Sobolev Spaces and Applications

T. Muthukumar
tmk@iitk.ac.in

April 22, 2016

Contents

Notations

Symbols

∇ $\left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right)$

Ω denotes an open subset of $\mathbb{R}^n$, not necessarily bounded

$\mathbb{R}$ denotes the real line

$\mathbb{R}^n$ denotes the n-dimensional Euclidean space

$|E|$ will denote the Lebesgue measure of $E \subset \mathbb{R}^n$

D^α $\frac{\partial^{\alpha_1}}{\partial {x_1}^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial {x_n}^{\alpha_n}}$ and $\alpha = (\alpha_1, \ldots, \alpha_n)$. In particular, for $\alpha = (1, 1, \ldots, 1)$, $D = \nabla = D^{(1,1,,\ldots,1)} = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right)$

D^α $\frac{\partial^{\alpha_1}}{\partial {x_1}^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial {x_n}^{\alpha_n}}$ and $\alpha = (\alpha_1, \ldots, \alpha_n)$

Function Spaces

$C^{k,\gamma}(\Omega)$ all functions in $C^k(\Omega)$ whose k-th partial derivatives ($k \geq 0$) are Hölder continuous with exponent γ

$\mathrm{Lip}(E)$ denotes the space of all Lipschitz functions on E

$C(\Omega)$ is the class of all continuous functions on Ω

$C^k(\Omega)$ is the class of all k-times ($k \geq 1$) continuously differentiable functions on Ω

$C^\infty(\overline{\Omega})$ denotes class of $C^\infty(\Omega)$ functions such that all its derivatives can be extended continuously to $\overline{\Omega}$

$C_c^\infty(\Omega)$ is the class of all infinitely differentiable functions on Ω with compact support

$C_0(\Omega)$ is the class of all continuous functions on Ω that vanishes at boundary (for bounded sets) or ∞

$H^k(\Omega)$ is the space $W^{k,2}(\Omega)$

$H_0^k(\Omega)$ is the space $W_0^{k,2}(\Omega)$

$\mathcal{E}(\Omega)$ is the class of all infinitely differentiable functions on Ω, also denoted as $C^\infty(\Omega)$, endowed with topology of uniform convergence on compact subsets

$\mathcal{D}(\Omega)$ is the class of all infinitely differentiable functions on Ω with compact support endowed with inductive limit topology

$W^{k,p}(\Omega)$ is the class of all functions in $L^p(\Omega)$ whose distributional derivative upto order k are also in $L^p(\Omega)$

General Conventions

$\mathcal{M}(\alpha)$ denotes, for $\alpha > 0$, the class of all $n \times n$ matrices, $A = A(x)$, with $L^\infty(\Omega)$ entries such that,

$$\alpha|\xi|^2 \leq A(x)\xi.\xi \quad a.e.\ x \quad \forall \xi \in \mathbb{R}^n$$

Chapter 1

Theory of Distributions

1.1 History

In late 1920's P. A. M. Dirac (cf. [Dir49]) derived the equation

$$\frac{d}{dx}\ln(x) = \frac{1}{x} - i\pi\delta(x) \tag{1.1.1}$$

in the study of quantum theory of collision processes, where δ is a "function" defined and continuous in real line $\mathbb{R}$ satisfying the following properties:

1. $\delta(x) = 0$ for $x \neq 0$.
2. $\int_{-\infty}^{\infty} \delta(x)\, dx = 1$.
3. For any continuous function defined on $\mathbb{R}$, $f(a) = \int_{-\infty}^{\infty} f(x)\delta(a-x)\, dx$ for all $a \in \mathbb{R}$.
4. δ is infinitely differentiable and for any k-times continuously differentiable function f on $\mathbb{R}$, $f^{(k)}(a) = \int_{-\infty}^{\infty} f(x)\delta^{(k)}(a-x)\, dx$ for all $a \in \mathbb{R}$.
5. Given the Heaviside function

   $$H(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x < 0, \end{cases} \tag{1.1.2}$$

   then the delta function is the "derivative" of H, i.e., $\delta(x) = H'(x)$.

The δ became known as *Dirac's delta function*. It had caused unrest among the mathematicians of the era, because δ did not adhere to the classical notion of function.

H. Lebesgue had laid the foundations of measure theory in his doctoral thesis in 1902. With this knowledge, δ can be viewed as a set function on the σ-algebra(a collection of subsets of $\mathbb{R}$). δ can be viewed as a positive measure, called *Dirac measure*, defined on subsets E from the σ-algebra of $\mathbb{R}$ as,

$$\delta(E) = \begin{cases} 1 & \text{if } 0 \in E \\ 0 & \text{if } 0 \notin E. \end{cases}$$

Thus, mathematically one has given sense to properties in 1, 2 and 3. However, the differentiability of δ and H was yet to be made precise, since classically, every differentiable function is continuous.

In 1944, G. Choquet and J. Deny published a work on polyharmonic functions in two dimensions (cf. [CD44]). Laurent Schwartz, in an attempt to generalise the work of Choquet and Deny in higher dimensions, published an article in 1945 (cf. [Sch45]). With this article, the theory of distributions was discovered and also settled the issue raised in properties 4 and 5, in addition to 1, 2 and 3.

Dirac won the physics Nobel prize (1933) and Schwartz won the Fields medal (1950) for their respective work.

1.2 Motivation

The theory of distribution is a concept that generalises the notion of function, hence is also called *generalised functions*. But the beauty of the theory lies in the fact that this new notion admits the concept of differentiation and every distribution is infinitely differentiable. The need for such a generalisation was felt at various instances in the history of mathematics.

1.2.1 Case 1

Recall that the wave equation $u_{tt}(x,t) = c^2 u_{xx}(x,t)$ on $\mathbb{R} \times (0,\infty)$, describing the vibration of an infinite string, has the general solution $u(x,t) = F(x+ct) + G(x-ct)$. In reality, it happens that even if F and G are *not* twice differentiable, they are "solution" to the wave equation. For instance, consider F to be the Heaviside function H (cf. (1.1.2)) and $G \equiv 0$ or vicev-ersa. Now, how do we mathematically describe this nature of "solution"?

1.2.2 Case 2

Consider the Burger's equation

$$\begin{cases} u_t(x,t) + u(x,t)u_x(x,t) & = 0 \quad & x \in \mathbb{R} \text{ and } t \in (0,\infty) \\ u(x,0) & = \phi(x) \quad & x \in \mathbb{R} \end{cases}$$

and the corresponding characteristic equations:

$$\frac{dx(r,s)}{ds} = z, \quad \frac{dt(r,s)}{ds} = 1, \text{ and } \frac{dz(r,s)}{ds} = 0$$

with initial conditions,

$$x(r,0) = r, \quad t(r,0) = 0, \text{ and } z(r,0) = \phi(r).$$

Solving the ODE corresponding to z, we get $z(r,s) = c_3(r)$ with initial conditions $z(r,0) = c_3(r) = \phi(r)$. Thus, $z(r,s) = \phi(r)$. Using this in the ODE of x, we get

$$\frac{dx(r,s)}{ds} = \phi(r).$$

Solving the ODE's, we get

$$x(r,s) = \phi(r)s + c_1(r), \quad t(r,s) = s + c_2(r)$$

with initial conditions

$$x(r,0) = c_1(r) = r$$

and

$$t(r,0) = c_2(r) = 0.$$

Therefore,

$$x(r,s) = \phi(r)s + r, \text{ and } t(r,s) = s.$$

Eliminating s, we get the characteristic curves on the $x - t$ plane on which we know u is constant. Thus, the characteristic curves are $x = \phi(x_0)t + x_0$ and value of u on this curve is $\phi(x_0)$. Hence $u(\phi(x_0)t + x_0, t) = \phi(x_0)$.

Suppose we choose ϕ to be the function

$$\phi(x) = \begin{cases} 1 & x < 0 \\ 1 - x & 0 \leq x \leq 1 \\ 0 & 1 \leq x. \end{cases}$$

Then the characteristic curves are

$$x = \begin{cases} t + x_0 & x_0 < 0 \\ (1 - x_0)t & 0 \leq x_0 \leq 1 \\ x_0 & 1 \leq x_0. \end{cases}$$

Note that the characteristic curves are intersecting. This is no issue, as long as, the solution u which is constant on these curves takes the same constant in each characteristic curve. Unfortunately, that is not the case here:

$$u(x,t) = \begin{cases} 1 & x < t \\ \frac{x}{t+1} & 0 \leq 1 - \frac{x}{t} \leq 1 \\ 0 & 1 \leq x. \end{cases}$$

This situation is, usually, referred to as *shock waves*.

