced p-norm on $\mathbb{F}^{\mathrm{n}}$, and $\mathbb{F}^{\mathrm{n}}$

Definition B.46. Let $A \in \mathbb{M}(\boldsymbol{m}, \mathrm{n}; \mathbb{F})$ with induced linear map $L : \mathbb{F}^{\mathrm{n}} \to \mathbb{F}^{\boldsymbol{m}}$. The p-norm of A, denoted by $\|A\|_p$, is the operator norm of $L : (\mathbb{F}^{\mathrm{n}}, \|\cdot\|_p) \to (\mathbb{F}^{\boldsymbol{m}}, \|\cdot\|_p)$ given by

$$\|A\|_p = \sup_{\|x\|_p=1} \|Lx\|_p = \sup_{x\neq 0} \frac{\|Lx\|_p}{\|x\|_p} .$$

Remark B.47. We can also choose different p in the domain and the co-domain. In other words, the (p,q)-norm of $A \in \mathbb{M}(\boldsymbol{m}, \mathrm{n}, \mathbb{F})$ is the operator norm of the induced linear map $L : (\mathbb{F}^{\mathrm{n}}, \|\cdot\|_p) \to (\mathbb{F}^{\boldsymbol{m}}, \|\cdot\|_q)$ given by

$$\|A\|_{(p,q)} = \sup_{\|x\|_p=1} \|Lx\|_q = \sup_{x\neq 0} \frac{\|Lx\|_q}{\|x\|_p} .$$

From now on, for notational simplicity we use Ax to denote $[Lx]_{\widetilde{\mathcal{B}}}$, where $\widetilde{\mathcal{B}}$ is the standard basis of the co-domain.

Example B.48. Consider the case $p = 1$ and $p = \infty$, respectively.

1. $p = \infty$: $\|A\|_\infty = \sup\limits_{\|x\|_\infty=1} \|Ax\|_\infty = \max\left\{\sum\limits_{j=1}^{\boldsymbol{m}} |a_{1j}|, \sum\limits_{j=1}^{\boldsymbol{m}} |a_{2j}|, \cdots \sum\limits_{j=1}^{\boldsymbol{m}} |a_{nj}|\right\}$.

   Reason: Let $x = (x_1, x_2, \cdots, x_{\mathrm{n}})^{\mathrm{T}}$ and $A = \big[a_{ij}\big]_{n\times\boldsymbol{m}}$. 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 $\max_{1\leqslant i\leqslant n}\sum_{j=1}^{m}|a_{ij}| = \sum_{j=1}^{m}|a_{kj}|$ for some $1\leqslant k\leqslant n$. Let

$$x = (\operatorname{sgn}(a_{k1}), \operatorname{sgn}(a_{k2}), \cdots, \operatorname{sgn}(a_{kn}))\ .$$

Then $\|x\|_\infty = 1$, and $\|Ax\|_\infty = \sum\limits_{j=1}^{m}|a_{kj}|$.

On the other hand, if $\|x\|_\infty = 1$, then

$$|a_{i1}x_1 + a_{i2}x_2 + \cdots a_{im}x_m| \leqslant \sum_{j=1}^{\boldsymbol{m}}|a_{ij}| \leqslant \max\left\{\sum_{j=1}^{\boldsymbol{m}}|a_{1j}|, \sum_{j=1}^{\boldsymbol{m}}|a_{2j}|, \cdots \sum_{j=1}^{\boldsymbol{m}}|a_{nj}|\right\};$$

thus $\|A\|_\infty = \max\left\{\sum\limits_{j=1}^{\boldsymbol{m}}|a_{1j}|, \sum\limits_{j=1}^{\boldsymbol{m}}|a_{2j}|, \cdots \sum\limits_{j=1}^{\boldsymbol{m}}|a_{nj}|\right\}$. In other words, $\|A\|_\infty$ is the largest sum of the absolute value of row entries.

2. $p = 1$: $\|A\|_1 = \max\left\{\sum\limits_{i=1}^{n}|a_{i1}|, \sum\limits_{i=1}^{n}|a_{i2}|, \cdots, \sum\limits_{i=1}^{n}|a_{im}|\right\}$.

B.1.6 Representation of linear transformations

In Section B.1.5, we see that any $\boldsymbol{m}\times \mathrm{n}$ matrix is associated with a linear map. On the other hand, Suppose that $\mathcal{V}$ is a n-dimensional vector space with basis $\{\boldsymbol{v}_i\}_{i=1}^{\mathrm{n}}$, and $\mathcal{W}$ is a m-dimensional vector space with basis $\{\boldsymbol{w}_j\}_{j=1}^{\boldsymbol{m}}$. Let $L\in\mathscr{L}(\mathcal{V},\mathcal{W})$ and $\boldsymbol{u}\in\mathcal{V}$. Then

$$\boldsymbol{u} = \sum_{i=1}^{\mathrm{n}}\alpha_i\boldsymbol{v}_i\ .$$

Since $L\boldsymbol{u}\in\mathcal{W}$, we can express $L\boldsymbol{u}$ in terms of the basis $\{\boldsymbol{w}_j\}_{j=1}^{\boldsymbol{m}}$. In other words, write

$$L\boldsymbol{u} = \sum_{j=1}^{\boldsymbol{m}}\beta_j\boldsymbol{w}_j\ .$$

On the other hand, by the linearity of L,

$$L\boldsymbol{u} = L\Big(\sum_{i=1}^{\mathrm{n}}\alpha_i\boldsymbol{v}_i\Big) = \alpha_i\sum_{i=1}^{\mathrm{n}}L\boldsymbol{v}_i.$$

Definition B.49 (Kronecker's delta)**.**

B.1.7 Change of basis

Let $\mathcal{B}_1 = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_\mathrm{n}\}$ and $\mathcal{B}_2 = \{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_\mathrm{n}\}$ be two bases of $\mathcal{V}$; that is, any vector $\boldsymbol{u} \in \mathcal{V}$ can be uniquely written as

$$\boldsymbol{u} = \alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_\mathrm{n} \boldsymbol{v}_\mathrm{n} = \beta_1 \boldsymbol{w}_1 + \cdots \beta_\mathrm{n} \boldsymbol{w}_\mathrm{n} \,.$$

Definition B.50.

Definition B.51 (Similarity of matrices)**.**

B.1.8 Elementary row operations and elementary matrices

Definition B.52 (Elementary row operations)**.** For an $\mathrm{n} \times \boldsymbol{m}$ matrix A, three types of elementary row operations can be performed on A:

1. The first type of row operation on A switches all matrix elements on the i-th row with their counterparts on j-th row.

2. The second type of row operation on A multiplies all elements on the i-th row by a non-zero scalar λ.

3. The third type of row operation on A adds j-th row multiplied by a scalar μ to the i-th row.

The elementary row operation on an $\mathrm{n} \times \boldsymbol{m}$ matrix A can be done by multiplying A by an $\mathrm{n} \times \mathrm{n}$ matrix, called an elementary matrix, on the left. The elementary matrices are defined in the following

Definition B.53 (Elementary matrices)**.** An elementary matrix is a matrix which differs from the identity matrix by one single elementary row operation.

1. Switching the i_0-th and j_0-th rows of A, where $i_0 \neq j_0$, is done by left multiplied A by the matrix $\mathrm{E} = [e_{ij}]_{\mathrm{n}\times\mathrm{n}}$ given by

$$e_{ij} = \begin{cases} 1 & \text{if } (i,j) = (i_0, j_0) \text{ or } (i,j) = (j_0, i_0) \\ & \text{or } i = j = k_0 \text{ for some } k_0 \neq i_0, j_0, \\ 0 & \text{otherwise,} \end{cases}$$

or in the matrix form,

$$
\mathrm{E} = \begin{array}{cl}
\left[\begin{array}{ccccccccccc}
1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & \ddots & 0 & & & & & & & & \vdots \\
\vdots & \ddots & 1 & \ddots & & & & & & & \vdots \\
\vdots & & 0 & 0 & 0 & & & 1 & & & \vdots \\
\vdots & & & \ddots & 1 & \ddots & & & & & \vdots \\
\vdots & & & & 0 & \ddots & 0 & & & & \vdots \\
\vdots & & & & & \ddots & 1 & \ddots & & & \vdots \\
\vdots & & & 1 & & & 0 & 0 & 0 & & \vdots \\
0 & & & & & & & \ddots & 1 & \ddots & \vdots \\
0 & & & & & & & & 0 & \ddots & 0 \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1
\end{array}\right]
&
\begin{array}{l}
\\ \\ \\ \leftarrow \text{the } i_0\text{-th row} \\ \\ \\ \\ \leftarrow \text{the } j_0\text{-th row} \\ \\ \\ \\
\end{array}
\end{array}
$$

2. Multiplying the k_0-th row of A by a non-zero scalar λ is done by left multiplied A by the matrix $\mathrm{E} = [e_{ij}]_{\mathrm{n}\times\mathrm{n}}$ given by

$$
e_{ij} = \begin{cases} 0 & \text{if } i \neq j, \\ \lambda & \text{if } i = j = k_0, \\ 1 & \text{otherwise,} \end{cases}
$$

or in the matrix form,

$$
\mathrm{E} = \begin{array}{cl}
\left[\begin{array}{ccccccccc}
1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & 1 & 0 & & & & & & \vdots \\
\vdots & \ddots & \ddots & \ddots & & & & & \vdots \\
\vdots & & 0 & 1 & 0 & & & & \vdots \\
\vdots & & & 0 & \lambda & 0 & & & \vdots \\
\vdots & & & & 0 & 1 & 0 & & \vdots \\
\vdots & & & & & \ddots & \ddots & \ddots & \vdots \\
\vdots & & & & & & 0 & 1 & 0 \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1
\end{array}\right]
&
\begin{array}{l}
\\ \\ \\ \\ \leftarrow \text{the } k_0\text{-th row} \\ \\ \\ \\ \\
\end{array}
\end{array}
$$

3. Adding the j_0-th row of A multiplied by a scalar μ to the i_0-th row, where $i_0 \neq j_0$, is done by left multiplied A by the matrix $\mathrm{E} = [e_{ij}]_{\mathrm{n}\times\mathrm{n}}$ given by

$$
e_{ij} = \begin{cases} 1 & \text{if } i = j, \\ \mu & \text{if } (i,j) = (i_0, j_0), \\ 0 & \text{otherwise,} \end{cases}
$$

or in the matrix form,

$$
\mathrm{E} = \begin{array}{c}
\left[\begin{array}{ccccccccc}
1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & 1 & 0 & & & & & & 0 \\
\vdots & \ddots & \ddots & \ddots & & & \mu & & 0 \\
\vdots & & \ddots & \ddots & \ddots & & & & 0 \\
\vdots & & & 0 & 1 & 0 & & & \vdots \\
\vdots & & & & \ddots & \ddots & \ddots & & \vdots \\
\vdots & & & & & \ddots & \ddots & \ddots & \vdots \\
\vdots & & & & & & 0 & 1 & 0 \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1
\end{array}\right]
\begin{array}{l}
\\ \\ \leftarrow \text{the } i_0\text{-th row} \\ \\ \\ \\ \\ \\ \\
\end{array} \\
\begin{array}{c}
\uparrow \\ \text{the } j_0\text{-th column}
\end{array}
\end{array}
$$

Theorem B.54. *Any invertible* $\mathrm{n} \times \mathrm{n}$ *matrix can be represented as the product of some elementary matrices. In other words, if* A *is an invertible* $\mathrm{n} \times \mathrm{n}$ *matrix, then*

$$\mathrm{A} = \mathrm{E}_k \mathrm{E}_{k-1} \cdots \mathrm{E}_2 \mathrm{E}_1 \,,$$

where $\mathrm{E}_1, \cdots, \mathrm{E}_k$ *are elementary matrices.*

B.1.9 Determinants

In order to introduce the notion of the determinant of square matrices, we need to talk about permutations first. Note that there are many other ways of defining determinants, but it is quite elegant to use the notion of permutations, and we can derive a lot of useful results via this definition.

Definition B.55 (Permutations)**.** A sequence $(k_1, k_2, \cdots, k_\mathrm{n})$ of positive integers not exceeding n, with the property that no two of the k_i are equal, is called a **permutation of degree** n. The collection of all permutations of degree n is denoted by $\mathbb{P}(\mathrm{n})$.

A sequence $(k_1, k_2, \cdots, k_\mathrm{n})$ can be obtained from the sequence $(1, 2, \cdots, \mathrm{n})$ by a finite number of interchanges of pairs of elements. For example, if $k_1 \neq 1$, we can transpose 1 and k_1, obtaining $(k_1, \cdots, 1, \cdots)$. Proceeding in this way we shall arrive at the sequence $(k_1, k_2, \cdots, k_\mathrm{n})$ after n or less such interchanges of pairs.

In general, a permutation $(k_1, k_2, \cdots, k_\mathrm{n})$ can be expressed as

$$\tau_{(i_N,j_N)} \cdots \tau_{(i_2,j_2)}\tau_{(i_1,j_1)}(1, 2, \cdots, \mathrm{n}) = (k_1, k_2, \cdots, k_\mathrm{n}),$$

where $\tau_{(i,j)}$ is a "pair-interchange operator" which swaps the i-th and the j-th elements (of the object fed into), and N is the number of pair interchanges. We call such pair-interchange operators the permutation operator. Since $\tau_{(i,j)}$ is the inverse operator of itself, we also have

$$\tau_{(i_1,j_1)}\tau_{(i_2,j_2)} \cdots \tau_{(i_N,j_N)}(k_1, k_2, \cdots, k_\mathrm{n}) = (1, 2, \cdots, \mathrm{n}).$$

We remark here that the number of pair interchanges (from $(1, 2, \cdots, \mathrm{n})$ to $(k_1, k_2, \cdots, k_\mathrm{n})$) is not unique; nevertheless, if two processes of pair interchanges lead to the same permutation, then the numbers of interchanges differ by an even number. This leads to the following

Definition B.56 (Even and odd permutations)**.** A permutation $(k_1, \cdots, k_\mathrm{n})$ is called an even (odd) permutation of degree n if the number requiblack to interchange pairs of $(1, 2, \cdots, \mathrm{n})$ in order to obtain $(k_1, k_2, \cdots, k_\mathrm{n})$ is even (odd).

Example B.57. If n = 3, the permutation $(3, 1, 2)$ can be obtained by interchanging pairs of $(1, 2, 3)$ twice:

$$(1, 2, 3) \xrightarrow{\tau_{(1,3)}} (3, 2, 1) \xrightarrow{\tau_{(2,3)}} (3, 1, 2);$$

thus $(3, 1, 2)$ is an even permutation of $(1, 2, 3)$. On the other hand, $(1, 3, 2)$ is obtained by interchanging pairs of $(1, 2, 3)$ once:

$$(1, 2, 3) \xrightarrow{\tau_{(2,3)}} (1, 3, 2);$$

thus $(1, 3, 2)$ is an odd permutation of $(1, 2, 3)$.

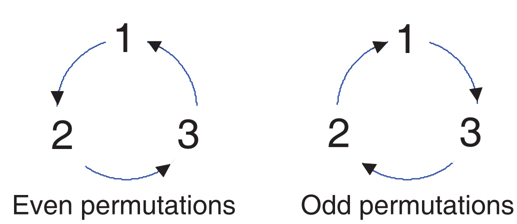


Figure B.1: Even and odd permutations of degree 3

For n = 3, the even and odd permutations can also be viewed as the orientation of the permutation (k_1, k_2, k_3). To be more precise, if $(1, 2, 3)$ is arranged in a counter-clockwise orientation (see Figure B.1), then an even permutation of degree 3 is a permutation in the counter-clockwise orientation, while an odd permutation of degree 3 is a permutation in the clockwise orientation. From figure B.1, it is easy to see that $(3, 1, 2)$ is an even permutation of degree 3 and $(1, 3, 2)$ is an odd permutation of degree 3.

Definition B.58 (The permutation symbol)**.** The permutation symbol $\varepsilon_{k_1k_2\cdots k_\mathrm{n}}$ is a function of permutations of degree n defined by

$$\varepsilon_{k_1k_2\cdots k_\mathrm{n}} = \begin{cases} 1 & \text{if } (k_1, k_2, \cdots, k_\mathrm{n}) \text{ is an even permutation of degree n,} \\ -1 & \text{if } (k_1, k_2, \cdots, k_\mathrm{n}) \text{ is an odd permutation of degree n.} \end{cases}$$

Remark B.59. One can extend the domain the permutation symbol to all the sequence $(k_1, k_2, \cdots, k_\mathrm{n})$ by defining that $\varepsilon_{k_1k_2\cdots k_\mathrm{n}} = 0$ if $(k_1, k_2, \cdots, k_\mathrm{n})$ is not a permutation of degree n.

Definition B.60 (Determinants)**.** Given an $\mathrm{n}\times\mathrm{n}$ matrix $\mathrm{A} = [a_{ij}]$, the determinants of A, denoted by det(A), is defined by

$$\det(\mathrm{A}) = \sum_{(k_1,\cdots,k_\mathrm{n})\in\mathbb{P}(\mathrm{n})} \varepsilon_{k_1k_2\cdots k_\mathrm{n}} \prod_{\ell=1}^{\mathrm{n}} a_{\ell k_\ell}.$$

We note that the product $\prod\limits_{\ell=1}^{\mathrm{n}} a_{\ell k_\ell}$ in the definition of the determinant is formed by multiplying n-elements which appears exactly once in each row and column.

Proposition B.61. *Let* E *be an elementary matrix. Then*

1. $\det(\mathrm{E}) \neq 0$.

2. $\det(\mathrm{E}) = \det(\mathrm{E}^\mathrm{T})$.

3. *If* A *is an* $\mathrm{n}\times\mathrm{n}$ *matrix, then* $\det(\mathrm{EA}) = \det(\mathrm{E})\det(\mathrm{A})$.

The proof of the proposition above is not difficult, and is left as an exercise.

Corollary B.62. *Let* $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_{\rm n} \in \mathbb{R}^{\rm n}$ *be (column) vectors,* $c \in \mathbb{R}$*, and*

$$
\begin{aligned}
&\mathrm{A} = [\boldsymbol{v}_1 \vdots \cdots \vdots \boldsymbol{v}_{\rm n}]\,,\\
&\mathrm{B} = [\boldsymbol{v}_1 \vdots \cdots \vdots \boldsymbol{v}_{j-1} \vdots c\boldsymbol{v}_j \vdots \boldsymbol{v}_{j+1} \vdots \cdots \vdots \boldsymbol{v}_{\rm n}]\,,\\
&\mathrm{C} = [\boldsymbol{v}_1 \vdots \cdots \vdots \boldsymbol{v}_{j-1} \vdots \boldsymbol{v}_j + \boldsymbol{v}_i \vdots \boldsymbol{v}_{j+1} \vdots \cdots \vdots \boldsymbol{v}_{\rm n}] \quad \textit{for some } i \neq j\,.
\end{aligned}
$$

Then $\det(\mathrm{B}) = c\det(\mathrm{A})$*, and* $\det(\mathrm{C}) = \det(\mathrm{A})$.

Proof. The corollary is easily concluded since $\mathrm{B} = \mathrm{E}_1\mathrm{A}$ and $\mathrm{C} = \mathrm{E}_2\mathrm{A}$ for some elementary matrices E_1 and E_2 with $\det(\mathrm{E}_1) = c$ and $\det(\mathrm{E}_2) = 1$. □

Corollary B.63. *Let* A *be an* $\mathrm{n} \times \mathrm{n}$ *invertible matrix. Then* $\det(\mathrm{A}) \neq 0$.

Proof. Since A is invertible, Theorem B.54 implies that

$$\mathrm{A} = \mathrm{E}_k\mathrm{E}_{k-1}\cdots\mathrm{E}_2\mathrm{E}_1$$

for some elementary matrices $\mathrm{E}_1, \cdots, \mathrm{E}_k$, and this corollary follows from Proposition B.61. □

Corollary B.64. *Let* A *be an* $\mathrm{n} \times \mathrm{n}$ *matrix. Then the determinant of* A *and* $\mathrm{A}^{\rm T}$*, the transpose of* A*, are the same; that is,*

$$\det(\mathrm{A}) = \det(\mathrm{A}^{\rm T}).$$

Proof. If A is not invertible, then $\mathrm{A}^{\rm T}$ is not invertible either because of Theorem B.45. Therefore, $\det(\mathrm{A}) = 0 = \det(\mathrm{A}^{\rm T})$.

Now suppose that A is invertible. Then Theorem B.54 implies that

$$\mathrm{A} = \mathrm{E}_k\mathrm{E}_{k-1}\cdots\mathrm{E}_2\mathrm{E}_1$$

for some elementary matrices $\mathrm{E}_1, \cdots, \mathrm{E}_k$. Since all $\mathrm{E}_j^{\rm T}$'s are also elementary matrices, by Proposition B.61 we conclude that

$$
\begin{aligned}
\det(\mathrm{A}^{\rm T}) = \det(\mathrm{E}_1^{\rm T}\cdots\mathrm{E}_k^{\rm T}) &= \det(\mathrm{E}_1^{\rm T})\cdots\det(\mathrm{E}_k^{\rm T})\\
&= \det(\mathrm{E}_k^{\rm T})\cdots\det(\mathrm{E}_1^{\rm T})\\
= \det(\mathrm{E}_k)\cdots\det(\mathrm{E}_1) &= \det(\mathrm{E}_k\cdots\mathrm{E}_1) = \det(\mathrm{A})\,.
\end{aligned}
$$
□

Corollary B.65. *Let* A, B *be* $\mathrm{n} \times \mathrm{n}$ *matrices. Then* $\det(\mathrm{AB}) = \det(\mathrm{A})\det(\mathrm{B})$.

Proof. If A is not invertible, then AB is not invertible either; thus in this case $\det(\mathrm{A})\det(\mathrm{B}) = 0 = \det(\mathrm{AB})$.

Now suppose that A is invertible. Then Theorem B.54 implies that

$$\mathrm{A} = \mathrm{E}_k\mathrm{E}_{k-1}\cdots\mathrm{E}_2\mathrm{E}_1$$

for some elementary matrices $\mathrm{E}_1, \cdots, \mathrm{E}_k$. As a consequence, Proposition B.61 implies that

$$\begin{aligned}\det(\mathrm{AB}) = \det(\mathrm{E}_k\cdots\mathrm{E}_1\mathrm{B}) &= \det(\mathrm{E}_k)\det(\mathrm{E}_{k-1}\cdots\mathrm{E}_1\mathrm{B})\\ &= \cdots = \det(\mathrm{E}_k)\cdots\det(\mathrm{E}_1)\det(\mathrm{B})\\ &= \det(\mathrm{E}_k\cdots\mathrm{E}_1)\det(\mathrm{B}) = \det(\mathrm{A})\det(\mathrm{B})\,.\end{aligned}$$ □

Definition B.66 (Cofactor matrices)**.** Let A be an $\mathrm{n}\times\mathrm{n}$ matrix, and $\mathrm{A}(\hat{i},\hat{j})$ be the $(\mathrm{n}-1)\times(\mathrm{n}-1)$ matrix obtained by eliminating the i-th row and j-th column of A; that is,

$$\mathrm{A}(\hat{i},\hat{j}) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1(j-1)} & a_{1(j+1)} & \cdots & a_{1\mathrm{n}}\\ \vdots & \ddots & & \vdots & \vdots & & \vdots\\ a_{(i-1)1} & a_{(i-1)2} & \cdots & a_{(i-1)(j-1)} & a_{(i-1)(j+1)} & \cdots & a_{(i-1)\mathrm{n}}\\ a_{(i+1)1} & a_{(i+1)2} & \cdots & a_{(i+1)(j-1)} & a_{(i+1)(j+1)} & \cdots & a_{(i+1)\mathrm{n}}\\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots\\ a_{\mathrm{n}1} & a_{\mathrm{n}2} & \cdots & a_{\mathrm{n}(j-1)} & a_{\mathrm{n}(j+1)} & \cdots & a_{\mathrm{nn}} \end{bmatrix}.$$

The cofactor matrix of A, denoted by Cof(A), is given by $\mathrm{Cof(A)} = [m_{ij}]$ with m_{ij} defined by

$$m_{ij} = (-1)^{i+j}\det\big(\mathrm{A}(\hat{j},\hat{i})\big).$$

Lemma B.67. *Let* A *be an* $\mathrm{n}\times\mathrm{n}$ *matrix. Then*

$$\det\big(\mathrm{A}(\hat{i},\hat{j})\big) = (-1)^{i+j}\sum_{(k_1,\cdots,k_\mathrm{n})\in\mathbb{P}(\mathrm{n}),\,k_i=j}\varepsilon_{k_1k_2\cdots k_\mathrm{n}}\prod_{\substack{1\le\ell\le\mathrm{n}\\ \ell\neq i}}a_{\ell k_\ell}.$$

Proof. Fix $(i,j)\in\{1,2,\cdots,\mathrm{n}\}\times\{1,2,\cdots,\mathrm{n}\}$. The matrix $\mathrm{A}(\hat{i},\hat{j})$ is given by $\mathrm{A}(\hat{i},\hat{j}) = [b_{\alpha\beta}]$, where $\alpha,\beta = 1,2,\cdots,\mathrm{n}-1$, and

$$b_{\alpha\beta} = \begin{cases} a_{\alpha\beta} & \text{if } \alpha<i \text{ and } \beta<j,\\ a_{(\alpha+1)\beta} & \text{if } \alpha>i \text{ and } \beta<j,\\ a_{\alpha(\beta+1)} & \text{if } \alpha<i \text{ and } \beta>j,\\ a_{(\alpha+1)(\beta+1)} & \text{if } \alpha>i \text{ and } \beta>j.\end{cases}$$

For every $1 \leqslant \tau \leqslant \mathrm{n}-1$ and permutations $(\sigma_1, \sigma_2, \cdots, \sigma_{\mathrm{n}-1})$ of degree $\mathrm{n}-1$, there exist a unique $\ell \neq i$ and a unique permutation $(k_1, k_2, \cdots, k_{\mathrm{n}})$ of degree n such that $k_i = j$ and $b_{\tau\sigma_\tau} = a_{\ell k_\ell}$. Moreover, if a process of pair interchanges of the permutation permutation $(\sigma_1, \sigma_2, \cdots, \sigma_{\mathrm{n}-1})$ leads to $(1, 2, \cdots, \mathrm{n}-1)$, then similar process of pair interchanges of the permutation $(k_1, k_2, \cdots, k_{i-1}, j, k_{i+1}, \cdots, k_{\mathrm{n}})$, by leaving the i-th slot fixed, leads to the permutation of degree n

$$\begin{cases} (1, 2, \cdots, j-1, j+1, \cdots, i-1, j, i, \cdots, \mathrm{n}) & \text{if } \ i > j, \\ (1, 2, \cdots, i-1, j, i, \cdots, j-1, j+1, \cdots, \mathrm{n}) & \text{if } \ i < j, \\ \qquad\qquad (1, 2, \cdots, \mathrm{n}) & \text{if } \ j = i. \end{cases}$$

For the case that $i \neq j$, another $|i-j|$-times of pair interchanges leads to $(1, 2, \cdots, \mathrm{n})$. To be more precise, suppose that $i > j$. We first interchange the $(i-2)$-th and the $(i-1)$-th components, and then interchange that $(i-3)$-th and the $(i-2)$-th components, and so on. After $(i-j)$-times of pair interchanges, we reach $(1, 2, \cdots, \mathrm{n})$. Symbolically,

$$\begin{array}{c} (1, 2, \cdots, j-1, j+1, \cdots, i-1, j, i, \cdots, \mathrm{n}) \\ \downarrow \tau_{(i-2,\, i-1)} \\ (1, 2, \cdots, j-1, j+1, \cdots, i-2, j, i-1, \cdots, \mathrm{n}) \\ \downarrow \tau_{(i-3,\, i-2)} \\ (1, 2, \cdots, j-1, j+1, \cdots, i-3, j, i-2 \cdots, \mathrm{n}) \\ \downarrow \\ \vdots \\ \downarrow \\ (1, 2, \cdots, \mathrm{n}). \end{array}$$

Similar argument applies to the case $i < j$; thus

$$\varepsilon_{\sigma_1\sigma_2\cdots\sigma_{\mathrm{n}-1}} = (-1)^{|i-j|} \varepsilon_{k_1 k_2 \cdots k_{\mathrm{n}}} = (-1)^{i+j} \varepsilon_{k_1 k_2 \cdots k_{\mathrm{n}}}.$$

As a consequence,

$$\begin{aligned} \det\big(\mathrm{A}(\hat{i}, \hat{j})\big) &= \sum_{(\sigma_1, \sigma_2, \cdots, \sigma_{\mathrm{n}-1}) \in \mathbb{P}(\mathrm{n}-1)} \varepsilon_{\sigma_1\sigma_2\cdots\sigma_{\mathrm{n}-1}} \prod_{\tau=1}^{\mathrm{n}-1} b_{\tau\sigma_\tau} \\ &= (-1)^{i+j} \sum_{(k_1, \cdots, k_{\mathrm{n}}) \in \mathbb{P}(\mathrm{n}),\, k_i = j} \varepsilon_{k_1 k_2 \cdots k_{\mathrm{n}}} \prod_{\substack{1 \leqslant \ell \leqslant \mathrm{n} \\ \ell \neq i}} a_{\ell k_\ell}\,. \end{aligned}$$ □

Theorem B.68. *Let* A *be an* $\mathrm{n} \times \mathrm{n}$ *matrix. Then*

$$\mathrm{Cof(A)A} = \mathrm{ACof(A)} = \det(\mathrm{A})\mathrm{I}_{\mathrm{n}}.$$

Proof. Let $\mathrm{A} = [a_{ij}]$. By definition of matrix multiplications,

$$\big(\mathrm{Cof}(\mathrm{A})\mathrm{A}\big)_{ij} = \sum_{m=1}^{\mathrm{n}} \Big[\sum_{(k_1,\cdots,k_\mathrm{n})\in\mathbb{P}(\mathrm{n}),\, k_m=i} \varepsilon_{k_1k_2\cdots k_\mathrm{n}} \prod_{\substack{1\leqslant \ell\leqslant \mathrm{n}\\ \ell\neq m}} a_{\ell k_\ell}\Big] a_{mj}$$
$$= \begin{cases} \displaystyle\sum_{(k_1,\cdots,k_\mathrm{n})\in\mathbb{P}(\mathrm{n})} \varepsilon_{k_1k_2\cdots k_\mathrm{n}} \prod_{\ell=1}^{\mathrm{n}} a_{\ell k_\ell} & \text{if } \ i=j, \\ 0 & \text{if } \ i\neq j. \end{cases}$$

The conclusion then follows from the definition of the determinant. □

Corollary B.69. *Let* $\mathrm{A} = [a_{ij}]$ *be an* $\mathrm{n}\times\mathrm{n}$ *matrix, and* $\mathrm{C} = [c_{ij}]$ *be the cofactor matrix of* A. *Then*

$$\det(\mathrm{A}) = \sum_{j=1}^{\mathrm{n}} a_{ij}c_{ji} = \sum_{j=1}^{\mathrm{n}} a_{ji}c_{ij} \quad \forall\, 1\leqslant i\leqslant \mathrm{n}\,.$$

Corollary B.70. *Let* A *be an* $\mathrm{n}\times\mathrm{n}$ *matrix and* $\det(\mathrm{A})\neq 0$. *Then the matrix* $\dfrac{\mathrm{Cof}(\mathrm{A})}{\det(\mathrm{A})}$ *is the inverse matrix of* A, *or equivalently,*

$$\mathrm{Cof}(\mathrm{A}) = \det(\mathrm{A})\mathrm{A}^{-1}. \tag{B.2}$$

Variations of determinants

Let δ be an operator satisfying the "product rule" $\delta(fg) = f\delta g + (\delta f)g$. Typically δ will be differential operators. By the definition of the determinant,

$$\delta\det(\mathrm{A}) = \sum_{(k_1,\cdots,k_\mathrm{n})\in\mathbb{P}(\mathrm{n})} \varepsilon_{k_1k_2\cdots k_\mathrm{n}}\delta \prod_{\ell=1}^{\mathrm{n}} a_{\ell k_\ell}$$
$$= \sum_{i=1}^{\mathrm{n}} \Big[\sum_{(k_1,\cdots,k_\mathrm{n})\in\mathbb{P}(\mathrm{n})} \varepsilon_{k_1k_2\cdots k_\mathrm{n}}\delta a_{ik_i} \prod_{\substack{1\leqslant \ell\leqslant \mathrm{n}\\ \ell\neq i}} a_{\ell k_\ell}\Big]$$
$$= \sum_{i,j=1}^{\mathrm{n}} \Big[\sum_{(k_1,\cdots,k_\mathrm{n})\in\mathbb{P}(\mathrm{n}),\, k_i=j} \varepsilon_{k_1k_2\cdots k_\mathrm{n}}\delta a_{ik_i} \prod_{\substack{1\leqslant \ell\leqslant \mathrm{n}\\ \ell\neq i}} a_{\ell k_\ell}\Big]$$
$$= \sum_{i,j=1}^{\mathrm{n}} (-1)^{i+j}\det\big(\mathrm{A}(\hat{i},\hat{j})\big)\delta a_{ij} = \sum_{i,j=1}^{\mathrm{n}} m_{ji}\delta a_{ij},$$

where $\mathrm{Cof}(\mathrm{A}) = [m_{ij}]_{\mathrm{n}\times\mathrm{n}}$ is the cofactor matrix of A. Therefore, we obtain the following

Theorem B.71. *Let* A *be an* $\mathrm{n} \times \mathrm{n}$ *matrix, and* δ *be an operator satisfying* $\delta(fg) = f\delta g + (\delta f)g$ *whenever the product makes sense. Then*

$$\delta \det(\mathrm{A}) = \mathrm{tr}\big(\mathrm{Cof}(\mathrm{A})\delta\mathrm{A}\big) = \det(\mathrm{A})\mathrm{tr}\big(\mathrm{A}^{-1}\delta\mathrm{A}\big), \tag{B.3}$$

where $\delta\mathrm{A} \equiv [\delta a_{ij}]_{\mathrm{n}\times\mathrm{n}}$ *if* $\mathrm{A} = [a_{ij}]_{\mathrm{n}\times\mathrm{n}}$.

The differentiation of the Jacobian

In this sub-section, we study the differentiation of a special determinant, the Jacobian.

Example B.72. Suppose that $\psi : \Omega \subseteq \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^{\mathrm{n}}$ is a given diffeomorphism (thus $\det(\nabla\psi) \neq 0$). Let $\mathrm{M} = \nabla\psi$, and $\mathrm{J} = \det(\mathrm{M})$. By Corollary B.70, the cofactor matrix of M is JM^{-1}. Letting δ be a (first order) partial differential operator which satisfies $\delta(fg) = f\delta g + (\delta f)g$, by Theorem B.71 we find that

$$\delta\mathrm{J} = \mathrm{tr}(\mathrm{JM}^{-1}\delta\mathrm{M}) = \sum_{i,j=1}^{\mathrm{n}} \mathrm{JA}^j_i \delta\psi^i_{,j}\,, \tag{B.4}$$

where $\mathrm{A}^j_i = a_{ji}$ with $\mathrm{M}^{-1} = [a_{ji}]_{\mathrm{n}\times\mathrm{n}}$, and $f_{,j} \equiv \dfrac{\partial f}{\partial x_j}$.

Remark B.73. From now on we sometimes write the row index of a matrix as a super-script for the following reason: if $\psi : \Omega \subseteq \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^{\mathrm{m}}$ is a diffeomorphism, then $\nabla\psi$ is usually expressed by

$$\nabla\psi = \begin{bmatrix} \dfrac{\partial\psi_1}{\partial x_1} & \dfrac{\partial\psi_1}{\partial x_2} & \cdots & \dfrac{\partial\psi_1}{\partial x_{\mathrm{n}}} \\ \dfrac{\partial\psi_2}{\partial x_1} & \dfrac{\partial\psi_2}{\partial x_2} & \cdots & \dfrac{\partial\psi_2}{\partial x_{\mathrm{n}}} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial\psi_{\mathrm{m}}}{\partial x_1} & \dfrac{\partial\psi_{\mathrm{m}}}{\partial x_2} & \cdots & \dfrac{\partial\psi_{\mathrm{m}}}{\partial x_{\mathrm{n}}} \end{bmatrix};$$

thus the (i,j) element of $\nabla\psi$ is $\dfrac{\partial\psi_i}{\partial x_j}$, and the row index i appears "above" the column index j.

Theorem B.74 (Piola's identity)**.** *Let* $\psi : \Omega \subseteq \mathbb{R}^{\mathrm{n}} \to \mathbb{R}^{\mathrm{n}}$ *be a diffeomorphism, and* $[a_{ij}]_{\mathrm{n}\times\mathrm{n}}$ *be the cofactor matrix of* $\nabla\psi$*. Then*

$$\sum_{j=1}^{\mathrm{n}} \frac{\partial}{\partial x_j} a_{ji} = 0. \tag{B.5}$$

Proof. Let $\mathrm{J} = \det(\nabla\psi)$ and $\mathrm{A} = (\nabla\psi)^{-1}$. Then $a_{ji} = \mathrm{J}\mathrm{A}^j_i$. Moreover, since $\mathrm{A}\nabla\psi = \mathrm{I}_n$, $\sum_{r=1}^{\mathrm{n}} \mathrm{A}^j_r \psi^r_{,s} = \delta_{js}$; thus

$$0 = \Big[\sum_{r=1}^{\mathrm{n}} \mathrm{A}^j_r \psi^r_{,s}\Big]_{,k} = \sum_{r=1}^{\mathrm{n}} \big[\mathrm{A}^j_{r,k}\psi^r_{,s} + \mathrm{A}^j_r \psi^r_{,sk}\big]$$

which, after multiplying the equality above by A^s_i and then summing over s, implies that

$$\mathrm{A}^j_{i,k} = -\sum_{r,s=1}^{\mathrm{n}} \mathrm{A}^j_r \psi^r_{,sk} \mathrm{A}^s_i . \tag{B.6}$$

As a consequence,

$$\sum_{j=1}^{\mathrm{n}} \frac{\partial}{\partial x_j}(\mathrm{J}\mathrm{A}^j_i) = \sum_{j=1}^{\mathrm{n}} \sum_{r,s=1}^{\mathrm{n}} \Big[\mathrm{J}\mathrm{A}^r_s \psi^s_{,rj} \mathrm{A}^j_i - \mathrm{J}\mathrm{A}^j_r \psi^r_{,sj} \mathrm{A}^s_i\Big] = 0\,. \qquad \square$$

B.1.10 Matrix diagonalization and the Jordan forms

Not yet completed!!!