1.3 Space of Test Functions

The space of test functions will be the argument for the 'generalised function' or distributions, a notion to be introduced subsequently. Let $(a,b) \subset \mathbb{R}$ be a non-empty open subset, not necessarily bounded. Let $C_c^\infty(a,b)$ denote the class of all infinitely differentiable functions on Ω with compact support in (a,b). The genius of L. Schwartz is the choice of this function space. The motivation of this choice is the following: If $f \in C^1(\Omega)$ and $\phi \in C_c^\infty(a,b)$, then by classical integration by parts we have

$$\int_a^b \phi(x) f'(x)\,dx = -\int_a^b f(x)\phi'(x)\,dx. \tag{1.3.1}$$

There are no boundary integrals above because ϕ has compact support in (a,b) and vanishes at the end-points a and b. This explains the choice of compact support. Suppose f was chosen from $C^k(a,b)$, then above integration by parts could be repeated k times to get

$$\int_a^b \phi(x) f^{(k)}(x)\,dx = (-1)^k \int_a^b f(x)\phi^{(k)}\,dx.$$

Observe that the maps $\phi \mapsto \int_a^b f\phi\,dx$ and $\phi \mapsto \int_a^b f\phi'\,dx$ are linear on $C_c^\infty(a,b)$. If there exists a suitable complete topology on $C_c^\infty(a,b)$ such that these linear maps are continuous, then f can be identified with a continuous linear functional on $C_c^\infty(a,b)$ given by

$$\phi \mapsto \int_a^b f\phi\,dx \tag{1.3.2}$$

and the derivative of f, f', can be identified with the mapping ϕ mapped to the right hand side of (1.3.1), even if $f \notin C^1(\Omega)$, as long as the integrals make sense.

Let $\Omega \subset \mathbb{R}^n$ be a non-empty open subset. Let $C_c^\infty(\Omega)$ denote the class of all infinitely differentiable functions on Ω with compact support. Observe that $C_c^\infty(\Omega)$ is a vector space under usual addition and scalar multiplication of real-valued functions. Of course, the zero function is in $C_c^\infty(\Omega)$. We know that polynomials, trigonometric functions, exponential are all C^∞ functions. Any function in $C_c^\infty(\Omega)$ falls to zero within a compact set.

Exercise 1. Show that any non-zero function in $C_c^\infty(\Omega)$, for $\Omega \subset \mathbb{R}^n$, cannot be an analytic function.

Proof. If ϕ were analytic then, for all $a \in \Omega$, we have the Taylor expansion $\phi(x) = \sum_{k=0}^\infty \frac{\phi^{(k)}(a)}{k!}(x-a)^k$. Analyticity of ϕ implies that the radius of convergence of the Taylor series at a is infinite. For $\phi \in C_c^\infty(\Omega)$, choose $a \notin \operatorname{supp}(\phi)$. The Taylor's expansion around a imply that ϕ is zero function, which is a contradiction. $\square$

Therefore, no non-zero analytic function can sit in $C_c^\infty(\Omega)$ because zero function is the only analytic function with compact support. The function $f(x) = e^{ax}$ is a smooth analytic function on $\mathbb{R}$ for all $a \in \mathbb{R}$. Do we have non-zero smooth (non-analytic) functions in $C_c^\infty(\Omega)$?

1.3.1 Smooth Non-Analytic (Bump) Functions

We are looking for a smooth non-analytic functions in $C_c^\infty(\Omega)$. This can be viewed as gluing of the zero function, along the boundary of a compact set, with smooth function inside the compact set. The question is : can they be glued such that the smoothness is preserved in all of domain.

Two continuous maps can be glued together to result in a continuous map. In a different view point: consider a connected subset (interval) $E = (a, b)$ of $\mathbb{R}$ and the function f on E^c which takes zero on the connected component, $x \leq a$ and 1 on the connected component $x \geq b$. This is a step function. A continuous extension is trivial: Consider the linear map $x \mapsto \frac{x-a}{b-a}$ from $[a, b]$ to $[0, 1]$ which glues. However, the function is not differentiable at a and b. Similarly, gluing of the smooth maps $x \mapsto -x$ on $(-\infty, 0]$ and $x \mapsto x$ on $[0, \infty)$ yields the non-differentiable map $|x|$ on $\mathbb{R}$. Thus the smoothness property need not be preserved when glued. A tool to accomplish the gluing of smooth functions in a smooth way is the partition of unity (cf. Appendix **??**).

Recall that we wish to define f on E such that the extended function is smooth or C^∞ on $\mathbb{R}$. We wish to glue a function in (a, b) such that on $\mathbb{R}$ it is smooth. Note that the smoothness can break only at the end-points a and b. Thus, if we can find a smooth function whose zero derivatives are at a point, then we hope to transit smoothly from 0 to 1 in (a, b).

Recall that the positive function $f(x) = e^{1/x}$ behaves badly at $x = 0$. From the right side it approaches $+\infty$ and from left side it approaches zero. However, for $x \neq 0$, $e^{1/x}$ is infinitely differentiable (smooth). In fact, the k-th derivative of $e^{1/x}$ is $P_k(1/x)e^{1/x}$. The proof is by induction. For $k = 0$, $P_0(t) \equiv 1$, the constant function 1. Let us assume $f^{(k)}(x) = P_k(1/x)e^{1/x}$ for $x \neq 0$. Then,

$$f^{(k+1)}(x) = P_k(1/x)\left(\frac{-1}{x^2}\right)e^{1/x} + e^{1/x}P'_k(1/x)\left(\frac{-1}{x^2}\right) = P_{k+1}(1/x)e^{1/x},$$

where $P_{k+1}(t) = -t^2(P_k(t) + P'_k(t))$.

Exercise 2. Show that the k-th derivative of $e^{1/x}$ for $x \neq 0$ is

$$\frac{d^k}{dx^k}(e^{1/x}) = e^{1/x}\left[(-1)^k \sum_{i=1}^{k} \binom{k}{i}\binom{k-1}{i-1}(k-i)!x^{-k-i}\right].$$

The coefficients of the polynomial appearing in the derivative are called the *Lah* numbers, after Ivo Lah who encountered these numbers in actuarial science.

Example 1.1. Consider the non-negative function $f_0 : (-\infty, 0] \to [0,1]$ defined as

$$f_0(x) = \begin{cases} \exp(1/x) & \text{if } x < 0 \\ 0 & \text{if } x = 0. \end{cases}$$

Example 1.2. A variant of the above example is

$$f_1(x) = \begin{cases} \exp(-1/|x|) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

Example 1.3. Another variant is the function $f_2 : \mathbb{R} \to \mathbb{R}$ defined as

$$f_2(x) = \begin{cases} \exp(-1/x) & \text{if } x > 0 \\ 0 & \text{if } x \leq 0. \end{cases}$$

It is clear that $0 \leq f_i(x) < 1$ and f_i is infinitely differentiable for all $x \neq 0$ and $i = 0, 1, 2$. We need to check differentiability only at $x = 0$. We shall do it only for $f = f_2$ as the proof is similar for other cases. Of course, the left side limit of f and its derivative is zero at $x = 0$. Thus, we consider

$$f'(0) = \lim_{h \to 0^+} \frac{f(h) - f(0)}{h} = \lim_{h \to 0^+} \frac{\exp(-1/h)}{h}.$$

For all $h > 0$ and any fixed positive integer $m > 0$, we have

$$\frac{1}{h} = h^m \frac{1}{h^{m+1}} \leq h^m (m+1)! \sum_{i=0}^{\infty} \frac{1}{i!} \left(\frac{1}{h}\right)^i = h^m (m+1)! \exp(1/h).$$

Thus,

$$\frac{\exp(-1/h)}{h} \leq h^m (m+1)!$$

and the larger the m, the sharper the estimate. Hence,

$$f'(0) = \lim_{h \to 0^+} \frac{\exp(-1/h)}{h} = 0.$$

The same argument follows for any $k+1$ derivative of f because

$$f^{(k+1)}(0) = \lim_{h \to 0^+} \frac{f^{(k)}(h) - f^{(k)}(0)}{h} = \lim_{h \to 0^+} \frac{P_{k-1}(h) \exp(-1/h)}{h^{2k+1}} = 0.$$

Therefore, $f \in C^\infty(\mathbb{R})$. We shall now observe that f is not analytic at $x = 0$. The Taylor series of f at $x = 0$,

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k = 0,$$

converges to the zero function for all $x \in \mathbb{R}$. But for $x > 0$, we know that $f(x) > 0$ and hence do not converge to the Taylor series at $x = 0$. Thus, f is not analytic.