B.2 Vector Calculus

B.2.1 The line integrals

Curves

Definition B.75. A subset $C \subseteq \mathbb{R}^{\mathrm{n}}$ is called a ***curve*** if C is the image of an interval $I \subseteq \mathbb{R}$ under the continuous map $\gamma : I \to \mathbb{R}^{\mathrm{n}}$ (that is, $C = \gamma(I)$). The continuous map $\gamma : I \to \mathbb{R}^{\mathrm{n}}$ is called a ***parametrization*** of the curve. A curve C with parametrization $\gamma : I \to \mathbb{R}^{\mathrm{n}}$ is called ***simple*** if γ is injective; that is, $\gamma(x) = \gamma(y)$ implies that $x = y$. A curve C with parametrization $\gamma : I \to \mathbb{R}^{\mathrm{n}}$ is called ***closed*** if $I = [a,b]$ for some closed interval $[a,b] \subseteq \mathbb{R}$ and $\gamma(a) = \gamma(b)$.

Example B.76.

Example B.77.

Definition B.78. The length of curve $C \subseteq \mathbb{R}^{\mathrm{n}}$ parametrized by $\gamma : [a,b] \to \mathbb{R}^{\mathrm{n}}$ is defined as the number

$$\ell(C) \equiv \sup\Big\{ \sum_{i=1}^{k} \big\|\gamma(t_i) - \gamma(t_{i-1})\big\|_{\mathbb{R}^{\mathrm{n}}} \,\Big|\, k \in \mathbb{N} \text{ and } a = t_0 < t_1 < \cdots < t_k = b \Big\}.$$

DEFINITION B.79 (Recifiable curves)**.** A curve $C \subseteq \mathbb{R}^{\mathrm{n}}$ with parametrization $\gamma : I \to \mathbb{R}^{\mathrm{n}}$ is call ***rectifiable*** if there is an homeomorphism $\varphi : \widetilde{I} \to I$, where $\widetilde{I}$ is again an interval, such that the map $\gamma \circ \varphi : \widetilde{I} \to \mathbb{R}^{\mathrm{n}}$ is Lipschitz.

REMARK B.80. 1. By an homeomorphism it means a continuous bijection whose inverse is also continuous.

2. We can think of a curve as an equivalence class of continuous maps $\gamma : I \to \mathbb{R}^{\mathrm{n}}$, where two parametrization $\gamma : I \to \mathbb{R}^{\mathrm{n}}$ and $\widetilde{\gamma} : \widetilde{I} \to \mathbb{R}^{\mathrm{n}}$ are equivalent if and only if there is an homeomorphism $\varphi : \widetilde{I} \to I$ such that $\widetilde{\gamma} = \gamma \circ \varphi$. Each element of the equivalence class is a parametrization of the curve and thus a rectifiable curve is a curve which has a Lipschitz continuous parametrization.

3. The length of a rectifiable curve parametrized by $\gamma : [a, b] \to \mathbb{R}^{\mathrm{n}}$ is finite since by choosing a Lipschitz parametrization $\widetilde{\gamma} : [c, d] \to \mathbb{R}^{\mathrm{n}}$, the number

$$\Big\{ \sum_{i=1}^{k} \big\| \widetilde{\gamma}(t_i) - \widetilde{\gamma}(t_{i-1}) \big\|_{\mathbb{R}^{\mathrm{n}}} \,\Big|\, k \in \mathbb{N} \text{ and } c = t_0 < t_1 < \cdots < t_k = d \Big\}$$

is bounded from above by $M(d - c)$, where M is the Lipschitz constant of $\widetilde{\gamma}$.

DEFINITION B.81. A curve $C \subseteq \mathbb{R}^{\mathrm{n}}$ is said to be of class $\mathscr{C}^k$ or a $\mathscr{C}^k$-curve if there exists a parametrization $\gamma : I \to \mathbb{R}^{\mathrm{n}}$ such that γ is k-times continuously differentiable. Such a parametrization is called a $\mathscr{C}^k$-parametrization of the curve. If there exists a parametrization $\gamma : I \to \mathbb{R}$ which is of class $\mathscr{C}^k$ for all $k \in \mathbb{N}$, then the curve is said to be ***smooth***. A curve $C \subseteq \mathbb{R}^{\mathrm{n}}$ is said to be ***regular*** if there exists a $\mathscr{C}^1$-parametrization $\gamma : I \to \mathbb{R}^{\mathrm{n}}$ such that $\gamma'(t) \neq \mathbf{0}$ for all $t \in I$.

THEOREM B.82. *Let $C \subseteq \mathbb{R}^{\mathrm{n}}$ be a simple curve with $\mathscr{C}^1$-parametrization $\gamma : [a, b] \to \mathbb{R}^{\mathrm{n}}$. Then*

$$\ell(C) = \int_a^b \|\gamma'(t)\|_{\mathbb{R}^{\mathrm{n}}} \, dt \,.$$

DEFINITION B.83. Let $C \subseteq \mathbb{R}^{\mathrm{n}}$ be a curve with finite length. An ***arc length parametrization*** of C is an injective parametrization $\gamma : [a, b] \to \mathbb{R}^{\mathrm{n}}$ such that the length of the curve $\gamma([a, s])$ is exactly s; that is,

$$\ell\big(\gamma([a, s])\big) = s \qquad \forall\, s \in [a, b] \,.$$

Note that the function $s(t) = \int_a^t \|\gamma'(t)\|_{\mathbb{R}^n}\, dt$ is increasing in t. If there exists a $\mathscr{C}^1$-parametrization $\gamma : [a,b] \to \mathbb{R}^n$ such that $\gamma'(t) \neq 0$ for all $t \in [a,b]$, then s has a $\mathscr{C}^1$-inverse by the Inverse Function Theorem. In general, the arc length parametrization of a rectifiable curve exists, and we have the following

Theorem B.84. *Let $C \subseteq \mathbb{R}^n$ be a rectifiable simple curve. Then there exists a arc length parametrization of C.*

Proof. Let C be parametrized by $\gamma : [a,b] \to \mathbb{R}^n$ for some Lipschitz injective map $\gamma : [a,b] \to \mathbb{R}^n$. Define

$$L(t) = \sup\Big\{ \sum_{i=1}^{k} \big\|\gamma(t_i) - \gamma(t_{i-1})\big\|_{\mathbb{R}^n} \,\Big|\, k \in \mathbb{N} \text{ and } a = t_0 < t_1 < \cdots < t_k = t \Big\};$$

that is, $L(t)$ is the length of the curve $\gamma([a,t])$. Then L is continuous and strictly increasing; thus the inverse $L^{-1} : [0, \ell(C)] \to [a,b]$ exists and is continuous which allows us to reparametrize the curve C as

$$\widetilde{\gamma} = \gamma \circ L^{-1} : [0, \ell(C)] \to \mathbb{R}^n .$$

The parametrization $\widetilde{\gamma} : [0, \ell(C)] \to \mathbb{R}^n$ is an arc length parametrization of C since

1. $C = \widetilde{\gamma}([0, \ell(C)])$.

2. **Not yet completed!!!**

□

The line element and the line integral

Not yet completed!!!

Definition B.85. Let $C \subseteq \mathbb{R}^n$ be a simple curve with parametrization $\gamma : [a,b] \to \mathbb{R}^n$, and $f : C \to \mathbb{R}$ be a non-negative real-valued function. The ***line integral*** of f over C, denoted by $\int_C f ds$, is the number

$$\sup\Big\{ \sum_{i=1}^{k} f(\xi_i)\big\|\gamma(t_i) - \gamma(t_{i-1})\big\|_{\mathbb{R}^n} \,\Big|\, k \in \mathbb{N}, \xi_i \in \gamma([t_{i-1}, t_i]), a = t_0 < t_1 < \cdots < t_k = b \Big\}$$

provided that the infimum and supremum are the same.

Remark B.86. In particular, if $f \equiv 1$, then $\ell(C) = \int_C 1 ds \equiv \int_C ds$.

Theorem B.87. *Let $C \subseteq \mathbb{R}^n$ be a simple curve with $\mathscr{C}^1$-parametrization $\gamma : [a,b] \to \mathbb{R}^n$, and $f : C \to \mathbb{R}$ be a real-valued continuous function. Then*

$$\int_C f ds = \int_a^b f\big(\gamma(t)\big) \|\gamma'(t)\|_{\mathbb{R}^n} \, dt \, .$$

Torsions and curvatures

Not yet completed!!!

B.2.2 The surface integrals

Surfaces

Definition B.88. A subset $\mathcal{S} \subseteq \mathbb{R}^3$ is called a surface if for each $p \in \mathcal{S}$, there exist an open neighborhood $\mathcal{U} \subseteq \mathcal{S}$ of p, an open set $\mathcal{V} \subseteq \mathbb{R}^2$, and a continuous map $\varphi : \mathcal{U} \to \mathcal{V}$ such that $\varphi : \mathcal{U} \to \mathcal{V}$ is one-to-one, onto, and its inverse $\psi = \varphi^{-1}$ is also continuous. Such a pair $\{\mathcal{U}, \varphi\}$ is called a coordinate chart (or simply chart) at p, and $\{\mathcal{V}, \psi\}$ is called a (local) parametrization at p.

Remark B.89. In some literatures the surface is defined in the following equivalent but reversed way: A subset $\mathcal{S} \subseteq \mathbb{R}^3$ is a surface if for each $p \in \mathcal{S}$, there exists a neighborhood $\mathcal{U} \subseteq \mathbb{R}^3$ of p and a map $\psi : \mathcal{V} \to \mathcal{U} \cap \mathcal{S}$ of an open set $\mathcal{V} \subseteq \mathbb{R}^2$ onto $\mathcal{U} \cap \mathcal{S} \subseteq \mathbb{R}^3$ such that ψ is a homeomorphism; that is, ψ has an inverse $\varphi = \psi^{-1} : \mathcal{U} \cap \mathcal{S} \to \mathcal{V}$ which is continuous. The mapping ψ is called a parametrization or a system of (local) coordinates in (a neighborhood of) p.

Definition B.90 (Regular surfaces)**.** A surface $\mathcal{S} \subseteq \mathbb{R}^3$ is said to be regular if for each $p \in \mathcal{S}$, there exists a differentiable local parametrization $\psi : \mathcal{V} \to \mathcal{U} \subseteq \mathbb{R}^3$ such that $D\psi(q)$, the derivative of ψ at q, has full rank for all $q \in \mathcal{V}$; that is, $D\psi(q) : \mathbb{R}^2 \to \mathbb{R}^3$ is one-to-one for all $q \in \mathcal{V}$.

Remark B.91. Write $\psi : \mathcal{V} \to \mathcal{U}$ as

$$\psi(u,v) = \big(x(u,v), y(u,v), z(u,v)\big) \, .$$

Then if $q = (u_0, v_0)$,

$$(D\psi)(q) = \begin{bmatrix} x_u(u_0, v_0) & x_v(u_0, v_0) \\ y_u(u_0, v_0) & y_v(u_0, v_0) \\ z_u(u_0, v_0) & z_v(u_0, v_0) \end{bmatrix}.$$

The injectivity of $D\psi(q)$ is then translated to that the two vectors

$$\psi_u(u_0, v_0) = \big(x_u(u_0, v_0), y_u(u_0, v_0), z_u(u_0, v_0)\big)$$
$$\psi_v(u_0, v_0) = \big(x_v(u_0, v_0), y_v(u_0, v_0), z_v(u_0, v_0)\big)$$

are linearly independent.

Example B.92. Let $\mathbb{S}^2 = \big\{(x, y, z) \,\big|\, x^2 + y^2 + z^2 = 1\big\}$ be the unit sphere. If $p = (x_0, y_0, z_0) \in \mathbb{S}^2$, then either x_0, y_0 or z_0 is non-zero. Suppose that $z_0 \neq 0$. Choose $r > 0$ such that $(x - x_0)^2 + (y - y_0)^2 < 1$. Define

$$\psi(x, y) = \begin{cases} (x, y, \sqrt{1 - x^2 - y^2}) & \text{if } z_0 > 0\,, \\ (x, y, -\sqrt{1 - x^2 - y^2}) & \text{if } z_0 < 0\,, \end{cases}$$

$\mathcal{V} = B\big((x_0, y_0), r\big)$, and $\mathcal{U} = \psi(\mathcal{V})$. Then $\psi : \mathcal{V} \to \mathcal{U}$ is a bijection. Let $\varphi = \psi^{-1}$. Then $\{\mathcal{U}, \varphi\}$ is a coordinate chart at p; thus $\mathbb{S}^2$ is a surface.

There exists another coordinate chart. Let $\mathcal{U}_1 = \mathbb{S}^2 \backslash (0, 0, -1)$ and $\mathcal{U}_2 = \mathbb{S}^2 \backslash (0, 0, 1)$. Define the map $\varphi_1 : \mathcal{U}_1 \to \mathbb{R}^2$ by that $\varphi_1(p)$ is the unique point on $\mathbb{R}^2$ such that $(0, 0, -1)$, $\varphi_1(p)$ and $(x, y, 0)$ are on the same straight line. Similarly, define $\varphi_2 : \mathcal{U}_2 \to \mathbb{R}^2$ by that $\varphi_2(p)$ is the unique point on $\mathbb{R}^2$ such that $(0, 0, 1)$, $\varphi_2(p)$ and $(x, y, 0)$ are on the same straight line. It is easy to check that if $p \in \mathbb{S}^2$, then either $\{\mathcal{U}_1, \varphi_1\}$ or $\{\mathcal{U}_2, \varphi_2\}$ is a coordinate chart at p.

A third kind of coordinate chart is given as follows. Let $\mathcal{U} = (0, 2\pi) \times (0, \pi)$, and define

$$\psi(\theta, \phi) = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi).$$

Then $\psi : \mathcal{U} \to \mathbb{S}^2 \backslash \{(x, 0, z) \,|\, 0 \leqslant x \leqslant 1, x^2 + z^2 = 1\}$ is a continuous bijection with a continuous inverse. We note that for any $\mathcal{U} = (\theta_0, \theta_0 + 2\pi) \times (\phi_0, \phi_0 + \pi)$, ψ is a homeomorphism between $\mathcal{U}$ and an open subset of $\mathbb{S}^2$.

The metric tensor

DEFINITION B.93 (Metric)**.** Let $\mathcal{S} \subseteq \mathbb{R}^3$ be a regular surface. The metric tensor associated to the local parametrization $\{\mathcal{V}, \psi\}$ (at $p \in \mathcal{S}$) is the matrix $g = [g_{\alpha\beta}]_{2\times 2}$ given by

$$g_{\alpha\beta} = \psi,_\alpha \cdot \psi,_\beta = \sum_{i=1}^{3} \frac{\partial \psi^i}{\partial y_\alpha} \frac{\partial \psi^i}{\partial y_\beta} \qquad \text{in} \quad \mathcal{V}\,.$$

PROPOSITION B.94. *Let $\mathcal{S} \subseteq \mathbb{R}^3$ be a regular surface, and $g = [g_{\alpha\beta}]_{2\times 2}$ be the metric tensor associated to the local parametrization $\{\mathcal{V}, \psi\}$ (at $p \in \mathcal{S}$) such that $D\psi$ has full rank on $\mathcal{V}$. Then the metric tensor g is positive definite; that is,*

$$\sum_{\alpha,\beta=1}^{2} g_{\alpha\beta} v^\alpha v^\beta > 0 \qquad \forall\, \boldsymbol{v} = \sum_{\gamma=1}^{2} v^\gamma \frac{\partial \psi}{\partial y^\gamma} \neq 0\,.$$

Proof. Since $D\psi$ has full rank on $\mathcal{V}$, every tangent vector $\boldsymbol{v}$ can be expressed as the linear combination of $\left\{\frac{\partial \psi}{\partial y_1}, \frac{\partial \psi}{\partial y_2}\right\}$. Write $\boldsymbol{v} = \sum_{\gamma=1}^{2} v^\gamma \frac{\partial \psi}{\partial y^\gamma}$. Then if $\boldsymbol{v} \neq \mathbf{0}$,

$$0 < \|\boldsymbol{v}\|^2 = \sum_{i=1}^{3} \sum_{\alpha,\beta=1}^{2} v^\alpha \frac{\partial \psi^i}{\partial y_\alpha} v^\beta \frac{\partial \psi^i}{\partial \psi_\beta} = \sum_{\alpha,\beta=1}^{2} g_{\alpha\beta} v^\alpha v^\beta\,. \qquad \square$$

The first fundamental form

DEFINITION B.95 (The first fundamental form)**.** Let $\mathcal{S} \subseteq \mathbb{R}^3$ be a regular surface. The first fundamental form associated to the local parametrization $\{\mathcal{V}, \psi\}$ (at $p \in \mathcal{S}$) is the scalar function $\mathrm{g} = \det(g)$.

THEOREM B.96. *Let $\mathcal{S} \subseteq \mathbb{R}^3$ be a regular surface, and $\{\mathcal{V}, \psi\}$ be a local parametrization at $p \in \mathcal{S}$. Then*

$$\sqrt{\mathrm{g}} = \left|\psi,_1 \times \psi,_2\right|. \tag{B.7}$$

Proof. Using the permutation symbol and Kronecker's delta, we have

$$\begin{aligned}
\left|\psi,_1 \times \psi,_2\right|^2 &= \sum_{i=1}^{3} \Big(\sum_{j,k=1}^{3} \varepsilon_{ijk} \psi^j,_1 \psi^k,_2 \Big) \Big(\sum_{r,s=1}^{3} \varepsilon_{irs} \psi^r,_1 \psi^s,_2 \Big) \\
&= \sum_{j,k,r,s=1}^{3} \Big[\Big(\sum_{i=1}^{3} \varepsilon_{ijk} \varepsilon_{irs} \Big) \psi^j,_1 \psi^k,_2 \psi^r,_1 \psi^s,_2 \Big] \\
&= \sum_{j,k,r,s=1}^{3} \big(\delta_{jr}\delta_{ks} - \delta_{js}\delta_{kr} \big) \psi^j,_1 \psi^k,_2 \psi^r,_1 \psi^s,_2 \,,
\end{aligned}$$

where we use the identity

$$\sum_{i=1}^{3} \varepsilon_{ijk}\varepsilon_{irs} = \delta_{jr}\delta_{ks} - \delta_{js}\delta_{kr} \tag{B.8}$$

to conclude the last equality. Therefore,

$$\begin{aligned} \left|\psi,_1 \times \psi,_2\right|^2 &= \sum_{j,k=1}^{3} \left(\psi^j,_1 \psi^k,_2 \psi^j,_1 \psi^k,_2 - \psi^j,_1 \psi^k,_2 \psi^j,_2 \psi^k,_1\right) \\ &= g_{11}g_{22} - g_{12}g_{21} = \det(g) = \mathrm{g}\,. \end{aligned}$$

Finally, (B.7) is concluded from the fact that g is positive definite. □

The surface element and the surface integral

Not yet completed!!!