Example 1.4. The function $f : \mathbb{R} \to \mathbb{R}$ defined as

$$f_3(x) = \begin{cases} \exp(-1/x^2) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

is infinitely differentiable but not analytic.

Example 1.5. The function $f : \mathbb{R} \to \mathbb{R}$, referred to as the Cauchy's exponential function, defined as

$$f_4(x) = \begin{cases} \exp(-1/x^2) & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

is infinitely differentiable but not analytic.

One can, in fact, construct bump functions where the function takes the value 1 on a given subset of the support of f.

Lemma 1.3.1 (Smooth Cut-off Function). *For any interval (a, b) of $\mathbb{R}$, there is a decreasing smooth function $g : \mathbb{R} \to \mathbb{R}$ such that $0 < g(x) < 1$ for all $x \in (a, b)$ and*

$$g(x) = \begin{cases} 1 & x \leq a \\ 0 & x \geq b. \end{cases}$$

Proof. Choose $f = f_i$, for $i = 2$ or $i = 4$ in the above examples. This is precisely the choice of function whose zero derivatives are all at a point. Note that the function $f(b-x) = 0$ for all $x \geq b$ and the function $f(x-a) = 0$ for all $x \leq a$. Consider the non-zero positive function

$$h(x) = f(b - x) + f(x - a)$$

on $\mathbb{R}$ which is

$$h(x) = \begin{cases} f(b-x) & x \leq a \\ f(x-a) & x \geq b. \end{cases}$$

Since $f(b-x) < f(b-x)+f(x-a)$, addition of positive function. Moreover, h is smooth. Define the function

$$g(x) = \frac{f(b-x)}{f(b-x)+f(x-a)}$$

and it satisfies all the desired properties. $\square$

In particular, when $f = f_2$, we have

$$g(x) = \frac{e^{-1/(b-x)}}{e^{-1/(b-x)} + e^{-1/(x-a)}}.$$

Note that the increasing function

$$1 - g(x) = \frac{f(x-a)}{f(b-x)+f(x-a)}$$

is identically zero for $x \leq a$ and one on $x \geq b$ satisfying the other properties of above lemma.

Lemma 1.3.2 (Smooth Bump Function)**.** *For any positive $a, b \in \mathbb{R}$ such that $a < b$, there is a $h \in C_c^\infty(\mathbb{R})$ such that $0 \leq h(x) \leq 1$, for all $x \in \mathbb{R}$, $h \equiv 1$ on $[-a, a]$ and supp(h)$= [-b, b]$.*

Proof. Recall the function g obtained in above lemma. Then the function $x \mapsto g(|x|)$ is identically one in $[-a, a]$ and zero outside $(-b, b)$. The desired function h is obtained by setting $h(x) = g(|x|)$. The function h is smooth because it is composed with the function $|x|$, which is smooth except at $x = 0$. But $h \equiv 1$ around the origin. $\square$

Lemma 1.3.3 (Smooth Bump Function on $\mathbb{R}^n$)**.** *For any $0 < a < b$, there is a $h \in C_c^\infty(\mathbb{R}^n)$ with range in $[0, 1]$ such that $0 \leq h(x) \leq 1$, for all $x \in \mathbb{R}$, $h \equiv 1$ on the closed ball $B_a(0)$ and supp(h) is the closure of $B_b(0)$.*

Proof. The same function $h(x) = g(|x|)$ works. $\square$

Theorem 1.3.4. *For any compact subset K of $\mathbb{R}^n$, then there exists a $\phi \in C_c^\infty(\mathbb{R}^n)$ such that $\phi \equiv 1$ on K.*

Proof. One can chose an open set $U \supset K$ such that $\overline{U}$ is compact. Then, U and K^c form an open cover of $\mathbb{R}^n$. By Theorem **??**, there is a partition of unity subordinate to the given cover of $\mathbb{R}^n$. Thus, there are smooth non-negative functions ϕ and ψ on $\mathbb{R}^n$ such that supp(ϕ) is in U, supp(ψ) is in K^c and $\phi + \psi = 1$ on $\mathbb{R}^n$. On K, $\psi = 0$ and hence $\phi = 1$. Also, supp(ϕ) is a closed subset of a compact subset $\overline{U}$, hence ϕ has compact set. Thus, ϕ is the desired function. □

1.3.2 Mollifiers

The functions f_2 and f_4 in the above examples can be tweaked to construct a function in $C_c^\infty(\mathbb{R})$ such that supp(f)$= [-b, b]$ for any positive $b \in \mathbb{R}$. Note that this time we do not demand that the function takes constant value 1 in a subset of the support. Consider the transformation $g_2^b(x) = f_2(1 - |x|/b)$. Thus,

$$g_2^b(x) = \begin{cases} \exp\left(\frac{-b}{b-|x|}\right) & \text{if } |x| < b \\ 0 & \text{if } |x| \geq b. \end{cases}$$

Also, using the function

$$f_5(x) = \begin{cases} \exp\left(\frac{-1}{1-x^2}\right) & \text{if } |x| < 1 \\ 0 & \text{if } |x| \geq 1 \end{cases}$$

one defines the function

$$g_5^b(x) = \begin{cases} \exp\left(\frac{-b^2}{b^2-|x|^2}\right) & \text{if } |x| < b \\ 0 & \text{if } |x| \geq b. \end{cases}$$

In fact, the functions g_2^b, g_5^b can be extended to n dimensions and are in $C_c^\infty(\mathbb{R}^n)$ with support in $\overline{B}(0; b)$, the disk with centre at origin and radius b.

We shall now introduce an important sequence of functions in $C_c^\infty(\mathbb{R}^n)$, called *mollifiers*. Recall the $C_c^\infty(\mathbb{R}^n)$ functions g_2^b and g_5^b introduced in the previous section. For $\varepsilon > 0$, we set $b = \varepsilon$ in g_5^b (or g_2^b) and

$$\begin{aligned} \int_{\mathbb{R}^n} g_5^\varepsilon(x)\, dx &= \int_{|x|<\varepsilon} \exp\left(\frac{-\varepsilon^2}{\varepsilon^2 - |x|^2}\right) dx \\ &= \varepsilon^n \int_{|y|<1} \exp\left(\frac{-1}{1 - |y|^2}\right) dy \quad \text{(by setting } y = x/\varepsilon) \\ &= \varepsilon^n c^{-1}, \end{aligned}$$

where

$$c^{-1} = \int_{|y|\leq 1} \exp\left(\frac{-1}{1-|y|^2}\right) dy.$$

Now, set $\rho_\varepsilon(x) = c\varepsilon^{-n} g_5^\varepsilon(x)$, equivalently,

$$\rho_\varepsilon(x) = \begin{cases} c\varepsilon^{-n}\exp\left(\frac{-\varepsilon^2}{\varepsilon^2-|x|^2}\right) & \text{if } |x| < \varepsilon \\ 0 & \text{if } |x| \geq \varepsilon. \end{cases} \tag{1.3.3}$$

Note that $\rho_\varepsilon \geq 0$ and is in $C_c^\infty(\mathbb{R}^n)$ with support in $B(0;\varepsilon)$. The sequence $\{\rho_\varepsilon\}$ is an example of mollifiers, a particular case of the *Dirac Sequence*.

Definition 1.3.5. *A sequence of functions $\{\rho_k\}$, say on $\mathbb{R}^n$, is said to be a* Dirac Sequence *if*

(i) $\rho_k \geq 0$ *for all k.*

(ii) $\int_{\mathbb{R}^n} \rho_k(x)\, dx = 1$ *for all k.*

(iii) For every given $r > 0$ and $\varepsilon > 0$, there exists a $N_0 \in \mathbb{N}$ such that

$$\int_{\mathbb{R}^n \setminus B(0;r)} \rho_k(x)\, dx < \varepsilon, \quad \forall k > N_0.$$

The connection between the sequence of mollifiers and Dirac delta function will become evident in the sequel. The notion of mollifiers is also an example for the *approximation of identity* concept in functional analysis and ring theory.