The second fundamental form

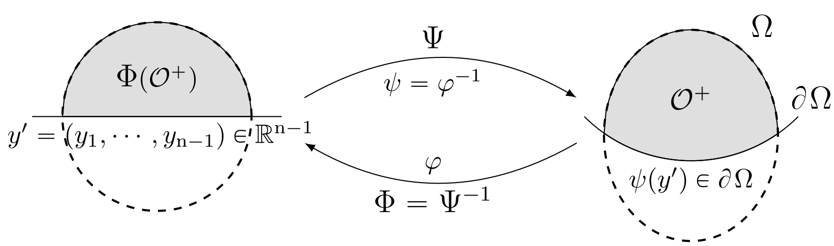
Not yet completed!!!

Some useful identities

Let $\mathcal{S} \subseteq \mathbb{R}^{\mathrm{n}}$ be an oriented surface, $\psi : \mathcal{U} \subseteq \mathbb{R}^{\mathrm{n}-1} \to \mathcal{S}$ be a parametrization of $\mathcal{S}$, and $\mathbf{N} : \mathcal{S} \to \mathbb{R}^{\mathrm{n}}$ be the normal vector on $\mathcal{S}$ 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 $\Psi : \mathcal{U} \times (-\varepsilon, \varepsilon) \to \mathcal{T}$ for some tubular neighborhood $\mathcal{T}$ of $\mathcal{S}$.



Denote $\psi,_\alpha \cdot \psi,_\beta$ by $g_{\alpha,\beta}$, and $\det([g_{\alpha\beta}])$ by g. Letting $\widetilde{\mathbf{N}} = \mathbf{N}\circ\psi$, we find that

$$\nabla\Psi\big|_{\{y_n=0\}} = \left[\psi,_1 \vdots \psi,_2 \vdots \cdots \vdots \psi,_{\mathrm{n}-1} \vdots \widetilde{\mathbf{N}}\right];$$

thus by Corollary B.64 and B.65,

$$\det(\nabla\Psi)^2\big|_{\{y_{\rm n}=0\}} = \Big[\det\big((\nabla\Psi)^{\rm T}\big)\det(\nabla\Psi)\Big]\Big|_{\{y_{\rm n}=0\}} = \det\big((\nabla\Psi)^{\rm T}\nabla\Psi\big)\Big|_{\{y_{\rm n}=0\}}$$
$$= \det\left(\begin{bmatrix} g_{11} & g_{12} & \cdots & g_{({\rm n}-1)1} & 0 \\ g_{21} & g_{22} & \cdots & g_{({\rm n}-1)2} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ g_{({\rm n}-1)1} & g_{({\rm n}-1)2} & \cdots & g_{({\rm n}-1)({\rm n}-1)} & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}\right) = {\rm g}\,.$$

As a consequence,

$$\det(\nabla\Psi) = \sqrt{\rm g} \qquad \text{on} \quad \{y_{\rm n}=0\}\,. \tag{B.9}$$

Moreover,

$$\mathrm{A}\big|_{\{y_n=0\}} = (\nabla\psi)^{-1}\big|_{\{y_n=0\}} = \begin{bmatrix} \displaystyle\sum_{\alpha=1}^{\rm n-1} g^{1\alpha}\psi_{,\alpha} \\ \cdots\cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots\cdots \\ \displaystyle\sum_{\alpha=1}^{\rm n-1} g^{({\rm n}-1)\alpha}\psi_{,\alpha} \\ \cdots\cdots\cdots\cdots \\ \widetilde{\mathbf{N}} \end{bmatrix}$$

in which $\big[g^{\alpha\beta}\big]_{({\rm n}-1)\times({\rm n}-1)}$ is the inverse matrix of $\big[g_{\alpha\beta}\big]_{({\rm n}-1)\times({\rm n}-1)}$. Therefore, writing J for $\det(\nabla\psi)$, we have

$$(\mathrm{J}\mathrm{A}^{\rm T}\mathrm{e}_{\rm n})\big|_{\{y_{\rm n}=0\}} = \sqrt{\rm g}\,\widetilde{\mathbf{N}}\,. \tag{B.10}$$

B.2.3 The divergence theorem

Two differential operators play important roles in vector calculus. The first one is called the **curl operator** which measures the circulation (the speed of rotation) of a vector field, and the second one is called the **divergence operator** which measures the flux of a vector field. We will study this two operators in the following two sections.

Measurements of the flux - the divergence operator

Not yet completed!!!

Definition B.97 (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 B.98. A vector field $\boldsymbol{u}$ is called *solenoidal* or *divergence-free* if $\operatorname{div}\boldsymbol{u} = 0$ in Ω.

The divergence theorem

Theorem B.99 (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. Let $\{\mathcal{U}_m\}_{m=0}^K$ denote an open cover of Ω such that $\{\mathcal{U}_m\}_{m=1}^K$ intersects the boundary $\partial\Omega$, and let $\{\theta_m\}_{m=1}^K$ denote a collection of charts such that

1. $\theta_m : B(0, r_m) \to \mathcal{U}_m$ is a $\mathscr{C}^{1,1}$-diffeomorphism,
2. $\theta_m : B(0, r_m) \cap \{x_n = 0\} \to \mathcal{U}_m \cap \partial\Omega$,
3. $\theta_m : B_m^+ \equiv B(0, r_m) \cap \{y_{\mathrm{n}} > 0\} \to \mathcal{U}_m \cap \Omega$.
4. The orientation of $\mathbf{N}$ is compatible with θ_m; that is,

$$\det\left(\left[\theta_{m,1} \vdots \cdots \vdots \theta_{m,\mathrm{n}-1} \vdots \tilde{\mathbf{N}}\right]\right) > 0\,.$$

Let $0 \leqslant \zeta_m \leqslant 1$ in $\mathscr{C}_c^\infty(\mathcal{U}_m)$ denote a partition of unity subordinate to the open covering $\mathcal{U}_m$; that is,

$$\sum_{m=0}^K \zeta_m = 1 \quad \text{and} \quad \operatorname{spt}(\zeta_m) \subseteq \mathcal{U}_m \quad \forall\, m\,.$$

Define $\mathrm{J}_m = \det(\nabla\theta_m)$ and $\mathrm{A}_m = (\nabla\theta_m)^{-1}$. Using (B.10), $\sqrt{g_m}(\mathbf{N}\circ\theta_m) = -\mathrm{J}_m(\mathrm{A}_m)^{\mathrm{T}}\mathrm{e}_{\mathrm{n}}$ for all $m = 1, \cdots, K$. We note that the minus sign is due to the fact that $\theta_m : B_m^+ \to \Omega$, while $\mathbf{N}$ is the outward-pointing unit normal of $\partial\Omega$. Therefore, making change of variable $x = \theta_m(y)$ in each $\mathcal{U}_m$ we find that

$$\begin{aligned}
\int_{\partial\Omega} \boldsymbol{v} \cdot \mathbf{N}\, dS &= \sum_{m=0}^{K} \int_{\partial\Omega\cap\mathcal{U}_m} \zeta_m(\boldsymbol{v} \cdot \mathbf{N})\, dS \\
&= \sum_{m=0}^{K} \sum_{i=1}^{\mathrm{n}} \int_{\mathbb{R}^{\mathrm{n}-1}\times\{y_{\mathrm{n}}=0\}} (\zeta_m \circ \theta_m)(\boldsymbol{v}^i \circ \theta_m)(\mathbf{N}^i \circ \theta_m)\sqrt{g_m} dy_1 \cdots dy_{\mathrm{n}-1} \\
&= -\sum_{m=0}^{K} \sum_{i=1}^{\mathrm{n}} \int_{\mathbb{R}^{\mathrm{n}-1}\times\{y_{\mathrm{n}}=0\}} (\zeta_m \circ \theta_m)(\boldsymbol{v}^i \circ \theta_m)\mathrm{J}_m(\mathrm{A}_m)^{\mathrm{n}}_i dy_1 \cdots dy_{\mathrm{n}-1} \\
&= \sum_{m=0}^{K} \sum_{i=1}^{\mathrm{n}} \int_{B_m^+} \frac{\partial}{\partial y_{\mathrm{n}}}\big[(\zeta_m \circ \theta_m)\mathrm{J}_m(\mathrm{A}_m)^{\mathrm{n}}_i(\boldsymbol{v}^i \circ \theta_m)\big] dy\,.
\end{aligned}$$

On the other hand, for $\alpha = 1, \cdots, \mathrm{n}-1$ and $i = 1, \cdots, \mathrm{n}$,

$$\int_{B_m^+} \frac{\partial}{\partial y_\alpha}\big[(\zeta_m \circ \theta_m)\mathrm{J}_m(\mathrm{A}_m)^{\alpha}_i(\boldsymbol{v}^i \circ \theta_m)\big] dy = 0\,;$$

thus the Piola identity implies that

$$\begin{aligned}
\int_{\partial\Omega} \boldsymbol{v} \cdot \mathbf{N}\, dS &= \sum_{m=0}^{K} \sum_{i,j=1}^{\mathrm{n}} \int_{B_m^+} \frac{\partial}{\partial y_j}\big[(\zeta_m \circ \theta_m)\mathrm{J}_m(\mathrm{A}_m)^{j}_i(\boldsymbol{v}^i \circ \theta_m)\big] dy \\
&= \sum_{m=0}^{K} \sum_{i,j=1}^{\mathrm{n}} \int_{B_m^+} \mathrm{J}_m(\mathrm{A}_m)^{j}_i(\zeta_m \circ \theta_m),_j (\boldsymbol{v}^i \circ \theta_m) dy \\
&\quad + \sum_{m=0}^{K} \sum_{i,j=1}^{\mathrm{n}} \int_{B_m^+} (\zeta_m \circ \theta_m)\mathrm{J}_m(\mathrm{A}_m)^{j}_i(\boldsymbol{v}^i \circ \theta_m),_j\, dy\,.
\end{aligned}$$

Making change of variable $x = \theta_m(y)$ in each $\theta_m(\mathcal{U}_m) \subseteq \mathbb{R}^{\mathrm{n}}_+$ again, we conclude that

$$\begin{aligned}
\int_{\partial\Omega} \boldsymbol{v} \cdot \mathbf{N}\, dS &= \sum_{m=0}^{K} \int_{\mathcal{U}_m} (\boldsymbol{v} \cdot \nabla_x)\zeta_m\, dx + \sum_{m=0}^{K} \int_{\mathcal{U}_m} \zeta_m \mathrm{div}\boldsymbol{v}\, dx \\
&= \int_{\Omega} (\boldsymbol{v} \cdot \nabla_x)1\, dx + \int_{\Omega} \mathrm{div}\boldsymbol{v}\, dx = \int_{\Omega} \mathrm{div}\boldsymbol{v}\, dx\,. \qquad \square
\end{aligned}$$

Letting $\boldsymbol{v} = (0, \cdots, 0, f, 0, \cdots, 0) = f\mathrm{e}_i$, we obtain the following

COROLLARY B.100. *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 B.99, we conclude the following

Corollary B.101. *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*

$$\int_\Omega \varphi \operatorname{div}\boldsymbol{v}\, dx = \int_{\partial\Omega} (\boldsymbol{v}\cdot\mathbf{N})\varphi\, dS - \int_\Omega \boldsymbol{v}\cdot\nabla\varphi\, dx\,, \tag{B.11}$$

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 $\mathbb{R}^{\mathrm{n}}$. 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 B.102. Let $\mathcal{S} \subseteq \mathbb{R}^3$ be an open $\mathscr{C}^1$-surface; that is, $\mathcal{S}$ is of class $\mathscr{C}^1$ and $\mathcal{S} \cap \partial\mathcal{S} = \varnothing$. A vector field $\boldsymbol{u}$ defined on $\mathcal{S}$ is called a tangent vector field on $\mathcal{S}$, denoted by $\boldsymbol{u} \in \mathbf{T}(\mathcal{S})$, if $\boldsymbol{u}\cdot\mathbf{N} = 0$ on $\mathcal{S}$, where $\mathbf{N} : \mathcal{S} \to \mathbb{S}^2$ is a unit normal vector field on $\mathcal{S}$.

Having established (B.11), we find that the divergence operator div is the formal adjoint of the operator $-\nabla$. The following definition is motivated by this observation.

Definition B.103 (The surface gradient and the surface divergence)**.** Let $\mathcal{S} \subseteq \mathbb{R}^{\mathrm{n}}$ be a $\mathscr{C}^1$-surface. We let $\nabla_{\mathcal{S}}$ denote the tangential gradient on $\mathcal{S}$. If $\varphi : \mathcal{S} \to \mathbb{R}$ is differentiable, then in local chart $(\mathcal{U},\theta)$, $\nabla_{\mathcal{S}}\varphi$ is given by

$$(\nabla_{\mathcal{S}}\varphi)\circ\theta = \sum_{\alpha,\beta=1}^{\mathrm{n}-1} g^{\alpha\beta}\frac{\partial(\varphi\circ\theta)}{\partial y_\alpha}\frac{\partial\theta}{\partial y_\beta}\,,$$

where $[g^{\alpha\beta}]$ is the inverse matrix of the induced metric $[g_{\alpha\beta}]$, and $\left\{\dfrac{\partial\theta}{\partial y_\beta}\right\}_{\beta=1}^2$ are tangent vectors to $\mathcal{S}$.

The surface divergence operator $\operatorname{div}_{\mathcal{S}}$ is defined as the formal adjoint of $-\nabla_{\mathcal{S}}$; that is, if $\boldsymbol{u} \in \mathbf{T}(\mathcal{S})$, then

$$-\int_{\mathcal{S}} \boldsymbol{u}\cdot\nabla_{\mathcal{S}}\varphi\, dS = \int_{\mathcal{S}} \varphi \operatorname{div}_{\mathcal{S}}\boldsymbol{u}\, dS \qquad \forall\,\varphi \in \mathscr{C}_c^1(\mathcal{S})\,.$$

In a local chart $(\mathcal{U}, \theta)$,

$$(\operatorname{div}_{\mathcal{S}} \boldsymbol{u}) \circ \theta = \frac{1}{\sqrt{\mathrm{g}}} \sum_{\alpha,\beta=1}^{\mathrm{n}-1} \frac{\partial}{\partial y_\alpha} \Big[\sqrt{\mathrm{g}} g^{\alpha\beta} \big((\boldsymbol{u} \circ \theta) \cdot \frac{\partial \theta}{\partial y_\beta} \big) \Big] ,$$

where $\mathrm{g} = \det(g)$ is the determinant of the induced metric $[g_{\alpha\beta}]$.

Now we are in the position of stating the divergence theorem on surfaces with boundary.

THEOREM B.104. *Let $\mathcal{S} \subseteq \mathbb{R}^3$ be an oriented $\mathscr{C}^1$-surface with $\mathscr{C}^1$-boundary $\partial\mathcal{S}$, $\mathbf{N} : \mathcal{S} \to \mathbb{S}^2$ be normal vector on $\mathcal{S}$, and $\mathbf{T} : \partial\mathcal{S} \to \mathbb{S}^2$ be tangent vector on $\partial\mathcal{S}$ such that $\mathbf{T}$ is compatible with $\mathbf{N}$. Then*

$$\int_{\partial\mathcal{S}} \boldsymbol{u} \cdot (\mathbf{T} \times \mathbf{N}) ds = \int_{\mathcal{S}} \operatorname{div}_{\mathcal{S}} \boldsymbol{u} \, dS \qquad \forall \, \boldsymbol{u} \in \mathbf{T}(\mathcal{S}) \cap \mathscr{C}^1(\mathcal{S}; \mathbb{R}^3) ,$$

where $\operatorname{div}_{\mathcal{S}}$ is the surface divergence operator.