Definition 1.3.6. *An* approximate identity *is a sequence (or net) $\{\rho_k$ in a Banach algebra or ring (possible with no identity), $(X, \star)$ such that for any element a in the algebra or ring, the limit of $a \star \rho_k$ (or $\rho_k \star a$) is a.*

1.3.3 Multi-Index Notations

A *multi-index* $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a n-tuple where α_i, for each $1 \leq i \leq n$, is a non-negative integer. Let $|\alpha| := \alpha_1 + \ldots + \alpha_n$. If α and β are two multi-indices, then $\alpha \leq \beta$ means $\alpha_i \leq \beta_i$, for all $1 \leq i \leq n$, and $\alpha \pm \beta = (\alpha_1 \pm \beta_1, \ldots, \alpha_n \pm \beta_n)$. Also, $\alpha! = \alpha_1! \ldots \alpha_n!$ and, for any $x \in \mathbb{R}^n$, $x^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}$. The multi-index notation, introduced by L. Schwartz, is quite handy in representing multi-variable equations in a concise form. For instance, a k-degree polynomial in n-variables can be written as

$$\sum_{|\alpha|\leq k} a_\alpha x^\alpha.$$

The partial differential operator of *order* α is denoted as

$$D^\alpha = \frac{\partial^{\alpha_1}}{\partial {x_1}^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial {x_n}^{\alpha_n}} = \frac{\partial^{|\alpha|}}{\partial {x_1}^{\alpha_1} \ldots \partial {x_n}^{\alpha_n}}.$$

One adopts the convention that, among the similar components of α, the order in which differentiation is performed is irrelevant. This is not a restrictive convention because the independence of order of differentiation is valid for smooth[1] functions. For instance, if $\alpha = (1, 1, 2)$ then one adopts the convention that

$$\frac{\partial^4}{\partial x_1 \partial x_2 \partial {x_3}^2} = \frac{\partial^4}{\partial x_2 \partial x_1 \partial {x_3}^2}.$$

If $|\alpha| = 0$, then $D^\alpha f = f$. For each $k \in \mathbb{N}$, $D^k u(x) := \{D^\alpha u(x) \mid |\alpha| = k\}$. The $k = 1$ case is

$$\begin{aligned} D^1 u(x) &= \left(D^{(1,0,\ldots,0)} u(x), D^{(0,1,0,\ldots,0)} u(x), \ldots, D^{(0,0,\ldots,0,1)} u(x)\right) \\ &= \left(\frac{\partial u(x)}{\partial x_1}, \frac{\partial u(x)}{\partial x_2}, \ldots, \frac{\partial u(x)}{\partial x_n}\right). \end{aligned}$$

The operator D^1 is called the *gradient* operator and is denoted as D or ∇. Thus, $\nabla u(x) = (u_{x_1}(x), u_{x_2}(x), \ldots, u_{x_n}(x))$. The $k = 2$ case is

$$D^2 u(x) = \begin{pmatrix} \frac{\partial^2 u(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 u(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 u(x)}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 u(x)}{\partial x_2 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 u(x)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 u(x)}{\partial x_n^2} \end{pmatrix}_{n\times n}.$$

The matrix $D^2 u$ is called the *Hessian* matrix. Observe that the Hessian matrix is symmetric due to the independence hypothesis of the order in which partial derivatives are taken. Further, for a k-times differentiable function u, the n^k-*tensor* $D^k u(x) := \{D^\alpha u(x) \mid |\alpha| = k\}$ may be viewed as a map from $\mathbb{R}^n$ to $\mathbb{R}^{n^k}$.

[1]smooth, usually, refers to as much differentiability as required.

Example 1.1. Let $u(x,y) : \mathbb{R}^2 \to \mathbb{R}$ be defined as $u(x,y) = ax^2 + by^2$. Then

$$\nabla u = (u_x, u_y) = (2ax, 2by)$$

and

$$D^2 u = \begin{pmatrix} u_{xx} & u_{yx} \\ u_{xy} & u_{yy} \end{pmatrix} = \begin{pmatrix} 2a & 0 \\ 0 & 2b \end{pmatrix}.$$

Observe that $\nabla u : \mathbb{R}^2 \to \mathbb{R}^2$ and $D^2 u : \mathbb{R}^2 \to \mathbb{R}^4 = \mathbb{R}^{2^2}$.

Thus, the magnitude of $D^k u(x)$ is

$$|D^k u(x)| := \left(\sum_{|\alpha|=k} |D^\alpha u(x)|^2 \right)^{\frac{1}{2}}.$$

In particular, $|\nabla u(x)| = (\sum_{i=1}^n u_{x_i}^2(x))^{\frac{1}{2}}$ or $|\nabla u(x)|^2 = \nabla u(x) \cdot \nabla u(x)$ and $|D^2 u(x)| = (\sum_{i,j=1}^n u_{x_i x_j}^2(x))^{\frac{1}{2}}$.

1.3.4 Topology on $C(\Omega)$, $C_0(\Omega)$ and $C_c(\Omega)$

Recall that (cf. § **??**) for an open subset $\Omega \subset \mathbb{R}^n$ and for any $\phi \in C_b(\Omega)$, $\|\phi\| = \sup_{x\in\Omega} |\phi(x)|$ defines a norm, called the *uniform norm* or *sup-norm*, in $C_b(\Omega)$. The space $C_b(\Omega)$ endowed with the uniform norm is a Banach space and a sequence ϕ_k converges to ϕ in the uniform norm is said to converge *uniformly* in Ω.

If Ω is a compact subset, the norm on the space of continuous functions on Ω, $C(\Omega)$, is defined as

$$\|\phi\|_{\infty,\Omega} := \max_{x\in\Omega} |\phi(x)|$$

is complete. Thus $C(\Omega)$ is a Banach space w.r.t the uniform norm. For a non-compact subset Ω of $\mathbb{R}^n$, the space $C(\Omega)$ is not normable but are metrizable[2]. They form a locally convex complete metric space called *Fréchet space*. We describe this metric briefly in the following paragraph.

For any *open* subset Ω of $\mathbb{R}^n$, there is a sequence K_j of non-empty compact subsets of Ω such that $\Omega = \cup_{j=0}^{\infty} K_j$ and $K_j \subset \mathrm{Int}(K_{j+1})$, for all j (exhaustion

[2] Any norm space is a metric space but the converse is not true always because $d(x,0)$ may fail to satisfy the properties of norm. A metric d induces a norm if $d(x+z, y+z) = d(x,y)$ and $d(\alpha x, \alpha y) = |\alpha| d(x,y)$.

of an open set by compact sets, cf. Lemma **??**). This property is called the *σ-compactness* of Ω. We define a countable family of semi-norms on $C(\Omega)$ as

$$p_j(\phi) = \|\phi\|_{\infty,K_j}.$$

Note that $p_0 \leq p_1 \leq p_2 \leq \ldots$. The sets

$$\{\phi \in C(\Omega) \mid p_j(\phi) < 1/j\}$$

form a local base for $C(\Omega)$. The metric induced by the family of semi-norms on $C(\Omega)$ is

$$d(\phi,\psi) = \max_{j \in \mathbb{N}\cup\{0\}} \frac{1}{2^j} \frac{p_j(\phi-\psi)}{1+p_j(\phi-\psi)}$$

and the metric is complete and $C(\Omega)$ is a Fréchet space. This is precisely the topology of compact convergence (uniform convergence on compact sets) or the compact-open topology in this case. If $\{\phi_m\}$ is a Cauchy sequence w.r.t d then $p_j(\phi_m - \phi_\ell) \to 0$, for all j and as m, ℓ tends to infinity. Thus $\{\phi_m\}$ converges uniformly on K_j to some $\phi \in C(\Omega)$. Then it is easy to see that $d(\phi, \phi_m) \to 0$. The metric defined above is called the *Fréchet* metric and is equivalent to the metric

$$d(\phi,\psi) = \sum_{j=0}^{\infty} \frac{1}{2^j} \frac{p_j(\phi-\psi)}{1+p_j(\phi-\psi)}.$$

The Fréchet space may be seen as a countable limit of Banch spaces. In our case, the Banach spaces are $C(K_i)$ w.r.t the uniform norm, the restriction of the semi-norm.

Exercise 3. Show that the topology given in $C(\Omega)$ is independent of the choice the exhaustion compact sets $\{K_j\}$ of Ω.