Proof. Let $\{\mathcal{U}_m\}_{m=1}^K$ denote an open cover of $\bar{\mathcal{S}}$ such that $\mathcal{U}_m \cap \partial\mathcal{S} = \varnothing$ for $1 \leqslant m \leqslant J$, and $\mathcal{U}_m \cap \partial\mathcal{S}$ is non-empty and connected for $J+1 \leqslant m \leqslant K$. Let $\{\theta_m\}_{m=1}^K$ denote a collection of charts such that

1. $\theta_m : B(0, r_m) \subseteq \mathbb{R}^2 \to \mathcal{U}_m \cap \mathcal{S}$ is a $\mathscr{C}^1$-diffeomorphism for each $1 \leqslant m \leqslant K$,
2. $\theta_m : B(0, r_m) \cap \{y_2 = 0\} \to \mathcal{U}_m \cap \partial\mathcal{S}$ for $J+1 \leqslant m \leqslant K$,
3. $\theta_m : B_m^+ \equiv B(0, r_m) \cap \{y_2 > 0\} \to \mathcal{U}_m \cap \mathcal{S}$ for $J+1 \leqslant m \leqslant K$,

and $\{g_m\}_{m=1}^K$ 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 subordinate to $\{\mathcal{U}_m\}_{m=1}^K$. Then

$$\begin{aligned}
\int_{\mathcal{S}} \operatorname{div}_{\mathcal{S}} \boldsymbol{u} \, dS &= \sum_{m=1}^K \int_{\mathcal{U}_m \cap \mathcal{S}} \zeta_m \operatorname{div}_{\mathcal{S}} \boldsymbol{u} \, dS \\
&= \sum_{m=1}^J \sum_{\alpha,\beta=1}^2 \int_{B(0,r_m)} (\zeta_m \circ \theta_m) \frac{\partial}{\partial y_\alpha} \Big[\sqrt{\mathrm{g}_m} g_m^{\alpha\beta} \big((\boldsymbol{u} \circ \theta_m) \cdot \frac{\partial \theta_m}{\partial y_\beta} \big) \Big] dy \\
&\quad + \sum_{m=J+1}^K \sum_{\alpha,\beta=1}^2 \int_{B_m^+} (\zeta_m \circ \theta_m) \frac{\partial}{\partial y_\alpha} \Big[\sqrt{\mathrm{g}_m} g_m^{\alpha\beta} \big((\boldsymbol{u} \circ \theta_m) \cdot \frac{\partial \theta_m}{\partial y_\beta} \big) \Big] dy .
\end{aligned}$$

Let $\mathbf{n}$ denote the outward-pointing unit normal on either $\partial B(0,r_m)$ for $1 \leqslant m \leqslant J$ or ∂B_m^+ for $J+1 \leqslant m \leqslant K$. Since $\zeta_m \circ \theta_m = 0$ on $\partial B(0,r_m)$ for $1 \leqslant m \leqslant J$, and $\zeta_m \circ \theta_m = 0$ 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

$$\begin{aligned}
&\int_{\mathcal{S}} \operatorname{div}_{\mathcal{S}} \boldsymbol{u}\, dS \\
&\quad = - \sum_{\alpha,\beta=1}^{2} \int_{B(0,r_m)} \Big[\sqrt{\mathrm{g}_m} g_m^{\alpha\beta}\big((\boldsymbol{u}\circ\theta_m)\cdot \frac{\partial \theta_m}{\partial y_\beta}\big)\Big] \frac{\partial}{\partial y_\alpha} \sum_{m=1}^{K} (\zeta_m \circ \theta_m) dy \\
&\qquad + \sum_{m=J+1}^{K} \sum_{\alpha,\beta=1}^{2} \int_{B(0,r_m)\cap\{y_2=0\}} (\zeta_m\circ\theta_m)\mathbf{n}_\alpha \Big[\sqrt{\mathrm{g}_m} g_m^{\alpha\beta}\big((\boldsymbol{u}\circ\theta_m)\cdot \frac{\partial \theta_m}{\partial y_\beta}\big)\Big] dy_1 \\
&\quad = \sum_{m=J+1}^{K} \int_{B(0,r_m)\cap\{y_2=0\}} (\zeta_m\circ\theta_m)(\boldsymbol{u}\circ\theta_m)\cdot \Big[\sum_{\alpha,\beta=1}^{2} \mathbf{n}_\alpha \sqrt{\mathrm{g}_m} g_m^{\alpha\beta} \frac{\partial \theta_m}{\partial y_\beta}\Big] dy_1 \,.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\int_{\partial\mathcal{S}} \boldsymbol{u}\cdot(\mathbf{T}\times\mathbf{N}) ds &= \sum_{m=J+1}^{K} \int_{\partial\mathcal{S}\cap\mathcal{U}_m} \zeta_m \boldsymbol{u}\cdot(\mathbf{T}\times\mathbf{N}) ds \\
&= \sum_{m=J+1}^{K} \int_{B(0,r_m)\cap\{y_2=0\}} (\zeta_m\circ\theta_m)(\boldsymbol{u}\circ\theta_m)\cdot \Big[(\mathbf{T}\times\mathbf{N})\circ\theta_m \Big|\frac{\partial\theta_m}{\partial y_1}\Big|\Big] dy_1 \,.
\end{aligned}$$

Therefore, the theorem can be concluded as long as we can show that

$$\sum_{\alpha,\beta=1}^{2} \mathbf{n}_\alpha \sqrt{\mathrm{g}_m}\, g_m^{\alpha\beta} \frac{\partial \theta_m}{\partial y_\beta} = (\mathbf{T}\times\mathbf{N})\circ\theta_m \Big|\frac{\partial\theta_m}{\partial y_1}\Big| \quad \text{on} \quad B(0,r_m)\cap\{y_2=0\}\,. \tag{B.12}$$

Let $\boldsymbol{\tau}_m = \sum_{\alpha,\beta=1}^{2} \mathbf{n}_\alpha \sqrt{\mathrm{g}_m} g_m^{\alpha\beta} \frac{\partial \theta_m}{\partial y_\beta}$ on $B(0,r_m)\cap\{y_2=0\}$. Since $\mathbf{n}_\alpha = -\delta_{2\alpha}$, we find that $\boldsymbol{\tau}_m \cdot \frac{\partial\theta_m}{\partial y_1} = 0$ on $B(0,r_m)\cap\{y_2=0\}$; thus

$$\boldsymbol{\tau}_m\cdot(\mathbf{T}\circ\theta_m) = 0 \qquad \text{on} \quad B(0,r_m)\cap\{y_2=0\}\,.$$

Moreover, noting that $\boldsymbol{\tau}_m$ is a linear combination of tangent vectors $\frac{\partial\theta_m}{\partial y_\beta}$, we must have

$$\boldsymbol{\tau}_m\cdot(\mathbf{N}\circ\theta_m) = 0 \qquad \text{on} \quad B(0,r_m)\cap\{y_2=0\}\,.$$

As a consequence,

$$\boldsymbol{\tau}_m \,/\!\!/\, (\mathbf{T}\times\mathbf{N})\circ\theta_m \qquad \text{on} \quad B(0,r_m)\cap\{y_2=0\}\,.$$

Since $(\mathbf{T} \times \mathbf{N})$ points away from $\mathcal{S}$, while $\frac{\partial \theta_m}{\partial y_2} \circ \theta_m^{-1}\Big|_{\partial \mathcal{S}}$ points toward $\mathcal{S}$, by the fact that

$$\boldsymbol{\tau}_m \cdot \frac{\partial \theta_m}{\partial y_2} = \sum_{\alpha,\beta=1}^{2} \mathbf{n}_\alpha \sqrt{\mathrm{g}_m} g_m^{\alpha\beta} \frac{\partial \theta_m}{\partial y_\beta} \cdot \frac{\partial \theta_m}{\partial y_2} = -\sqrt{\mathrm{g}_m} g_m^{22} < 0\,,$$

we must have $\boldsymbol{\tau}_m \cdot (\mathbf{T} \times \mathbf{N}) \circ \theta_m > 0$ on $B(0, r_m) \cap \{y_2 = 0\}$. In other words,

$$\boldsymbol{\tau}_m = |\boldsymbol{\tau}_m| (\mathbf{T} \times \mathbf{N}) \circ \theta_m \qquad \text{on} \quad B(0, r_m) \cap \{y_2 = 0\}\,.$$

Finally, since

$$\boldsymbol{\tau}_m \cdot \boldsymbol{\tau}_m = \sum_{\alpha,\beta,\gamma,\delta=1}^{2} \mathrm{g}_m \, \mathbf{n}_\alpha \, \mathbf{n}_\gamma \, g_m^{\alpha\beta} g_m^{\gamma\delta} \, \frac{\partial \theta_m}{\partial y_\beta} \cdot \frac{\partial \theta_m}{\partial y_\delta} = \mathrm{g}_m g_m^{22} = g_{m11} = \Big| \frac{\partial \theta_m}{\partial y_1} \Big|^2,$$

we conclude that $\boldsymbol{\tau}_m = \Big| \frac{\partial \theta_m}{\partial y_1} \Big| (\mathbf{T} \times \mathbf{N}) \circ \theta_m$ on $\{y_2 = 0\}$; thus (B.12) is established. □

REMARK B.105. On $\partial \mathcal{S}$, the vector $\mathbf{T} \times \mathbf{N}$ is "tangent" to $\mathcal{S}$ and points away from $\mathcal{S}$. In other words, $\mathbf{T} \times \mathbf{N}$ can be treated as the "outward-pointing" unit "normal" of $\partial \mathcal{S}$ which makes the divergence theorem on surfaces more intuitive.

B.2.4 The Stokes theorem

Measurements of the circulation - the curl operator

We consider the circulation or the speed of rotation of a vector field u about an axis in the direction $\mathbf{N}$. Let P be a plane passing thorough a point a and having normal $\mathbf{N}$, and C_r be a circle on the plane P centeblack at a with radius r. Pick the orientation of the unit tangent vector $\mathbf{T}$ which is compatible with the unit normal $\mathbf{N}$ (see Figure B.2 for reference).

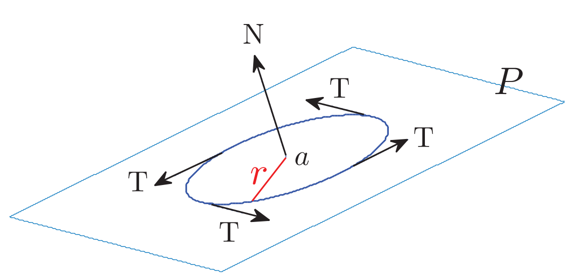


Figure B.2: the circulation about an axis in direction **N**

Since the instantaneous angular velocity of a vector field u along the circle C_r is measublack by $\dfrac{\boldsymbol{u}\cdot\mathbf{T}}{r}$, it is quite reasonable to measure the circulation of u along C_r by averaging the angular velocity; that is, we consider the quantity

$$\frac{1}{2\pi r}\oint_{C_r}\frac{\boldsymbol{u}\cdot\mathbf{T}}{r}\,ds \tag{B.13}$$

as a (constant multiple of) measurement of the speed of rotation. The limit of the quantity above, as $r\to 0$, is then a good measurement of the rotation speed of $\boldsymbol{u}$ at the point a about the axis in the direction $\mathbf{N}$.

Since we expect that this measurement does not depend on the choice of coordinate systems, we start from letting P be the x_1x_2-plane, and $\mathbf{N}=(0,0,1)$, $\mathbf{T}=(-\sin\theta,\cos\theta,0)$. By the change of variable $ds=rd\theta$ and the L'Hôspital rule,

$$\begin{aligned}
&\lim_{r\to 0}\frac{1}{2\pi r}\oint_{C_r}\frac{\boldsymbol{u}\cdot\mathbf{T}}{r}\,ds\\
&=\lim_{r\to 0}\int_0^{2\pi}\frac{\boldsymbol{u}^2\big(a+(r\cos\theta,r\sin\theta,0)\big)\cos\theta-\boldsymbol{u}^1\big(a+(r\cos\theta,r\sin\theta,0)\big)\sin\theta}{2\pi r}\,d\theta\\
&=\frac{1}{2\pi}\int_0^{2\pi}\Big[\boldsymbol{u}^2_{,1}(a)\cos^2\theta+\boldsymbol{u}^2_{,2}(a)\cos\theta\sin\theta-\boldsymbol{u}^1_{,1}(a)\cos\theta\sin\theta-\boldsymbol{u}^1_{,2}(a)\sin^2\theta\Big]d\theta\\
&=\frac{1}{2}\big[\boldsymbol{u}^2_{,1}(a)-\boldsymbol{u}^1_{,2}(a)\big]=\frac{1}{2}\sum_{i,j=1}^{2}\varepsilon_{3ij}\boldsymbol{u}^j_{,i}(a).
\end{aligned}\tag{B.14}$$

Now suppose the general case that $\mathbf{N}\neq e_3$. There is an orthonormal matrix $\mathrm{O}=[\widehat{\mathrm{e}}_1|\widehat{\mathrm{e}}_2|\widehat{\mathrm{e}}_3]$ so that $\mathrm{O}\mathrm{e}_3=\mathbf{N}$. As a consequence, $\widehat{\mathrm{e}}_3=\mathbf{N}$, $\widehat{\mathrm{e}}_j=\mathrm{O}\mathrm{e}_j$, $\mathbf{T}=\mathrm{O}\boldsymbol{\tau}$ with $\boldsymbol{\tau}=(-\sin\theta,\cos\theta,0)$, and the limit of the quantity in (B.13) is given by

$$\begin{aligned}
&\lim_{r\to 0}\frac{1}{2\pi r}\int_0^{2\pi}\boldsymbol{u}(a+r\cos\theta\widehat{\mathrm{e}}_1+r\sin\theta\widehat{\mathrm{e}}_2)\cdot(\mathrm{O}\boldsymbol{\tau})\,d\theta\\
&\qquad=\lim_{r\to 0}\frac{1}{2\pi r}\int_0^{2\pi}(\mathrm{O}\boldsymbol{v})(a+r\cos\theta\widehat{\mathrm{e}}_1+r\sin\theta\widehat{\mathrm{e}}_2)\cdot(\mathrm{O}\boldsymbol{\tau})\,d\theta\\
&\qquad=\lim_{r\to 0}\frac{1}{2\pi r}\int_0^{2\pi}\boldsymbol{v}(a+r\cos\theta\widehat{\mathrm{e}}_1+r\sin\theta\widehat{\mathrm{e}}_2)\cdot(-\sin\theta,\cos\theta,0)\,d\theta,
\end{aligned}$$

where $\boldsymbol{v}=\mathrm{O}^{\mathbf{T}}\boldsymbol{u}$, and the identity that $(\mathrm{O}\boldsymbol{v})\cdot(\mathrm{O}\boldsymbol{\tau})=v\cdot\boldsymbol{\tau}$ is used to deduce the last equality. By the L'Hôspital rule again,

$$\lim_{r\to 0}\frac{1}{2\pi r}\oint_{C_r}\frac{\boldsymbol{u}\cdot\mathbf{T}}{r}\,ds=\frac{1}{2}\sum_{j=1}^{3}\big[\boldsymbol{v}^2_{,j}(a)\,\widehat{\mathrm{e}}^j_1-\boldsymbol{v}^1_{,j}(a)\,\widehat{\mathrm{e}}^j_2\big].$$

In fact, we expect this to hold since if using $x' = \mathrm{O}^{\mathbf{T}}x$ as the new coordinate, by (B.14) and the chain rule we obtain that

$$\lim_{r\to 0}\frac{1}{2\pi r}\oint_{C_r}\frac{\boldsymbol{u}\cdot\mathbf{T}}{r}\,ds = \frac{1}{2}\Big[\frac{\partial \boldsymbol{v}^2}{\partial x_1'} - \frac{\partial \boldsymbol{v}^1}{\partial x_2'}\Big](a') = \frac{1}{2}\sum_{j=1}^{3}\Big[\frac{\partial \boldsymbol{v}^2}{\partial x_j}(a)\,\widehat{\mathrm{e}}_1^j - \frac{\partial \boldsymbol{v}^1}{\partial x_j}(a)\,\widehat{\mathrm{e}}_2^j\Big].$$

Finally, we note that $\boldsymbol{v}^\ell_{,j} = \sum\limits_{k=1}^{3}\boldsymbol{u}^k_{,j}\,\mathrm{e}_k\cdot\widehat{\mathrm{e}}_\ell = \sum\limits_{k=1}^{3}\boldsymbol{u}^k_{,j}\,\widehat{\mathrm{e}}^k_\ell$ for $\ell = 1,2,3$; thus

$$\begin{aligned}\lim_{r\to 0}\frac{1}{2\pi r}\oint_{C_r}\frac{\boldsymbol{u}\cdot\mathbf{T}}{r}\,ds &= \frac{1}{2}\sum_{j=1}^{3}\big[\boldsymbol{v}^2_{,j}\,\widehat{\mathrm{e}}_1^j - \boldsymbol{v}^1_{,j}\,\widehat{\mathrm{e}}_2^j\big] = \frac{1}{2}\sum_{j,k=1}^{3}\boldsymbol{u}^k_{,j}\big[\widehat{\mathrm{e}}_2^k\,\widehat{\mathrm{e}}_1^j - \widehat{\mathrm{e}}_1^k\,\widehat{\mathrm{e}}_2^j\big]\\ &= \frac{1}{2}\sum_{j,k,r,s=1}^{3}(\delta_{jr}\delta_{ks} - \delta_{js}\delta_{kr})\,\boldsymbol{u}^k_{,j}\,\widehat{\mathrm{e}}_1^r\,\widehat{\mathrm{e}}_2^s\,,\end{aligned}$$

where $\delta_{\cdot\cdot}$'s are the Kronecker deltas. Due to the following useful identity

$$\sum_{i=1}^{3}\varepsilon_{ijk}\varepsilon_{irs} = \delta_{jr}\delta_{ks} - \delta_{js}\delta_{kr}\,, \tag{B.15}$$

we conclude that

$$\begin{aligned}\lim_{r\to 0}\frac{1}{2\pi r}\oint_{C_r}\frac{\boldsymbol{u}\cdot\mathbf{T}}{r}\,ds &= \frac{1}{2}\sum_{i,j,k,r,s=1}^{3}\varepsilon_{ijk}\varepsilon_{irs}\boldsymbol{u}^k_{,j}\,\widehat{\mathrm{e}}_1^r\,\widehat{\mathrm{e}}_2^s = \frac{1}{2}\sum_{i,j,k=1}^{3}\varepsilon_{ijk}\boldsymbol{u}^k_{,j}(\widehat{\mathrm{e}}_1\times\widehat{\mathrm{e}}_2)^i\\ &= \frac{1}{2}\sum_{i,j,k=1}^{3}\varepsilon_{ijk}\boldsymbol{u}^k_{,j}\,\widehat{\mathrm{e}}_3^i = \frac{1}{2}\sum_{i,j,k=1}^{3}\varepsilon_{ijk}\boldsymbol{u}^k_{,j}\mathbf{N}_i.\end{aligned}$$

This motivates the following

Definition B.106 (The curl operator)**.** Let $u:\Omega\subseteq\mathbb{R}^{\mathrm{n}}\to\mathbb{R}^{\mathrm{n}}$, $\mathrm{n}=2$ or $\mathrm{n}=3$, be a vector field.

1. For $\mathrm{n}=2$, the curl of u is a scalar function defined by

$$\mathrm{curl}\boldsymbol{u} = \sum_{i,j=1}^{2}\varepsilon_{3ij}\boldsymbol{u}^j_{,i}\,.$$

2. For $\mathrm{n}=3$, the curl of u is a vector-valued function defined by

$$(\mathrm{curl}\boldsymbol{u})^i = \sum_{j,k=1}^{3}\varepsilon_{ijk}\boldsymbol{u}^k_{,j}\,.$$

The function $\operatorname{curl}\boldsymbol{u}$ is also called the *vorticity* of u, and is usually denoted by one single Greek letter ω.

Having the curl operator defined, for the three-dimensional case the circulation of a vector field u on the plane with normal $\mathbf{N}$ is given by $\dfrac{\operatorname{curl}\boldsymbol{u}\cdot\mathbf{N}}{2}$.

The Stokes theorem

The path we choose to circle around the point a does not have to be a circle. However, in such a case the average of the angular velocity no longer makes sense (since $\boldsymbol{u}\cdot\mathbf{T}$ might not contribute to the motion in the angular direction), and we instead consider the limit of the following quantity

$$\lim_{\mathrm{A}\to 0}\frac{1}{\mathrm{A}}\oint_C \boldsymbol{u}\cdot\mathbf{T}ds,$$

where A is the area enclosed by C. This limit is always $\operatorname{curl}\boldsymbol{u}\cdot\mathbf{N}$ because of the famous Stokes' theorem.