We say a function $\phi : \Omega \to \mathbb{R}$ vanishes at infinity (or on the boundary), if for every $\varepsilon > 0$ there exists a compact set $K \subset \Omega$ (depending on ε) such that $|\phi(x)| < \varepsilon$ for all $x \in \Omega \setminus K$. Let $C_0(\Omega)$ be the set of all continuous functions on Ω vanishing at infinity (or on the boundary). Let $C_c(\Omega)$ be the set of all continuous functions with compact support in Ω. Let $C_b(\Omega)$ be the space of all bounded continuous functions on Ω. We have the inclusion $C_c(\Omega) \subset C_0(\Omega) \subset C_b(\Omega) \subset C(\Omega)$. One may assign the uniform norm on $C_b(\Omega)$, $C_0(\Omega)$ and $C_c(\Omega)$. Under the uniform norm $C_c(\Omega)$ is dense in $C_0(\Omega)$ which is a closed subspace of the Banach space $C_b(\Omega)$.

We shall construct a complete (non-metrizable) topology on $C_c(\Omega)$. For every compact subset $K \subset \Omega$, let $D(K)$ denote the class of all continuous functions in Ω such that their support is in K. The space $D(K)$ is a Banach space w.r.t the uniform norm. Note that

$$C_c(\Omega) = \cup_{K \subset \Omega} D(K)$$

where the union is over all compact subsets of Ω. We declare a map T on $C_c(\Omega)$ is continuous if T restricted to $D(K)$, for each compact $K \subset \Omega$, is continuous. Such a topology is called the *inductive limit topology* of $D(K)$ with uniform norm. We say a sequence $\{\phi_m\}$ converges to ϕ w.r.t inductive limit topology if there exists a compact set K such that $\mathrm{supp}(\phi_m) \subset K$ and ϕ_m converges uniformly to ϕ. The space $C_c(\Omega)$ is complete with respect to the inductive limit topology.

1.3.5 Algebraic and Topological Dual of $C_c(\Omega)$

Definition 1.3.7. *A linear map $T : C_c(\Omega) \to \mathbb{R}$ is said to be* positive *if $T(\phi) \geq 0$, for all $\phi \geq 0$.*

Note that given a measure μ, the map

$$T(\phi) := \int_\Omega \phi \, d\mu \tag{1.3.4}$$

defines a positive, linear functional on $C_c(\Omega)$.

Theorem 1.3.8. *Let μ be a positive Radon measure on Ω and T be as defined in (1.3.4). Let, for every open subset $U \subset \Omega$, $S_U := \{\phi \in D(U) \mid 0 \leq \phi \leq 1\}$ and, for every compact subset $K \subset \Omega$, $S_K := \{\phi \in C_c(\Omega) \mid \phi \geq \chi_K\}$. Then*

$$\mu(U) = \sup_{\phi \in S_U} T(\phi) \quad \forall U \subset \Omega \tag{1.3.5}$$

and

$$\mu(K) = \inf_{\phi \in S_K} T(\phi) \quad \forall K \subset \Omega. \tag{1.3.6}$$

Theorem 1.3.9. *Let T be a positive, linear functional on $C_c(\Omega)$. Then there exists a unique positive Radon measure defined as in (1.3.5) or (1.3.6) and satisfies (1.3.4).*

Definition 1.3.10. *A linear map $T : C_c(\Omega) \to \mathbb{R}$ is* locally bounded *if, for every compact subset $K \subset \Omega$, there exists a positive constant $C_K > 0$ such that*

$$|T(\phi)| \leq C_K \|\phi\|_\infty \quad \forall \phi \in D(K).$$

Note that every positive linear functional on $C_c(\Omega)$ is locally bounded.

Theorem 1.3.11. *Let T be a locally bounded, linear functional on $C_c(\Omega)$. Then there exists two positive, linear functionals T^+ and T^- on $C_c(\Omega)$ such that $T = T^+ - T^-$.*

Theorem 1.3.12. *Let T be a locally bounded, linear functional on $C_c(\Omega)$. Then there exists two Radon measures μ_1 and μ_2 on Ω such that $T(\phi) = \int_\Omega \phi \, d\mu_1 - \int_\Omega \phi \, d\mu_2$, for all $\phi \in C_c(\Omega)$.*

Note that any linear continuous (bounded) functional $C_c(\Omega)$ is also locally bounded and, hence, the above result is true.

For a locally compact Hausdroff space Ω, recall (cf. § 1.3.4) that the spaces $C_c(\Omega) \subset C_0(\Omega) \subset C_b(\Omega)$ can be endowed with the uniform norm where the first inclusion is dense and $C_0(\Omega)$ is closed subspace of $C_b(\Omega)$. Further, one may also endow the inductive limit topology on $C_c(\Omega)$ which completes it.

Theorem 1.3.13. *Consider $C_b(\Omega)$ endowed with the uniform topology. Then there is an isometric isomorphism between the dual of $C_b(\Omega)$ and the space of bounded finitely additive (not necessarily countably additive) measures, i.e. for any continuous linear functional $T : C_b(\Omega) \to \mathbb{R}$ there is a unique bounded, finitely additive measure μ such that*

$$T(\phi) = \int_\Omega \phi(x) \, d\mu \quad \forall \phi \in C_b(\Omega).$$

This association $T \mapsto \mu$ defines an isometry, i.e., $\|T\| = |\mu|$.

Note that bounded countably additive measures are a subspace of bounded finitely additive measures.

Theorem 1.3.14 (Riesz-Alexandrov)**.** *Consider $C_0(\Omega)$ endowed with the uniform topology. Then there is an isometric isomorphism between the dual of*

$C_0(\Omega)$ and $\mathcal{M}_b(\Omega)$, i.e. for any continuous linear functional $T : C_0(\Omega) \to \mathbb{R}$ there is a unique finite regular Borel measure $\mu \in \mathcal{M}_b(\Omega)$ such that

$$T(\phi) = \int_\Omega \phi(x)\, d\mu \quad \forall \phi \in C_0(\Omega).$$

This association $T \mapsto \mu$ defines an isometry, i.e., $\|T\| = |\mu|$.

Note that due to the closed inclusion of $C_0(\Omega)$ in $C_b(\Omega)$ the dual space inclusion are not reversed. In fact they preserve the inclusion.

Theorem 1.3.15 (Riesz-Markov)**.** *Consider $C_c(\Omega)$ endowed with the inductive limit topology. Then there is an isometric isomorphism between the dual of $C_c(\Omega)$ and $\mathcal{R}(\Omega)$, i.e. for any continuous linear functional $T : C_c(\Omega) \to \mathbb{R}$ there is a unique Radon measure $\mu \in \mathcal{R}(\Omega)$ such that*

$$T(\phi) = \int_\Omega \phi(x)\, d\mu \quad \forall \phi \in C_c(\Omega).$$

This association $T \mapsto \mu$ defines an isometry, i.e., $\|T\| = |\mu|$.

1.3.6 Topology on $C^\infty(\Omega)$

Recall that our aim is to define a suitable topology in $C_c^\infty(\Omega)$ such that the operations described in (1.3.2) are continuous. For a compact Ω, the space of k-times differentiable functions on Ω, $C^k(\Omega)$, is a Banach space w.r.t the norm

$$\|\phi\|_{k,\Omega} := \sum_{|\alpha|=0}^{k} \|D^\alpha \phi\|_{\infty,\Omega}, \forall k \geq 1.$$

For any open subset Ω of $\mathbb{R}^n$, there is a sequence K_j of non-empty compact subsets of Ω such that $\Omega = \cup_{j=0}^\infty K_j$ and $K_j \subset \mathrm{Int}(K_{j+1})$, for all j (exhaustion of an open set by compact sets, cf. Lemma **??**). Similar to the Frëchet metric constructed on the space $C(\Omega)$, we shall construct one for $C^\infty(\Omega)$. We define a countable family of semi-norms on $C^\infty(\Omega)$ as

$$p_j(\phi) = \sum_{|\alpha|=0}^{j} \|D^\alpha \phi\|_{\infty,K_j} = \|\phi\|_{j,K_j}.$$

Again, as before, the family of semi-norms induces a locally convex complete metric on $C^\infty(\Omega)$ making it a Fréchet space.

Exercise 4. Show that the topology given in $C^\infty(\Omega)$ is independent of the choice the exhaustion compact sets $\{K_j\}$ of Ω.

The space $C_c^\infty(\Omega)$ is a subset of $C^\infty(\Omega)$ and the semi-norms defined in $C^\infty(\Omega)$ restricted to $C_c^\infty(\Omega)$ becomes a norm. For any $\phi \in C_c^\infty(\Omega)$, the family of norms ($j = 0, 1, 2, \ldots$)

$$\|\phi\|_j = \sum_{|\alpha|=0}^{j} \|D^\alpha \phi\|_{\infty,\Omega} = \|\phi\|_{j,\Omega}$$

induces the same topology as the one inherited from $C^\infty(\Omega)$. However, this norm induced topology on $C_c^\infty(\Omega)$ is not complete and its completion is $C^\infty(\Omega)$. For more details and proof of the semi-norm induced locally convex topology refer to [Rud91].

Exercise 5. Let $\phi \in C_c^\infty(\mathbb{R})$ with supp(ϕ)= $[0, 1]$ and $\phi > 0$ in $(0, 1)$. Then the sequence

$$\psi_m(x) = \sum_{i=1}^{m} \frac{1}{i}\phi(x - i) = \phi(x - 1) + \frac{1}{2}\phi(x - 2) + \ldots + \frac{1}{m}\phi(x - m)$$

is Cauchy in the topology induced by the norms, but $\lim \psi_m \notin C_c^\infty(\mathbb{R})$.

1.3.7 Inductive Limit Topology on $C_c^\infty(\Omega)$

We shall construct a complete (non-metrizable) topology on $C_c^\infty(\Omega)$ different from the one inherited from $C^\infty(\Omega)$. For every compact subset $K \subset \Omega$, let $D^\infty(K)$ denote the class of all functions in $C^\infty(\Omega)$ such that their support is in K. The space $D^\infty(K)$ is given the topology inherited from $C^\infty(\Omega)$, the same induced by the family of norms defined at the end of previous section,

$$\|\phi\|_{j,K} = \sum_{|\alpha|=0}^{j} \|D^\alpha \phi\|_{\infty,K} \quad \forall j \geq 0.$$

Exercise 6. $D^\infty(K)$ is a closed subspace of $C^\infty(\Omega)$ under the inherited topology of $C^\infty(\Omega)$.

Proof. For each $x \in \Omega$, define the functional $T_x : C^\infty(\Omega) \to \mathbb{R}$ defined as $T_x(\phi) = \phi(x)$. For each $x \in \Omega$, there is a j_0 such that $x \in K_j$ for all $j \geq j_0$. Then,

$$|T_x(\phi)| = |\phi(x)| \leq p_j(\phi) \quad \forall j \geq j_0.$$

The functional T_x is continuous[3] because uniform convergence implies pointwise convergence. The topology on $C^\infty(\Omega)$ is uniform convergence on compact subsets. Therefore the kernel of T_x,

$$\ker(T_x) := \{\phi \in C^\infty(\Omega) \mid T_x(\phi) = 0\},$$

is a closed subspace of $C^\infty(\Omega)$. Note that $\ker(T_x)$ is precisely those $\phi \in C^\infty(\Omega)$ such that $\phi(x) = 0$. We claim that

$$D^\infty(K) = \cap_{x\in K^c}\ker(T_x).$$

If $\phi \in D^\infty(K)$ then ϕ is in the intersection because $\phi(x) = 0$ for all $x \in K^c$. Conversely, If $\phi(x) = 0$ for all $x \in K^c$, then $\mathrm{supp}(\phi)\subseteq K$. Thus, for any compact subset K of Ω, we have our claim that $D^\infty(K)$ is an arbitrary intersection of closed sets. Thus, $D^\infty(K)$ is closed in $C^\infty(\Omega)$. □

Recall that K_j is a sequence of non-empty compact subsets of Ω such that $\Omega = \cup_{j=1}^\infty K_j$ and $K_j \subset \mathrm{Int}(K_{j+1})$, for all j. Note that

$$C_c^\infty(\Omega) = \cup_{j=1}^\infty D^\infty(K_j)$$

and $D^\infty(K_\ell) \subset D^\infty(K_m)$ for all $\ell < m$. With the given topology on these spaces, inherited from $C^\infty(\Omega)$, the inclusion map $I_{k\ell} : D^\infty(K_\ell) \to D^\infty(K_m)$ is continuous. This is because the local base in $D^\infty(K_m)$ is

$$\left\{\phi \in D^\infty(K_m) \mid \|\phi\|_m < \frac{1}{m}\right\}.$$

For any such ϕ in the local base, we have $\|\phi\|_\ell < 1/\ell$ and is in the local base of $D^\infty(K_\ell)$. Thus, we endow C_c^∞ with the finest[4] topology that makes the inclusion maps $I_j : D^\infty(K_j) \to C_c^\infty(\Omega)$ continuous, for all j. In other words, a set U in $C_c^\infty(\Omega)$ is said to be open if and only if $I_j^{-1}(U)$ is open in $D^\infty(K_j)$ for all $j \geq 1$. Such a topology is called the *inductive limit topology* with respect to $D^\infty(K_j)$ and the maps $I_{k\ell}$. The space $C_c^\infty(\Omega)$ is complete with respect to the inductive limit topology because any Cauchy sequence is Cauchy in $D^\infty(K_j)$, for some j. Since $D^\infty(K_j)$ is closed, the space $C_c^\infty(\Omega)$ is complete w.r.t the inductive limit topology. Though each $D^\infty(K)$ is metrizable, the space $C_c^\infty(\Omega)$ is not metrizable (cf. Exercise 8).

[3] Compare the functional T_x with Dirac distribution to be introduced later

[4] strongest or largest topology, the one with more open sets

Exercise 7. Every proper subspace of a topological vector space has empty interior.

Proof. Let X be a topological vector space and $V \subsetneq X$ be a vector subspace of X. We need to show that for $x \in V$ there is an open set U containing x such that $U \subset V$. It is enough to show the claim for $x = 0$ $(0 \in V)$ because if U contains 0, then $U + \{x\} \subset V$ is an open set (due to continuity of addition) containing x. Due to vector space structure $U + \{x\}$ is in V. For every $x \in X$, we can define a function $f_x : \mathbb{R} \to X$ as $f_x(\lambda) = \lambda x$. The function f_x is continuous because the scalar multiplication map from $\mathbb{R} \times X$ to X is continuous. Suppose that there is an open set V containing 0 such that $V \subset X$. Then $f_x^{-1}(V)$ will be an open set containing $0 \in \mathbb{R}$. Thus, for any $\lambda \in f_x^{-1}(V)$, we have $\lambda x \in V$. Thus, $x \in V$. This argument is true for all $x \in X$, thus $X \subset V$. This implies $X = V$, a contradiction. □

Exercise 8. The inductive limit topology on $C_c^\infty(\Omega)$ is not metrizable.

Proof. Recall that $C_c^\infty(\Omega) = \cup_{j=1}^\infty C^\infty(K_j)$, where each closed set $C^\infty(K_j)$ has empty interior (cf. Exercise 7). Therefore, the complete space $C_c^\infty(\Omega)$ is a countable union of no-where dense sets. If $C_c^\infty(\Omega)$ was metrizable then it would contradict the Baire's category theorem. □

Definition 1.3.16. *The space $C_c^\infty(\Omega)$ endowed with the inductive limit topology, and denoted as $\mathcal{D}(\Omega)$, is called the space of* test functions*.*

Exercise 9 (cf. Exercise 4). Show that the topology defined on $\mathcal{D}(\Omega)$ is independent of the choice of (exhaustion sets) the sequence of compact sets K_j of Ω.

The sequential characterisation of the inductive limit topology on $\mathcal{D}(\Omega)$ is given below. We define only for the zero converging sequence due to continuity of the addition operation.

Proposition 1.3.17. *A sequence of functions $\{\phi_m\} \subset \mathcal{D}(\Omega)$ converges to zero iff there exists a compact set $K \subset \Omega$ such that supp$(\phi_m) \subset K$, for all m, and ϕ_m and all its derivatives converge uniformly to zero on K.*

1.3.8 Regularization and Cut-off Technique

In this section, we show that that space of test functions is densely contained in L^p space. This observation is very useful since most conjectures can be checked for test functions and then carried on to the required function in L^p. As usual, let Ω be an open subset of $\mathbb{R}^n$. We already know the following results from the theory of Lebesgue measure on $\mathbb{R}^n$.