Theorem B.107 (The Stokes theorem)**.** *Let $\boldsymbol{u}:\Omega\subseteq\mathbb{R}^3\to\mathbb{R}^3$ be a smooth vector field, and $\mathcal{S}$ be a $\mathscr{C}^1$-surface with $\mathscr{C}^1$-boundary $\partial\mathcal{S}$ in Ω. Then*

$$\int_{\partial\mathcal{S}}\boldsymbol{u}\cdot\mathbf{T}ds=\int_{\mathcal{S}}\operatorname{curl}\boldsymbol{u}\cdot\mathbf{N}\,dS\,,$$

where $\mathbf{N}$ and $\mathbf{T}$ are compatible normal and tangent vector fields.

To prove the Stokes theorem, we first establish the following

Lemma B.108. *Let $\Omega\subseteq\mathbb{R}^3$ be a bounded Lipschitz domain, and $\boldsymbol{w}:\Omega\to\mathbb{R}^{\mathrm{n}}$ be a mooth vector-valued function. If $\mathcal{S}\subseteq\Omega$ is an oriented $\mathscr{C}^1$-surface with normal $\mathbf{N}$, then*

$$\operatorname{curl}\boldsymbol{w}\cdot\mathbf{N}=\operatorname{div}_{\mathcal{S}}(\boldsymbol{w}\times\mathbf{N})\quad\text{on}\quad\mathcal{S}\,. \tag{B.16}$$

Proof. Let $\mathcal{O}\subseteq\Omega$ be a $\mathscr{C}^1$-domain such that $\mathcal{S}\subseteq\partial\mathcal{O}$. In other words, $\mathcal{S}$ is part of the boundary of $\mathcal{O}$. Since

$$(\nabla\varphi)^i=\frac{\partial\varphi}{\partial\mathbf{N}}\mathbf{N}^i+(\nabla_{\partial\mathcal{O}}\varphi)^i\quad\text{on}\quad\partial\mathcal{O}\,,$$

by the divergence theorem we conclude that for all $\varphi\in\mathscr{C}^1(\mathcal{O})$ such that $\varphi\big|_{\mathcal{S}}\in\mathscr{C}^1_c(\mathcal{S})$,

$$\begin{aligned}\int_{\partial\mathcal{O}}(\operatorname{curl}\boldsymbol{w}\cdot\mathbf{N})\varphi\,dS&=\int_{\mathcal{O}}\operatorname{curl}\boldsymbol{w}\cdot\nabla\varphi\,dx=\int_{\partial\mathcal{O}}(\mathbf{N}\times\boldsymbol{w})\cdot\nabla\varphi\,dS\\&=\int_{\partial\mathcal{O}}(\mathbf{N}\times\boldsymbol{w})\cdot\nabla_{\partial\mathcal{O}}\varphi\,dS=\int_{\partial\mathcal{O}}\operatorname{div}_{\partial\mathcal{O}}(\boldsymbol{w}\times\mathbf{N})\varphi\,dS\,.\end{aligned}$$

Identity (B.16) is concluded since φ can be chosen arbitrarily on $\mathcal{S}$. ◻

Proof of the Stokes theorem. Using (B.16) and then applying the divergence theorem on surfaces with boundary (Theorem B.104), we find that

$$\begin{aligned}\int_{\mathcal{S}} \operatorname{curl}\boldsymbol{u}\cdot\mathbf{N}\,dS &= \int_{\mathcal{S}} \operatorname{div}_{\mathcal{S}}(\boldsymbol{u}\times\mathbf{N})\,dS = \int_{\partial\mathcal{S}} (\boldsymbol{u}\times\mathbf{N})\cdot(\mathbf{T}\times\mathbf{N})ds \\ &= \int_{\partial\mathcal{S}} (\boldsymbol{u}\cdot\mathbf{T})ds\end{aligned}$$

in which the identity $(\boldsymbol{u}\times\mathbf{N})\cdot(\mathbf{T}\times\mathbf{N}) = \boldsymbol{u}\cdot\mathbf{T}$ is used. □

B.2.5 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 $\psi : \begin{cases} \Omega_1 \to \Omega_2 \\ \partial\Omega_1 \to \partial\Omega_2 \\ y \mapsto x = \psi(y) \end{cases}$ 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 $\mathrm{A} = (\nabla\psi)^{-1}$,

$$\frac{\partial f}{\partial x_i}\circ\psi = \sum_{j=1}^{\mathrm{n}} \mathrm{A}_i^j \frac{\partial F}{\partial y_j}\,. \tag{B.17}$$

Letting $\mathrm{J} = \det(\nabla\psi)$ be the Jacobian of ψ, by the change of variable $y = \psi(y)$ and the Piola identity,

$$\int_{\Omega_2} \frac{\partial f}{\partial x_i}(x)dx = \int_{\Omega_1} \frac{\partial f}{\partial x_i}\big(\psi(y)\big)\det(\nabla\psi)(y)dy = \sum_{j=1}^{\mathrm{n}}\int_{\Omega_1} \frac{\partial}{\partial y_j}(\mathrm{J}\mathrm{A}_i^j F)\,dy.$$

The divergence theorem again implies that

$$\int_{\Omega_2} \frac{\partial f}{\partial x_i}(x)dx = \sum_{j=1}^{\mathrm{n}}\int_{\Omega_1} \mathrm{J}\mathrm{A}_i^j F\mathbf{N}_j\,dS_y$$

which further implies that

$$\int_{\partial\Omega_2} (f\mathbf{n})(x)\, dS_x = \int_{\partial\Omega_1} F \frac{\mathrm{JA}^{\mathrm{T}}\mathbf{N}}{|\mathrm{JA}^{\mathrm{T}}\mathbf{N}|} |\mathrm{JA}^{\mathrm{T}}\mathbf{N}|\, dS_y\,. \tag{B.18}$$

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, (B.18) suggests that

$$\int_{\partial\Omega_2} f\mathbf{n}\, dS = \int_{\partial\Omega_1} \big[(f\mathbf{n})\circ\psi\big](y)\psi^*(dS_x) = \int_{\partial\Omega_1} (f\circ\psi)\frac{\mathrm{JA}^{T}\mathbf{N}}{|\mathrm{JA}^{\mathrm{T}}\mathbf{N}|} |\mathrm{JA}^{\mathrm{T}}\mathbf{N}|\, dS_y\,.$$

Since f can be chosen arbitrarily, the equality above suggests that

$$\mathbf{n}\circ\psi = \frac{\mathrm{JA}^{T}\mathbf{N}}{|\mathrm{JA}^{\mathrm{T}}\mathbf{N}|} = \frac{\mathrm{A}^{\mathrm{T}}\mathbf{N}}{|\mathrm{A}^{\mathrm{T}}\mathbf{N}|} \tag{B.19}$$

and

$$\psi^*(dS_x) = |\mathrm{JA}^{\mathrm{T}}\mathbf{N}|\, dS_y\,. \tag{B.20}$$

We finish this section by the following

Theorem B.109 (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*

$$\frac{d}{dt}\int_{\Omega(t)} f(x,t)dx = \int_{\Omega(t)} f_t(x,t)dx + \int_{\partial\Omega(t)} (\sigma f)(x,t)\, dS_x\,, \tag{B.21}$$

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 (B.3) and (B.17), 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 (B.5) 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 (B.19) and (B.20) we conclude that

$$\frac{d}{dt}\int_{\Omega(t)} f(x,t)dx = \int_{\Omega(t)} f_t(x,t)dx + \sum_{i,j=1}^{\mathrm{n}}\int_{\partial\Omega(t)}(\sigma f)(x,t)\,dS_x\,. \qquad \square$$

B.3 The Einstein Summation Convention

In mathematics, especially in applications of linear algebra to physics, the Einstein summation convention is a notational convention that implies summation over a set of indexed terms in a formula, thus achieving notational brevity. According to this convention, when an index variable appears twice in a single term it implies summation of that term over all the values of the index. For example, with this convention, the inner product $\boldsymbol{u}\cdot\boldsymbol{v}$ of two vectors $\boldsymbol{u},\boldsymbol{v}\in\mathbb{R}^{\mathrm{n}}$, where $\boldsymbol{u} = (u_1,\cdots,u_{\mathrm{n}})$ and $\boldsymbol{v} = (v_1,\cdots,v_{\mathrm{n}})$, can be expressed as u_iv_i, and the i-th component of the cross product $\boldsymbol{u}\times\boldsymbol{v}$ of two vectors $\boldsymbol{u},\boldsymbol{v}\in\mathbb{R}^3$ can be expressed as $\varepsilon_{ijk}u^jv^k$.

In this book, we make a further convention that repeated Latin indices are summed from 1 to n, and repeated Greek indices are summed from 1 to n $-$ 1, where n is the space dimension. In other words, we use the symbol f_ig_i to denote the sum $\sum\limits_{i=1}^{\mathrm{n}} f_ig_i$, and the symbol $f_\alpha g_\alpha$ to denote the sum $\sum_{i=1}^{\mathrm{n}-1} f_\alpha g_\alpha$. Starting from the next Chapter, we use such summation convention for notational simplicity.

B.4 Exercises

In this set of exercise, the Einstein summation convention is used.

Problem B.1. Complete the following.

1. Let $\delta_{\cdot\cdot}$'s are the Kronecker deltas. Prove (B.8); that is, show that

$$\varepsilon_{ijk}\varepsilon_{irs} = \delta_{jr}\delta_{ks} - \delta_{js}\delta_{kr}\,. \tag{B.8}$$

2. Use (B.8) to show the following identities:

   (a) $u \times (v \times w) = (u \cdot w)v - (u \cdot v)w$ if u, v, w are three 3-vectors.

   (b) $\mathrm{curlcurl}u = -\Delta u + \nabla \mathrm{div} u$ if $u : \Omega \to \mathbb{R}^3$ is smooth.

   (c) $u \times \mathrm{curl}u = \dfrac{1}{2}\nabla(|u|^2) - (u \cdot \nabla)u$ if $u : \Omega \to \mathbb{R}^3$ is smooth.

Problem B.2. Let $\psi(\cdot, t) : \Omega \to \Omega(t)$ be a diffeomorphism as defined in Theorem B.109, and $\mathrm{J} = \det(\nabla\psi)$ and $\mathrm{A} = (\nabla\psi)^{-1}$. Complete the proof of the Piola identity, identities (B.6), (B.19) and (B.20) by the following argument:

1. Let $u(\cdot, t) : \Omega(t) \to \mathbb{R}^{\mathrm{n}}$ be a smooth vector field. Show that

$$\int_{\Omega(t)} \mathrm{div} u \, dx = \int_\Omega \mathrm{J}\mathrm{A}^j_i (u \circ \psi)^i_{,j} \, dy\,;$$

   thus by the divergence theorem,

$$\int_{\partial\Omega(t)} u \cdot n \, dS_x = \int_{\partial\Omega} \mathrm{J}\mathrm{A}^j_i (u \circ \psi)^i \mathbf{N}_j \, dS_y - \int_\Omega (\mathrm{J}\mathrm{A}^j_i)_{,j}\, (u \circ \psi)^i dy\,. \tag{B.22}$$

2. By (B.22),

$$\int_\Omega (\mathrm{J}\mathrm{A}^j_i)_{,j}\, (u \circ \psi)^i dy = 0 \quad \forall\, u(\cdot, t) : \Omega(t) \to \mathbb{R}^{\mathrm{n}} \text{ vanishing on } \partial\Omega(t).$$

   As a consequence, the Piola identity is valid.

3. By the Piola identity, (B.22) implies that

$$\int_{\partial\Omega(t)} u \cdot n \, dS_x = \int_{\partial\Omega} \mathrm{J}\mathrm{A}^j_i (u \circ \psi)^i \mathbf{N}_j \, dS_y \quad \forall\, u(\cdot, t) : \Omega(t) \to \mathbb{R}^{\mathrm{n}} \text{ smooth.}$$

   Therefore, identities (B.19) and (B.20) are also valid.

4. By identity (B.6) (which is obtained independent of), show that

$$\mathrm{J},_k = \mathrm{J}\mathrm{A}^j_i \psi^i_{,jk} \, .$$

Appendix C

Important Topics in Functional Analysis

C.1 The Hahn-Banach Theorem

Definition C.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 C.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 C.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 C.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 C.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 C.6. *If X is a normed space and $x_0 \in X$, there exists $T \in X'$ such that*

$$Tx_0 = \|x_0\|_X \qquad \text{and} \qquad |Tx| \leqslant \|x\|_X \qquad \forall\ x \in X\,.$$

Theorem C.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 C.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 C.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 C.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 $Tx_0 > 1$.*

C.2 The Open Mapping and Closed Graph Theorem

Theorem C.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 $\bigcap_{n=1}^{\infty} U_n$ is dense in X.*

(b) *X is not a countable union of nowhere dense sets.*

Definition C.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 C.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 C.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 C.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 C.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 $T \in \mathscr{B}(X, Y)$.*

C.3 Compact Operators

Definition C.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 C.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 C.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 C.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 C.21. *The spectrum of a bounded operator $T \in \mathscr{B}(X)$ is bounded.*

Theorem C.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 C.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 C.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 C.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 C.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 C.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*

$$\|x\|_X < r \qquad \text{but} \qquad \|x - y\|_X \geqslant 1 \quad \forall\, y \in \mathcal{M}\,.$$

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 C.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 C.22, $\dim N(T - \lambda I) < \infty$. By (a) of Lemma C.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 C.29. *The continuous spectrum of a compact operator $T \in \mathscr{B}(X)$ contains at most one point, namely 0.*

THEOREM C.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

$$\begin{aligned} &M_1 \subsetneq M_2 \subsetneq M_3 \subsetneq \cdots, && (1)\\ &T(M_n) \subseteq M_n \qquad \text{for } n \geqslant 1, && (2)\\ &(T - \lambda_n I)(M_n) \subseteq M_{n-1} \qquad \text{for } n \geqslant 2. && (3) \end{aligned}$$

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 C.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 C.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 C.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*

$$\alpha = \dim N(T - \lambda I) \qquad \alpha^* = \dim N(T^* - \lambda I)$$
$$\beta = \dim X/R(T - \lambda I) \qquad \beta^* = \dim X/R(T^* - \lambda I)$$

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 C.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 C.22, $\alpha < \infty$. By Lemma C.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 C.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 C.30 states that $R(\Phi - \lambda I)$ cannot be all of X. □

Corollary C.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 C.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 C.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 C.36. A much deeper result states that the spectrum of a bounded operator $T \in \mathscr{B}(X)$ is also compact.

C.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 C.37 (Symmetry)**.** The operator $T \in \mathscr{B}(\mathcal{H})$ is called symmetric if $T = T'$.

Lemma C.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

$$\langle \mathcal{L}u, u \rangle_H = (\lambda u - Tu, u)_{\mathcal{H}} \geqslant (\lambda - M)\|u\|_{\mathcal{H}}^2 .$$

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 B.15) 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

$$\|Mu - Tu\|_{\mathcal{H}} \leqslant C(Mu - Tu, u)_{\mathcal{H}}^{1/2} \qquad \forall\, u \in \mathcal{H} \tag{C.1}$$

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 (C.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 C.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 $\lambda_0 = 0$. □

C.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 C.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*

$$C_1 \|u\|_X \leqslant \|Au\|_Y + \|Ku\|_Z \leqslant C_2 \|u\|_X \qquad \forall\, u \in X \tag{C.2}$$

for some positive constants C_1 and C_2. Then

1. *The dimension of* $\operatorname{Ker}(A)$ *is finite, the mapping A is an isomorphism from* $X/\operatorname{Ker}(A)$ *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* $\operatorname{Ker}(A)$*, then*

$$\|L_1 u\|_F \leqslant C_0 \|L_1\|_{\mathscr{B}(X,F)} \|Au\|_Y \qquad \forall\, u \in X\,. \tag{C.3}$$

3. *If G is a Banach space and $L_2 \in \mathscr{B}(X,G)$ satisfies*

$$L_2 u \neq 0 \qquad \forall\, u \in \operatorname{Ker}(A)\backslash\{0\}\,, \tag{C.4}$$

then

$$C_3\|u\|_X \leqslant \|Au\|_Y + \|L_2 u\|_G \leqslant C_4\|u\|_X \qquad \forall\, u \in X \tag{C.5}$$

for some positive constants C_3 and C_4.

Proof. 1. Because of (C.2), we find that

$$C_1\|u\|_X \leqslant \|Ku\|_Z \leqslant C_2\|u\|_X \qquad \forall\, u \in \operatorname{Ker}(A)\,. \tag{C.6}$$

Let $\{u_n\}_{n=1}^\infty$ be a bounded sequence in $\operatorname{Ker}(A) \subseteq X$. 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 (C.6) we find that $\{u_{n_k}\}_{k=1}^\infty$ converges in X. In other words, the identity map $\iota : \operatorname{Ker}(A) \to X$ is compact; thus (b) of Theorem C.22 implies that $\operatorname{Ker}(A)$ is finite dimensional.

Consider the quotient space $M = X/\operatorname{Ker}(A)$ 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 $\operatorname{Ker}(A)$ 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 C.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 $\operatorname{Ker}(A)$, 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 (C.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 $\big\{u_{n_k}\big\}_{k=1}^\infty$ such that $\big\{Ku_{n_k}\big\}_{k=1}^\infty$ converges in Z. Moreover, $\{Au_n\}_{n=1}^\infty$ converges to 0; thus using (C.2) we find that $\big\{u_{n_k}\big\}_{k=1}^\infty$ 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 (C.4) we conclude that $u = 0$ which contradicts to that
$\|u\|_X = \lim\limits_{k\to\infty} \|u_{n_k}\|_X = 1$. □

Example C.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, $\operatorname{Ker}(A) = \mathbb{R}$. 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 (2.28).

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

$$L_2 u = ku \qquad \forall\, u \in \operatorname{Ker}(A)\backslash\{0\}\,,$$

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 (2.28).

EXAMPLE C.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

$$L_2 \boldsymbol{u} \neq 0 \qquad \forall\, \boldsymbol{u} \in \operatorname{Ker}(A)\backslash\{\mathbf{0}\}\,.$$

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 suggests the Poincaré inequality (**??**).

Appendix D

Fourier Series and its Applications

D.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 D.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.