Theorem 1.3.18 (cf. Theorem **??**). *For $1 \le p < \infty$, the class of all simple*[5] *functions $S(\Omega)$ is dense in $L^p(\Omega)$.*

Theorem 1.3.19 (cf. Theorem **??**). *For $1 \le p < \infty$, $C_c(\Omega)$ is dense in $L^p(\Omega)$.*

Further, a step ahead, we show that the space of test functions, $C_c^\infty(\Omega)$ is densely contained in $L^p(\Omega)$. This result is established by using the *technique of regularization by convolution* introduced by Leray and Friedrichs.

Definition 1.3.20. *Let $f, g \in L^1(\mathbb{R}^n)$. The convolution $f * g$ is defined as,*

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\,dy \quad \forall x \in \mathbb{R}^n.$$

The integral on RHS is well-defined, since by Fubini's Theorem and the translation invariance of the Lebesgue measure, we have

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} |f(x-y)g(y)|\,dx\,dy = \int_{\mathbb{R}^n} |g(y)|\,dy \int_{\mathbb{R}^n} |f(x-y)|\,dx = \|g\|_1 \|f\|_1.$$

Thus, for a fixed x, $f(x-y)g(y) \in L^1(\mathbb{R}^n)$.

Theorem 1.3.21. *The convolution operation on $L^1(\mathbb{R}^n)$ is both commutative and associative.*

---

[5]By our definition, simple function is non-zero on a finite measure. A simple function ϕ is a non-zero function on $\mathbb{R}^n$ having the (canonical) form

$$\phi(x) = \sum_{i=1}^{k} a_i 1_{E_i}$$

with disjoint measurable subsets $E_i \subset \mathbb{R}^n$ with $\mu(E_i) < +\infty$ and $a_i \neq 0$, for all i, and $a_i \neq a_j$ for $i \neq j$.

Theorem 1.3.22 (Young's inequality). *Let $1 \leq p, q, r < \infty$ such that $(1/p) + (1/q) = 1 + (1/r)$. If $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$, then the convolution $f * g \in L^r(\mathbb{R}^n)$ and*

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

In particular, for $1 \leq p < \infty$, if $f \in L^1(\mathbb{R}^n)$ and $g \in L^p(\mathbb{R}^n)$, then the convolution $f * g \in L^p(\mathbb{R}^n)$ and

$$\|f * g\|_p \leq \|f\|_1 \|g\|_p.$$

Proposition 1.3.23. [6] *Let $f \in L^1(\mathbb{R}^n)$ and $g \in L^p(\mathbb{R}^n)$, for $1 \leq p \leq \infty$. Then*

$$supp(f * g) \subset \overline{supp(f) + supp(g)}$$

If both f and g have compact support, then support of $f * g$ is also compact. The convolution operation preserves smoothness.

Theorem 1.3.24. *Let $\Omega \subset \mathbb{R}^n$ be an open subset of $\mathbb{R}^n$ and let*

$$\Omega_\varepsilon := \{x \in \Omega \mid dist(x, \partial\Omega) > \varepsilon\}.$$

If $f \in L^1_{loc}(\Omega)$ then $f_\varepsilon := \rho_\varepsilon * f$ is in $C^\infty(\Omega_\varepsilon)$.

Proof. Fix $x \in \Omega_\varepsilon$. Consider

$$\begin{aligned}
\frac{f_\varepsilon(x + he_i) - f_\varepsilon(x)}{h} &= \frac{1}{h}\int_\Omega [\rho_\varepsilon(x + he_i - y) - \rho_\varepsilon(x - y)]\, f(y)\, dy \\
&= \int_{B_\varepsilon(x)} \frac{1}{h}[\rho_\varepsilon(x + he_i - y) - \rho_\varepsilon(x - y)] f(y)\, dy.
\end{aligned}$$

Now, taking $\lim_{h \to 0}$ both sides, we get

$$\begin{aligned}
\frac{\partial f_\varepsilon(x)}{\partial x_i} &= \lim_{h \to 0} \int_{B_\varepsilon(x)} \frac{1}{h}[\rho_\varepsilon(x + he_i - y) - \rho_\varepsilon(x - y)] f(y)\, dy \\
&= \int_{B_\varepsilon(x)} \frac{\partial \rho_\varepsilon(x - y)}{\partial x_i} f(y)\, dy \\
&\qquad \text{(interchange of limits is due to the uniform convergence)} \\
&= \int_\Omega \frac{\partial \rho_\varepsilon(x - y)}{\partial x_i} f(y)\, dy = \frac{\partial \rho_\varepsilon}{\partial x_i} * f.
\end{aligned}$$

Similarly, one can show that, for any tuple α, $D^\alpha f_\varepsilon(x) = (D^\alpha \rho_\varepsilon * f)(x)$. Thus, $u_\varepsilon \in C^\infty(\Omega_\varepsilon)$. □

[6] Refer Brezis for proof

Proposition 1.3.25. [7] *Let $f \in C_c^k(\mathbb{R}^n)$ $(k \geq 1)$ and let $g \in L^1_{loc}(\mathbb{R}^n)$. Then $f * g \in C^k(\mathbb{R}^n)$ and for all $|\alpha| \leq k$*

$$D^\alpha(f * g) = D^\alpha f * g = f * D^\alpha g.$$

Theorem 1.3.26 (Regularization technique)**.** *$C^\infty(\mathbb{R}^n)$ is dense in $C(\mathbb{R}^n)$ under the uniform convergence on compact sets topology.*

Proof. Let $g \in C(\mathbb{R}^n)$ and $K \subset \mathbb{R}^n$ be a compact subset. Note that g is uniformly continuous on K. Hence, for every $\eta > 0$, there exist a $\delta > 0$ (independent of x and dependent on K and η) such that $|g(x-y) - g(x)| < \eta$ whenever $|y| < \delta$ for all $x \in K$. For each $m \in \mathbb{N}$, set $\rho_m := \rho_{1/m}$, the sequence of mollifiers. Define $g_m := \rho_m * g$. Note that $g_m \in C^\infty(\mathbb{R}^n)$ $(D^\alpha g_m = D^\alpha \rho_m * g)$. Now, for all $x \in \mathbb{R}^n$,

$$\begin{aligned}
|g_m(x) - g(x)| &= \left| \int_{|y| \leq 1/m} g(x-y)\rho_m(y)\,dy - g(x) \int_{|y|\leq 1/m} \rho_m(y)\,dy \right| \\
&\leq \int_{|y|\leq 1/m} |g(x-y) - g(x)|\rho_m(y)\,dy
\end{aligned}$$

Hence, for all $x \in K$ and $m > 1/\delta$, we have

$$\begin{aligned}
|g_m(x) - g(x)| &\leq \int_{|y|<\delta} |g(x-y) - g(x)|\rho_m(y)\,dy \\
&\leq \eta \int_{|y|<\delta} \rho_m(y)\,dy = \eta
\end{aligned}$$

Since the δ is independent of $x \in K$, we have $\|g_m - g\|_\infty < \eta$ for all $m > 1/\delta$. Hence, $g_m \to g$ uniformly on K. $\square$

Theorem 1.3.27. *For any $\Omega \subseteq \mathbb{R}^n$, $C_c^\infty(\Omega)$ is dense in $C_c(\Omega)$ under the uniform topology.*

Proof. Let $g \in C_c(\Omega)$ and $K := \operatorname{supp}(g)$. One can view $C_c(\Omega)$ as a subset of $C_c(\mathbb{R}^n)$ under the following identification: Each $g \in C_c(\Omega)$ is extended to $\mathbb{R}^n$ as $\tilde{g}$

$$\tilde{g}(x) = \begin{cases} g(x) & x \in K \\ 0 & x \in \mathbb{R}^n \setminus K. \end{cases}$$

[7] Refer Brezis for proof

By Theorem 1.3.26, the sequence $g_m := \rho_m * \tilde{g}$ in $C^\infty(\mathbb{R}^n)$ converges to $\tilde{g}$ uniformly on every compact subsets of $\mathbb{R}^n$. Note that $\operatorname{supp}(g_m) \subset K + B(0; 1/m)$ is compact because K is compact. Since we want $g_m \in C_c^\infty(\Omega)$, we choose $m_0 \in \mathbb{N}$ such that $1/m_0 < \operatorname{dist}(K, \Omega^c)$. Thus, $\operatorname{supp}(g_m) \subset \Omega$ and $g_m \in C_c^\infty(\Omega)$, for all $m \geq m_0$. The proof of the uniform convergence of g_m to g on Ω is same as in Theorem 1.3.26. □