D.1.1 Trigonometric polynomials

DEFINITION D.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

$$p(x) = \sum_{k=-n}^{n} a_k e^{ikx} \quad \text{with} \quad a_k = \frac{c_{|k|} - i s_{|k|}}{2}\,,$$

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 D.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 $\big\{\widehat{f}(k)\big\}_{k=-\infty}$, $\{c_k\}_{k=0}^{\infty}$ 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 D.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 D.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. □

D.1.2 Approximations of the identity

Definition D.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 D.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 D.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 D.5 to obtain the last equality. Now given $\varepsilon > 0$. Since $f \in \mathscr{C}(\mathbb{T})$, there exists $\delta > 0$ such that $|f(x) - f(y)| < \frac{\varepsilon}{2}$ whenever $|x-y| < \delta$. Therefore,

$$\begin{aligned}
&|(\varphi_n \star f)(x) - f(x)| \\
&\qquad \leqslant \int_{|x-y|<\delta} \varphi_n(x-y)\big|f(x)-f(y)\big|dy + \int_{\delta \leqslant |x-y|} \varphi_n(x-y)\big|f(x)-f(y)\big|dy \\
&\qquad \leqslant \frac{\varepsilon}{2}\int_{\mathbb{T}} \varphi_n(x-y)\,dy + 2\max_{\mathbb{T}}|f| \int_{\delta\leqslant|z|\leqslant\pi} \varphi_n(z)\,dz\,.
\end{aligned}$$

By (3) of Definition D.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 D.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 $\int_{\mathbb{T}} \varphi_n(x)\,dx = 1$. 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 D.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 $\lim\limits_{n\to\infty} \dfrac{n!}{\sqrt{2\pi n}n^n e^{-n}} = 1$,

$$\begin{aligned}
\lim_{n\to\infty}\int_{\delta\leqslant|x|\leqslant\pi} \varphi_n(x)\,dx &\leqslant \lim_{n\to\infty} 2^{2n}\Big(\frac{1+\cos\delta}{2}\Big)^n \frac{\big(\sqrt{2\pi n}n^n e^{-n}\big)^2}{\sqrt{2\pi(2n)}(2n)^{2n}e^{-2n}} \\
&= \lim_{n\to\infty}\sqrt{\pi n}\Big(\frac{1+\cos\delta}{2}\Big)^n = 0\,.
\end{aligned}$$

So $\{\varphi_n\}_{n=1}^\infty$ is an approximation of the identity. By Theorem D.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 D.9. *For any $f \in L^2(\mathbb{T})$, $\lim_{n\to\infty} \|s_n(f,\cdot) - f\|_{L^2(\mathbb{T})} = 0$, and*

$$\|f\|^2_{L^2(\mathbb{T})} = \sum_{k=-\infty}^{\infty} |\widehat{f}(k)|^2 = \pi\Big[\frac{c_0^2}{2} + \sum_{k=1}^{\infty}(c_k^2 + s_k^2)\Big]. \quad (\textit{Parseval's identity}) \tag{D.1}$$

Proof. By Theorem D.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 D.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 (D.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 D.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 (D.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 D.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

$k \geqslant N$,

$$\begin{aligned}|\widehat{f}(k)| \leqslant |\widehat{f}(k) - \widehat{f}_\varepsilon(k)| + |\widehat{f}_\varepsilon(k)| &= \left| \int_{\mathbb{T}} \big[f(x) - f_\varepsilon(x)\big] \mathrm{e}_k(x)\, dx \right| + |\widehat{f}_\varepsilon(k)| \\ &\leqslant \|f - f_\varepsilon\|_{L^1(\mathbb{T})} + |\widehat{f}_\varepsilon(k)| < \varepsilon\end{aligned}$$

which implies that $\lim\limits_{|k|\to\infty} \widehat{f}(k) = 0$. □

Definition D.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 D.13. *The Fourier basis* e_k *converges weakly to* 0 *in* $L^2(\mathbb{T})$*.*

D.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, $\left\{\frac{1}{\sqrt{\pi}}, \sqrt{\frac{2}{\pi}}\cos kx\right\}_{k=1}^{\infty}$, $\left\{\sqrt{\frac{2}{\pi}}\sin kx\right\}_{k=1}^{\infty}$ are both maximal orthonormal sets on $L^2(0,\pi)$. So is $\left\{\frac{1}{\sqrt{\pi}}, \sqrt{\frac{2}{\pi}}\cos 2kx, \sqrt{\frac{2}{\pi}}\sin 2kx\right\}_{k=1}^{\infty}$.

We note that $\left\{\pm\frac{1}{\sqrt{\pi}}, \pm\sqrt{\frac{2}{\pi}}\cos kx\right\}_{k=1}^{\infty}$ 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 $\left\{\pm\sqrt{\frac{2}{\pi}}\sin kx\right\}_{k=1}^{\infty}$ is the collection of all non-trivial functions with unit $L^2(0,\pi)$-norm satisfying

$$
\begin{aligned}
u_{xx} &= \lambda u \qquad \text{for some } \lambda \in \mathbb{R}\,,\\
u(0) &= u(\pi) = 0\,.
\end{aligned}
$$

D.2 Uniform Convergence of the Fourier Series

Given $f \in L^2(\mathbb{T})$, by Corollary D.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 D.9 states that $D_n \star f \to f$ in $L^2(\mathbb{T})$ for all $f \in L^2(\mathbb{T})$.

Definition D.14. The function

$$D_n(x) = \frac{\sin(n+\frac{1}{2})x}{2\pi \sin\frac{x}{2}} \tag{D.2}$$

is called the *Dirichlet kernel*.

D.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 D.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

$$\begin{aligned} \Big|\int_{|y|\leqslant\delta} D_n(y)\big(f(x+y)-f(x)\big)dy\Big| &\leqslant \int_{-\delta}^{\delta} \frac{\big|f(x+y)-f(x)\big|}{\big|\sin\frac{y}{2}\big|}dy \\ &\leqslant \|f'\|_{L^\infty(\mathbb{T})} \int_{-\delta}^{\delta} \frac{y}{\sin\dfrac{y}{2}}dy \leqslant C\delta\,. \end{aligned} \tag{D.3}$$

As for the integral over $\delta < |y| \leqslant \pi$, we have

$$\int_\delta^\pi \sin\big(n+\frac{1}{2}\big)y\frac{f(x+y)-f(x)}{\sin\frac{y}{2}}dy = -\frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{f(x+y)-f(x)}{\sin\frac{y}{2}}\Big|_{y=\delta}^{y=\pi} + \int_\delta^\pi \frac{\cos\big(n+\frac{1}{2}\big)y}{n+\frac{1}{2}}\frac{d}{dy}\frac{f(x+y)-f(x)}{\sin\frac{y}{2}}dy\,.$$

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 (D.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

$$\big\|f - s_n(f,\cdot)\big\|_{L^\infty(\mathbb{T})} \leqslant C_n \inf_{p\in\mathscr{P}_n(\mathbb{T})}\|f-p\|_{L^\infty(\mathbb{T})} \tag{D.4}$$

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 D.16. *The Dirichlet kernel D_n satisfies that for all $n \in \mathbb{N}$,*

$$\int_{\mathbb{T}}\big|D_n(x)\big|dx \leqslant 2 + \log n\,. \tag{D.5}$$

Proof. The validity of (D.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

$$\int_0^{\frac{1}{n}} 2\big|D_n(x)\big| dx \leqslant \frac{1}{\pi}\frac{2n+1}{n}\,. \tag{D.6}$$

Since $\dfrac{2x}{\pi} \leqslant \sin x$ for $0 \leqslant x \leqslant \dfrac{\pi}{2}$, the second integral can be estimated by

$$\int_{\frac{1}{n}}^{\pi} \Big|\frac{\sin(n+\frac{1}{2})x}{\pi \sin\frac{x}{2}}\Big| dx \leqslant \int_{\frac{1}{n}}^{\pi} \frac{1}{x}\, dx = \log\pi + \log n\,. \tag{D.7}$$

We then conclude (D.5) from (D.6) and (D.7) by noting that $\log\pi + \dfrac{2n+1}{n\pi} \leqslant 2$ for all $n \geqslant 2$. $\square$

Remark D.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 D.16, we are able to prove the inverse inequality (D.4). The following theorem is a direct consequence of Proposition D.16.

Theorem D.18. *Let $f \in \mathscr{C}(\mathbb{T})$; that is, f is a continuous function with period 2π. Then*

$$\big\|f - s_n(f,\cdot)\big\|_{L^\infty(\mathbb{T})} \leqslant (3 + \log n) \inf_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_{L^\infty(\mathbb{T})}\,. \tag{D.8}$$

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

$$\begin{aligned}\big\|f - s_n(f,\cdot)\big\|_{L^\infty(\mathbb{T})} &\leqslant \big\|f-p\big\|_{L^\infty(\mathbb{T})} + \big\|p - s_n(f,\cdot)\big\|_{L^\infty(\mathbb{T})} \\ &\leqslant \big\|f-p\big\|_{L^\infty(\mathbb{T})} + \big\|s_n(f-p,\cdot)\big\|_{L^\infty(\mathbb{T})} \\ &\leqslant \big\|f-p\big\|_{L^\infty(\mathbb{T})} + (2+\log n)\|f-p\|_{L^\infty(\mathbb{T})} \\ &\leqslant (3+\log n)\Big[\inf_{p\in\mathscr{P}_n(\mathbb{T})}\|f-p\|_{L^\infty(\mathbb{T})} + \epsilon\Big],\end{aligned}$$

and (D.8) is obtained by passing to the limit as $\epsilon \to 0$. $\square$

Having established Theorem D.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 D.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 D.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 D.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 D.19. *If f is a continuous function on $[a,b]$, then for all $\delta_1 > 0$,*

$$\sup_{|x-y|\leqslant\delta_1}\big|f(x)-f(y)\big| \leqslant \Big(1+\frac{\delta_1}{\delta_2}\Big)\sup_{|x-y|\leqslant\delta_2}\big|f(x)-f(y)\big|.$$

The proof of Lemma D.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 D.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 D.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 $\lim_{n\to\infty} n^{-\alpha}\log n = 0$ for all $\alpha \in (0,1]$, we conclude the following

Theorem D.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 D.22. The converse of Theorem D.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 $\alpha \in (0,1)$,

$$\inf_{p\in\mathscr{P}_n(\mathbb{T})} \|f-p\|_{L^\infty(\mathbb{T})} \leqslant Cn^{-\alpha} \tag{D.9}$$

for all $n \in \mathbb{N}$, then $f \in \mathscr{C}^{0,\alpha}(\mathbb{T})$. In other words, (D.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 D.4.

D.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 D.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, $c = \frac{1}{\pi}\int_0^\pi \frac{\sin x}{x}\,dx - \frac{1}{2} \approx 0.089490$), such that

$$\lim_{n\to\infty} s_n\big(f, x_0 + \frac{L}{2n}\big) = f(x_0^+) + ca\,, \tag{D.10a}$$

$$\lim_{n\to\infty} s_n\big(f, x_0 - \frac{L}{2n}\big) = f(x_0^-) - ca\,. \tag{D.10b}$$

Moreover,

$$\lim_{n\to\infty} s_n(f, x_0) = \frac{f(x_0^+) + f(x_0^-)}{2}\,. \tag{D.11}$$

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

$$g(x) = \begin{cases} \dfrac{a}{2\pi}(x+\pi) & \text{if } -\pi \leqslant x < 0\,, \\ 0 & \text{if } x = 0\,, \\ \dfrac{a}{2\pi}(x-\pi) & \text{if } 0 < x \leqslant \pi\,. \end{cases}$$

Then $F = f + g$ is Lipchitz continuous on $\mathbb{T}$, thus by Theorem D.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 (D.11).

By $\widehat{g}(k) = \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\pi}^{\pi} g(x)\, e^{-ikx} dx = \dfrac{ia}{\sqrt{2\pi}k}$ 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 (D.10a) follows from $\frac{f(0^+)+f(0^-)}{2} + \frac{a}{2} = f(0^+)$. (D.10b) can be proved in the same fashion, and is left as an exercise. □

D.3 The Sobolev Space $H^s(\mathbb{T})$

Definition D.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 D.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

$$\widehat{H}(k) = \frac{1}{\sqrt{2\pi}} \int_0^\pi e^{-ikx} dx = \begin{cases} \sqrt{\frac{\pi}{2}} & \text{if } k = 0, \\ 0 & \text{if } k \text{ is even, } k \neq 0, \\ \sqrt{\frac{2}{\pi}}\,\frac{1}{ik} & \text{if } k \text{ is odd,} \end{cases}$$

hence $H \in H^s(\mathbb{T})$ if $s < \dfrac{1}{2}$.

Proposition D.26. *If $0 < s < r$, then $H^r(\mathbb{T}) \subseteq H^s(\mathbb{T})$.*

Proposition D.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 D.28. *Let $0 < r < t < \infty$, and $s = \alpha r + (1-\alpha)t$ for some $\alpha \in (0,1)$. Then*

$$\|u\|_{H^s(\mathbb{T})} \leqslant \|u\|^{\alpha}_{H^r(\mathbb{T})} \|u\|^{1-\alpha}_{H^t(\mathbb{T})} . \tag{D.12}$$

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 (D.12). ◻

Theorem D.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*

$$\|f\|_{L^\infty(\mathbb{T})} \leqslant C_s \|f\|_{H^s(\mathbb{T})} \qquad \forall\, f \in H^s(\mathbb{T}) . \tag{D.13}$$

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, $\|s_n(f,\cdot) - s_m(f,\cdot)\|_{L^\infty(\mathbb{T})} \to 0$ as $n, m \to \infty$, which implies that $s_n(f,\cdot)$ converges uniformly; hence $\widetilde{f} \equiv \lim_{n\to\infty} s_n(f,\cdot)$ is continuous.

The constant C_s in (D.13) can be chosen as $\dfrac{1}{\sqrt{2\pi}}\Big[\sum_{k=-\infty}^{\infty} \dfrac{1}{(1+|k|^2)^s}\Big]^{\frac{1}{2}}$. □

D.3.1 Characterization of $H^1(\mathbb{T})$

Definition D.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 D.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

$$-\int_{\mathbb{T}} f'(x)\varphi(x)\, dx = \int_{\mathbb{T}} f(x)\varphi'(x)\, dx\,, \tag{D.14}$$

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 D.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*

$$-\int_{\mathbb{T}} g(x)\varphi(x)dx = \int_{\mathbb{T}} f(x)\varphi'(x)dx \qquad \forall\, \varphi \in \mathscr{C}^1(\mathbb{T})\,. \tag{D.15}$$

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 (D.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

$$\sum_{k=-\infty}^{\infty} |k|^2|\widehat{f}(k)|^2 = \sum_{k=-\infty}^{\infty} |\widehat{g}(k)|^2 = \|g\|^2_{L^2(\mathbb{T})} < \infty\,. \qquad \square$$

Corollary D.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 D.33. The proof of Theorem D.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 D.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 D.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 D.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*

$$C_1\|f\|_{H^k(\mathbb{T})} \leqslant \sum_{j=0}^{k}\|f^{(j)}\|_{L^2(\mathbb{T})} \leqslant C_2\|f\|_{H^k(\mathbb{T})}\,.$$

D.3.2 The space $H^k(0,\pi)$

Motivated by Theorem D.31 and Corollary D.32, we look for a qualitative description of the H^k-space using the language of weak derivatives.

Definition D.37. A function $u \in L^1_{\text{loc}}(0,\pi)$ is said to be weakly differentiable if there exists a function $g \in L^1_{\text{loc}}(0,\pi)$ such that

$$-\int_0^\pi g(x)\varphi(x)dx = \int_0^\pi f(x)\varphi'(x)dx \qquad \forall\, \varphi \in \mathscr{C}^1_c((0,\pi)) . \tag{D.16}$$

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 $L^1_{\text{loc}}(0,\pi)$ is chosen so that the integrals in (D.16) make sense. Moreover, the test function φ in (D.16) is compactly supported in $(0,\pi)$; that is, $\text{spt}(\varphi) \subseteq (0,\pi)$.

Definition D.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 D.39. As mentioned in Section D.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 D.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.

D.4 1-Dimensional Heat Equations with Periodic Boundary Condition

In this section, we consider the heat equation:

$$u_t(x,t) - u_{xx}(x,t) = f(x,t) \qquad \text{for all } (x,t) \in (0,2\pi) \times (0,T), \tag{D.17a}$$

$$u(0,t) = u(2\pi,t) \qquad \text{for all } t \in (0,T), \tag{D.17b}$$

$$u(x,0) = g(x) \qquad \text{for all } x \in (0,2\pi). \tag{D.17c}$$

Condition (D.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

$$\int_0^T \|f(\cdot,t)\|^2_{H^1(\mathbb{T})} dt \equiv \int_0^T \int_{\mathbb{T}} \big(|f(x,t)|^2 + |f_x(x,t)|^2\big) dx dt < \infty\,.$$

D.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 (D.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 (D.17a), for almost all $t \in [0,T]$,

$$\frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \Big[d_k'(t) + k^2 d_k(t) - \widehat{f}_k(t) \Big] e^{ikx} = 0 \,. \tag{D.18}$$

Since $\{\mathrm{e}_k\}_{k=-\infty}^{\infty}$ is maximal, we find that $d_k(t)$ solves the ODE

$$d_k'(t) + k^2 d_k(t) = \widehat{f}_k(t) \,. \tag{D.19}$$

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

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \Big[e^{-k^2 t} \widehat{g}(k) + \int_0^t \widehat{f}_k(s) e^{-k^2(t-s)} ds \Big] e^{ikx} \,. \tag{D.20}$$

D.4.2 Rigorous approaches

Before proceeding, we state a very important theorem in the study of differential equations.

Theorem D.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*

$$x(t) \leqslant e^{Ct} x(0) + \frac{M}{C}(e^{Ct} - 1) \qquad \forall\, t \in [0,T] \,. \tag{D.21}$$

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 D.41. *Let $y(t)$ be a non-negative, integrable function on the interval $[0,T]$. If $y(t)$ satisfies*

$$y(t) \leqslant M + C\int_0^t y(s)\,ds \qquad \forall\, t \in [0,T]\,, \tag{D.22}$$

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 D.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

$$u_{nt}(x,t) - u_{nxx}(x,t) = f_n(x,t) \qquad \text{for all } (x,t) \in (0,2\pi) \times (0,T), \tag{D.23a}$$

$$u_n(0,t) = u_n(2\pi,t) \qquad \text{for all } t \in (0,T), \tag{D.23b}$$

$$u_n(x,0) = g_n(x) \qquad \text{for all } x \in (0,2\pi). \tag{D.23c}$$

The same procedure as the formal approach implies that

$$u_n(x,t) = \frac{1}{\sqrt{2\pi}} \sum_{k=-n}^{n} \Big[e^{-k^2 t}\widehat{g}(k) + \int_0^t \widehat{f}_k(s) e^{-k^2(t-s)} ds \Big] e^{ikx}$$

is a solution to (D.23). Our goal is to show that u_n converges to the solution of (D.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

$$v_t^{n,m}(x,t) - v_{xx}^{n,m}(x,t) = f^{n,m}(x,t) \qquad \text{for all } (x,t) \in (0,2\pi) \times (0,T), \tag{D.24a}$$

$$v^{n,m}(0,t) = v^{n,m}(2\pi,t) \qquad \text{for all } t \in (0,T), \tag{D.24b}$$

$$v^{n,m}(x,0) = g^{n,m}(x) \qquad \text{for all } x \in (0,2\pi). \tag{D.24c}$$

Multiplying (D.24a) by $v_{xxxx}^{n,m}(x,t)$ and integrating over $\mathbb{T}$,

$$\int_{\mathbb{T}} \Big[v_t^{n,m}(x,t) - v_{xx}^{n,m}(x,t) \Big] v_{xxxx}^{n,m}(x,t)\,dx = \int_{\mathbb{T}} f^{n,m}(x,t) v_{xxxx}^{n,m}(x,t)\,dx\,. \tag{D.25}$$

Integrating by parts in x, we find that

$$\int_{\mathbb{T}} v_t^{n,m}(x,t) v_{xxxx}^{n,m}(x,t)\,dx = \int_{\mathbb{T}} v_{xxt}^{n,m}(x,t) v_{xx}^{n,m}(x,t)\,dx$$
$$= \int_{\mathbb{T}} \frac{1}{2}\frac{\partial}{\partial t}|v_{xx}^{n,m}(x,t)|^2 dx = \frac{1}{2}\frac{d}{dt}\|v_{xx}^{n,m}(\cdot,t)\|^2_{L^2(\mathbb{T})}$$

and

$$-\int_{\mathbb{T}} v_{xx}^{n,m}(x,t) u_{nxxxx}(x,t)\,dx = \int_{\mathbb{T}} v_{xxx}^{n,m}(x,t) v_{xxx}^{n,m}(x,t)\,dx = \|v_{xxx}^{n,m}(\cdot,t)\|^2_{L^2(\mathbb{T})}\,.$$

Moreover, by Hölder's and Young's inequality,

$$\int_{\mathbb{T}} f^{n,m}(x,s) v_{xxxx}^{n,m}(x,s)\,dx = -\int_{\mathbb{T}} f_x^{n,m}(x,s) v_{xxx}^{n,m}(x,s)\,dx$$
$$\leqslant \|f_x^{n,m}(\cdot,s)\|_{L^2(\mathbb{T})}\|v_{xxx}^{n,m}(\cdot,s)\|_{L^2(\mathbb{T})} \leqslant \frac{1}{2}\Big[\|f_x^{n,m}(\cdot,s)\|^2_{L^2(\mathbb{T})} + \|v_{xxx}^{n,m}(\cdot,s)\|^2_{L^2(\mathbb{T})}\Big]\,.$$

As a consequence, (D.25) implies that

$$\frac{d}{dt}\|v_{xx}^{n,m}(\cdot,t)\|^2_{L^2(\mathbb{T})} + \|v_{xxx}^{n,m}(\cdot,t)\|^2_{L^2(\mathbb{T})} \leqslant \|f_x^{n,m}(\cdot,t)\|^2_{L^2(\mathbb{T})}$$

and integrating in t over the time interval $(0,t)$ further implies that

$$\begin{aligned}\|v_{xx}^{n,m}(\cdot,t)\|^2_{L^2(\mathbb{T})} &+ \int_0^t \|v_{xxx}^{n,m}(\cdot,t)\|^2_{L^2(\mathbb{T})} \\ &\leqslant \|g_{xx}^{n,m}\|^2_{L^2(\mathbb{T})} + \int_0^t \|f_x^{n,m}(\cdot,s)\|^2_{L^2(\mathbb{T})}\,ds\,.\end{aligned} \tag{D.26}$$

Similarly, multiplying (D.24a) by $v^{n,m}$ or $v_{xx}^{n,m}$ and then integrating over $\mathbb{T}$, we obtain that

$$\begin{aligned}\|v_x^{n,m}(\cdot,t)\|^2_{L^2(\mathbb{T})} &+ \int_0^t \|v_{xx}^{n,m}(\cdot,t)\|^2_{L^2(\mathbb{T})} \\ &\leqslant \|g_x^{n,m}\|^2_{L^2(\mathbb{T})} + \int_0^t \|f^{n,m}(\cdot,s)\|^2_{L^2(\mathbb{T})}\,ds\,.\end{aligned} \tag{D.27}$$

and

$$\begin{aligned}\|v^{n,m}(\cdot,t)\|^2_{L^2(\mathbb{T})} + 2\int_0^t \|v_x^{n,m}(\cdot,t)\|^2_{L^2(\mathbb{T})} &\leqslant \|g^{n,m}\|^2_{H^2(\mathbb{T})} \\ + \int_0^t \|f^{n,m}(\cdot,s)\|^2_{H^1(\mathbb{T})}\,ds &+ \int_0^t \|v^{n,m}(\cdot,s)\|^2_{L^2(\mathbb{T})} ds\,.\end{aligned} \tag{D.28}$$

Summing (D.26), (D.27) and (D.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

$$\max_{t\in[0,T]} \|v^{n,m}(t)\|^2_{H^2(\mathbb{T})} \leqslant 2\Big[\|g^{n,m}\|^2_{H^2(\mathbb{T})} + \int_0^T \|f^{n,m}(\cdot,s)\|^2_{H^1(\mathbb{T})}\,ds\Big]e^T\,. \tag{D.29}$$

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 (D.29) we obtain that $u \in \mathscr{C}([0,T];H^2(\mathbb{T}))$ satisfies

$$\max_{t\in[0,T]} \|u(t)\|^2_{H^2(\mathbb{T})} \leqslant C\Big[\|g\|^2_{H^2(\mathbb{T})} + \int_0^T \|f(t)\|^2_{H^1(\mathbb{T})}dt\Big]\,. \tag{D.30}$$

So we conclude the following

Theorem D.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*

$$u(x,t) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \Big[e^{-k^2t}\widehat{g}(k) + \int_0^t \widehat{f}_k(s)e^{-k^2(t-s)}ds\Big]e^{ikx} \tag{D.31}$$

solves (D.17) (*in the sense of weak spatial derivatives*). *Moreover, u belongs to the space $\mathscr{C}([0,T];H^2(\mathbb{T}))$ and satisfies* (D.30).

REMARK D.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 (D.31) is

$$u(x,t) = \frac{1}{2\pi}\Big[\big(\frac{4\pi}{3}t + \frac{2}{\pi}t^2\big) + \frac{4}{\pi}\sum_{k\neq 0}\frac{(e^{-k^2t}-1)}{k^4}e^{ikx}\Big].$$

D.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

$$\frac{d}{dt}\big\|u(\cdot,t)\big\|^p_{L^p(\mathbb{T})} + p(p-1)\int_{\mathbb{T}} u_x^2(x,t)u^{p-2}(x,t)\,dx = 0\,.$$

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

$$\big\|u(\cdot,t)\big\|_{L^\infty(\mathbb{T})} \leqslant \|g\|_{L^\infty(\mathbb{T})} \qquad \forall\, t > 0\,. \tag{D.32}$$

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).

D.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 (D.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 D.44. *Let u be the solution to the heat equation* (D.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.*

D.5 1-Dimensional Heat Equations with Dirichlet Boundary Condition

In this section, we consider the following initial-boundary value problem for the heat equation

$$
\begin{aligned}
u_t(x,t) - u_{xx}(x,t) &= f(x,t) && \text{for all } (x,t) \in (0,L) \times (0,T), && \text{(D.33a)}\\
u(0,t) = u(L,t) &= 0 && \text{for all } t \in (0,T)\,, && \text{(D.33b)}\\
u(x,0) &= g(x) && \text{for all } x \in [0,L]. && \text{(D.33c)}
\end{aligned}
$$

Because of the boundary condition (D.33b), we use the orthonormal basis $\left\{\sqrt{\frac{2}{L}}\sin\frac{k\pi x}{L}\right\}_{k=1}^{\infty}$. 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 (D.33) can be expressed by

$$
u(x,t) = \sum_{k=1}^{\infty}\left[d_k(0)e^{-\frac{\pi^2k^2}{L^2}t} + \int_0^t f_k(s)e^{-\frac{\pi^2k^2}{L^2}(t-s)}ds\right]\sin\frac{k\pi x}{L}\,. \tag{D.34}
$$

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

$$v_t^{n,m}(x,t) - v_{xx}^{n,m}(x,t) = f^{n,m}(x,t) \quad \text{for all } (x,t) \in (0,L) \times (0,T), \tag{D.35a}$$

$$v^{n,m}(0,t) = v^{n,m}(L,t) = 0 \quad \text{for all } t \in (0,T), \tag{D.35b}$$

$$v^{n,m}(x,0) = g^{n,m}(x) \quad \text{for all } x \in (0,L). \tag{D.35c}$$

Unlike the case in Section D.4.2, this time we cannot multiply (D.35a) by $v_{xxxx}^{n,m}(x,t)$ then integrating over $(0,L)$ since non-vanishing uncontrollable boundary terms pop out after integrating by parts. To overcome this, we differentiate (D.35a) with respect to t and then multiply the resulting equation with $v_t^{n,m}(x,t)$ and obtain that

$$\big(v_{tt}^{n,m}(t), v_t^{n,m}(t)\big)_{L^2(0,L)} - \big(v_{xxt}^{n,m}(t), v_t^{n,m}(t)\big)_{L^2(0,L)} = \big(f_t^{n,m}(t), v_t^{n,m}(t)\big)_{L^2(0,L)}.$$

It is easy to see that

$$\big(v_{tt}^{n,m}(t), v_t^{n,m}(t)\big)_{L^2(0,L)} = \frac{1}{2}\frac{d}{dt}\|v_t^{n,m}(\cdot,t)\|_{L^2(0,L)}^2.$$

Since $v_t^{n,m}(0,t) = v_t^{n,m}(L,t) = 0$, integrating by parts we have

$$-\big(v_{xxt}^{n,m}(t), v_t^{n,m}(t)\big)_{L^2(0,L)} = \|v_{xt}^{n,m}(\cdot,t)\|_{L^2(0,L)}^2.$$

As a consequence,

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}&\|v_t^{n,m}(\cdot,t)\|_{L^2(0,L)}^2 + \|v_{xt}^{n,m}(\cdot,t)\|_{L^2(0,L)}^2 \\ &\leqslant \frac{1}{2}\|f_t^{n,m}(\cdot,t)\|_{L^2(0,L)}^2 + \frac{1}{2}\|v_t^{n,m}(\cdot,t)\|_{L^2(0,L)}^2;\end{aligned}$$

thus the Gronwall inequality implies that

$$\max_{t\in[0,T]} \|v_t^{n,m}(\cdot,t)\|_{L^2(0,L)}^2 \leqslant \Big[\|v_t^{n,m}(\cdot,0)\|_{L^2(0,L)}^2 + \int_0^T \|f_t^{n,m}(\cdot,t)\|_{L^2(0,L)}^2 dt\Big] e^T.$$

Using (D.35a,c), we further conclude that

$$\begin{aligned}\max_{t\in[0,T]} &\|v_t^{n,m}(\cdot,t)\|_{L^2(0,L)}^2 \\ &\leqslant C\Big[\|v_{xx}^{n,m}(\cdot,0)\|_{L^2(0,T)}^2 + \|f^{n,m}(\cdot,0)\|_{L^2(0,T)}^2 + \int_0^T \|f_t^{n,m}(\cdot,t)\|_{L^2(0,L)}^2 dt\Big] e^T \\ &\leqslant C\Big[\|g_{xx}^{n,m}\|_{L^2(0,T)}^2 + \|f^{n,m}(\cdot,0)\|_{L^2(0,T)}^2 + \int_0^T \|f_t^{n,m}(\cdot,t)\|_{L^2(0,L)}^2 dt\Big] e^T,\end{aligned} \tag{D.36}$$

where we emphasize that $v_{xx}^{n,m} = (v^{n,m})_{xx}$ and $g_{xx}^{n,m} = (g^{n,m})_{xx}$.

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

$$\|f^{n,m}(\cdot,0)\|^2_{L^2(0,L)} + \int_0^T \|f_t^{n,m}(\cdot,t)\|^2_{L^2(0,L)} dt \to 0 \quad \text{as } n,m \to \infty\,.$$

The convergence of $\|g_{xx}^{n,m}\|_{L^2(0,L)}$ 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 $\|g_{xx}^{n,m}\|_{L^2(0,L)} \to 0$. 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 $\|g_{xx}^{n,m}\|_{L^2(0,L)} \to \infty$ 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

$$\Big(\frac{d}{dx}\Big)^2 \sum_{k=1}^{\infty} a_k \sin\frac{k\pi x}{L} = \sum_{k=1}^{\infty} a_k \Big(\frac{d}{dx}\Big)^2 \sin\frac{k\pi x}{L}\,, \tag{D.37}$$

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 (D.37) to hold, the function $\sum\limits_{k=1}^{\infty} a_k \sin\frac{k\pi x}{L}$ and its weak derivative has to vanish at $x = 0$ and $x = L$. In fact, we have the following

Lemma D.45. *If $g \in H^2(0,L)$ and $g(0) = g'(0) = g(L) = g'(L) = 0$, then the partial sum*

$$g_n(x) = \sqrt{\frac{2}{L}}\sum_{k=1}^{n} \widehat{g}(k)\sin\frac{k\pi x}{L}\,, \quad \textit{where } \ \widehat{g}(k) = \sqrt{\frac{2}{L}}\int_0^L g(x)\sin\frac{k\pi x}{L}dx\,,$$

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. $g_n''(x) = -\sqrt{\frac{2}{L}}\sum\limits_{k=1}^{n}\frac{k^2\pi^2}{L^2}\widehat{g}(k)\sin\frac{k\pi x}{L}$;

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 $m > n$,

$$\big\|(g_n - g_m)''\big\|^2_{L^2(0,L)} = \sum_{k=n+1}^{m}\left|\frac{k^2\pi^2}{L^2}\widehat{g}(k)\right|^2$$

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 D.45 suggests that $v_t^{n,m}$ is a Cauchy sequence in $\mathscr{C}([0,T]; L^2(0,L))$. Similarly, multiplying (D.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 (D.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

$$\begin{aligned}&\max_{t\in[0,T]}\Big[\|u_t(t)\|_{L^2(0,L)} + \|u(t)\|_{H^2(0,L)}\Big]\\ &\qquad\leqslant C\Big[\|g\|_{H^2(0,L)} + \max_{t\in[0,T]}\|f(t)\|^2_{L^2(0,L)} + \int_0^T \|f_t(t)\|^2_{L^2(0,L)}dt\Big].\end{aligned} \tag{D.38}$$

Therefore, we establish the following

Theorem D.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* (D.34) *is a solution to* (D.33). *Moreover, u satisfies* (D.38).

D.6 Exercises

Problem D.1. Prove Lemma D.10.

Problem D.2. Let f be a 2π-periodic Lipchitz function. Show that for $n \geqslant 2$,

$$\|f - F_{n+1} \star f\|_{L^\infty(\mathbb{T})} \leqslant \frac{1 + 2\log n}{2n} \|f\|_{\mathscr{C}^{0,1}(\mathbb{T})} \tag{D.39}$$

and

$$\big\|f - s_n(f, \cdot)\big\|_{L^\infty(\mathbb{T})} \leqslant \frac{2\pi(1 + \log n)^2}{n} \|f\|_{\mathscr{C}^{0,1}(\mathbb{T})}\,. \tag{D.40}$$

Hint: For (D.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 (D.40), use (D.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 D.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 $\bigcup_i \bar{I}_i = \mathbb{T}$. Show that $D_n \star f$ converges to f uniformly as $n \to \infty$ on any compact subset of I_i.

Problem D.4. In this problem, we are concerned with the following

Theorem D.47 (Bernstein)**.** *Suppose that f is a 2π-periodic function such that for some constant C and $\alpha \in (0,1)$,*

$$\inf_{p \in \mathscr{P}_n(\mathbb{T})} \|f - p\|_{L^\infty(\mathbb{T})} \leqslant C n^{-\alpha}$$

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 $\alpha_k = \gamma + \frac{\pi}{n}(k + \frac{1}{2})$ 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

$$\|p'\|_{L^\infty(\mathbb{T})} \leqslant n\|p\|_{L^\infty(\mathbb{T})} \qquad \forall\, p \in \mathscr{P}_n(\mathbb{T}). \tag{D.41}$$

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 D.5. Show that $\{w_k\}_{k=0}^\infty$ defined in Remark D.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 D.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 $\sum_{n=-\infty}^{\infty} |n|^2 |\widehat{f}(n)|^2 = \infty$, so by "definition", it does not seem to be a function in $H^1(-\pi,\pi)$. What is wrong with the argument?

PROBLEM D.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 D.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 D.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 D.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 D.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 D.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

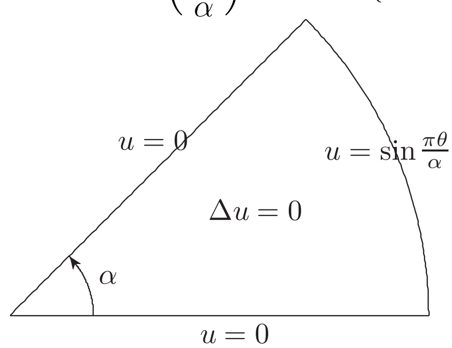
$$\Omega_\alpha = \{(r,\theta) \mid 0 < r < 1\,,\ 0 < \theta < \alpha\}\,.$$

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

$$\begin{aligned} \Delta u &= 0 && \text{in} \quad \Omega_\alpha\,,\\ u &= 0 && \text{on} \quad \{\theta = 0, \alpha\}\,,\\ u &= \sin\left(\frac{\pi\theta}{\alpha}\right) && \text{on} \quad \{r = 1\}\,. \end{aligned}$$



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 $u \in \mathscr{C}^2(\overline{\Omega_\alpha})$.

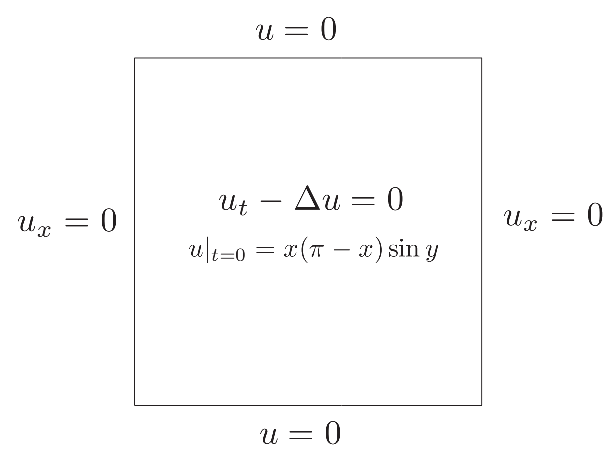
Problem D.13. Complete the following.

(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:

$$\begin{aligned}
u_t(x,y,t) - \Delta u(x,y,t) &= 0 \qquad && (x,y) \in (0,\pi)\times(0,\pi)\,, t > 0\,,\\
u(x,y,0) &= x(\pi - x)\sin y && (x,y) \in (0,\pi)\times(0,\pi)\,,\\
u_x(0,y,t) = u_x(\pi,y,t) &= 0 && y \in (0,\pi)\,, t > 0\,,\\
u(x,0,t) = u(x,\pi,t) &= 0 && x \in (0,\pi)\,, t > 0\,.
\end{aligned}$$



(3) Show that for all $t \geqslant 0$, u from (b) satisfies

$$\int_0^\pi\int_0^\pi |u(x,y,t)|^2 dxdy + 2\int_0^t\int_0^\pi\int_0^\pi |\nabla u(x,y,s)|^2 dxdyds = \frac{\pi^6}{60}\,.$$

Index