Corollary 1.3.28. *For any $\Omega \subseteq \mathbb{R}^n$, $C_c^\infty(\Omega)$ is dense in $C(\Omega)$ under the uniform convergence on compact sets topology.*

Theorem 1.3.29 (Regularization technique)**.** *The space $C^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$, for $1 \leq p < \infty$, under the p-norm.*

Proof. Let $f \in L^p(\mathbb{R}^n)$. For each $m \in \mathbb{N}$, set $\rho_m := \rho_{1/m}$, the sequence of mollifiers. Then the sequence $f_m := \rho_m * f$ is in $C^\infty(\mathbb{R}^n)$. Since $\rho_m \in L^1(\mathbb{R}^n)$, by Young's inequality, $f_m \in L^p(\mathbb{R}^n)$. We shall prove that f_m converges to f in p-norm. For any given $\varepsilon > 0$, by Theorem **??**, we choose a $g \in C_c(\mathbb{R}^n)$ such that $\|g - f\|_p < \varepsilon/3$. Therefore, by Theorem 1.3.27, there is a compact subset $K \subset \mathbb{R}^n$ such that $\|\rho_m * g - g\|_\infty < \varepsilon/3(\mu(K))^{1/p}$. Hence, $\|\rho_m * g - g\|_p < \varepsilon/3$. Thus, for sufficiently large m, we have

$$\begin{aligned}
\|f_m - f\|_p &\leq \|\rho_m * f - \rho_m * g\|_p + \|\rho_m * g - g\|_p + \|g - f\|_p \\
&< \|\rho_m * (f - g)\|_p + \frac{2\varepsilon}{3} \leq \|f - g\|_p \|\rho_m\|_1 + \frac{2\varepsilon}{3} \\
&< \frac{\varepsilon}{3} + \frac{2\varepsilon}{3} = \varepsilon.
\end{aligned}$$

The first term has been handled using Young's inequality. □

Theorem 1.3.30 (Cut-Off Technique)**.** *For $1 \leq p < \infty$ and $\Omega \subseteq \mathbb{R}^n$, $C_c^\infty(\Omega)$ is dense in $L^p(\Omega)$.*

Proof. Any $f \in L^p(\Omega)$ can be viewed as an element in $L^p(\mathbb{R}^n)$ under the extension

$$\tilde{f}(x) = \begin{cases} f(x) & x \in \Omega \\ 0 & x \in \Omega^c. \end{cases}$$

By Theorem 1.3.29, the sequence $f_m := \rho_m * \tilde{f}$ converges to $\tilde{f}$ in p-norm. The sequence $\{f_m\}$ may fail to have compact support in Ω because support of $\tilde{f}$ is not necessarily compact in Ω. To fix this issue, we shall multiply the sequence with suitable choice of test functions in $C_c^\infty(\Omega)$. Choose the sequence of exhaustion compact sets $\{K_m\}$ in Ω (cf. Lemma **??**). In particular, for $\Omega = \mathbb{R}^n$, we can choose $K_m = B(0;m)$. Note that $\Omega = \cup_m K_m$. Consider[8] $\{\phi_m\} \subset C_c^\infty(\Omega)$ such that $\phi_m \equiv 1$ on K_m and $0 \leq \phi_m \leq 1$, for all m. We extend ϕ_m by zero on Ω^c. Define $F_m := \phi_m f_m$ and, hence, $F_m \in C_c^\infty(\Omega)$. Also, $F_m = f_m$ on K_m and $|F_m| \leq |f_m|$ in $\mathbb{R}^n$. Thus,

$$\begin{aligned}
\|F_m - f\|_{p,\Omega} &= \|F_m - \tilde{f}\|_{p,\mathbb{R}^n} \leq \|\phi_m f_m - \phi_m \tilde{f}\|_{p,\mathbb{R}^n} + \|\phi_m \tilde{f} - \tilde{f}\|_{p,\mathbb{R}^n} \\
&\leq \|f_m - \tilde{f}\|_{p,\mathbb{R}^n} + \|\phi_m \tilde{f} - \tilde{f}\|_{p,\mathbb{R}^n}.
\end{aligned}$$

The first term converges to zero by Theorem 1.3.29 and the second term converges to zero by Dominated convergence theorem. □

The case $p = \infty$ is ignored in the above results, because the L^∞-limit of $\rho_m * f$ is continuous and we do have discontinuous functions in $L^\infty(\Omega)$.

Theorem 1.3.31 (Kolmogorov Compactness Criteria). *Let $p \in [1,\infty)$ and let A be a subset of $L^p(\mathbb{R}^n)$. Then A is relatively compact in $L^p(\mathbb{R}^n)$ iff the following conditions are satisfied:*

(i) A is bounded in $L^p(\mathbb{R}^n)$;

(ii) $\lim_{r\to+\infty} \int_{\{|x|>r\}} |f(x)|^p\,dx = 0$ uniformly with respect to $f \in A$;

(iii) $\lim_{h\to 0} \|\tau_h f - f\|_p = 0$ uniformly with respect to $f \in A$, where $\tau_h f$ is the translated function $(\tau_h f)(x) := f(x-h)$.

Proof. We shall prove the 'only if' part, i.e, (i), (ii), (iii) implies that A is relatively compact in $L^p(\mathbb{R}^n)$. Equivalently, we have to prove that A is precompact, which means that for any $\varepsilon > 0$, there exists a finite number of balls $B_\varepsilon(f_1), \dots, B_\varepsilon(f_k)$ which cover A. Let us choose $\varepsilon > 0$. By (ii) there exists a $r > 0$ such that

$$\int_{|x|>r} |f(x)|^p\,dx < \varepsilon \quad \forall f \in A.$$

Let $(\rho_n)_{n\in\mathbb{N}}$ be a mollifier. It follows from Theorem 1.3.30 that, for all $n \geq 1$ and $f \in L^p(\mathbb{R}^n)$

$$\|f - f * \rho_n\|_p^p \leq \int_{\mathbb{R}^n} \rho_n(y) \|f - \tau_y f\|_p^p\,dy.$$

[8]The type of functions, ϕ_k, are called cut-off functions

Hence
$$\|f - f * \rho_n\|_p \le \sup_{|y| \le \frac{1}{n}} \|f - \tau_y f\|_p.$$
By (iii), there exists an integer $N(\varepsilon) \in \mathbb{N}$ such that, for all $f \in A$,
$$\|f - f * \rho_{N(\varepsilon)}\|_p < \varepsilon.$$
On the other hand, for any $x, z \in \mathbb{R}^n$, $f \in L^p(\mathbb{R}^n)$ and $n \in \mathbb{N}$,
$$\begin{aligned} |(f * \rho_n)(x) - (f * \rho_n)(z)| &\le \int_{\mathbb{R}^n} |f(x-y) - f(z-y)| \rho_n(y)\, dy \\ &\le \|\tau_x \check{f} - \tau_z \check{f}\|_p \|\rho_n\|_q \\ &\le \|\tau_{x-z} f - f\|_p \|\rho_n\|_q. \end{aligned}$$
The last inequality follows from the invariance property of the Lebesgue measure. Moreover,
$$|(f * \rho_n)(x)| \le \|f\|_p \|\rho_n\|_q.$$
Let us consider the family $\mathcal{A} = \{f * \rho_{N(\varepsilon)} : B_r(0) \to \mathbb{R} \mid f \in A\}$. By using (i) and (iii), and Ascoli-Arzela result, we observe that $\mathcal{A}$ is relatively compact w.r.t the uniform topology on $B_r(0)$. Hence, there exists a finite set $\{f_1, \dots, f_k\} \subset A$ such that
$$\mathcal{A} \subset \cup_{i=1}^k B_{\varepsilon r^{-n/p}}(f_i * \rho_{N(\varepsilon)}).$$
Thus, for all $f \in A$, there exists some $j \in \{1, 2, \dots, k\}$ such that, for all $x \in B_r(0)$
$$|f * \rho_{N(\varepsilon)}(x) - f_j * \rho_{N(\varepsilon)}(x)| \le \varepsilon |B_r(0)|^{-1/p}.$$
Hence,
$$\begin{aligned} \|f - f_j\|_p &\le \left(\int_{|x|>r} |f|^p\, dx\right)^{1/p} + \left(\int_{|x|>r} |f_j|^p\, dx\right)^{1/p} \\ &\quad + \|f - f * \rho_{N(\varepsilon)}\|_p + \|f_j - f_j * \rho_{N(\varepsilon)}\|_p \\ &\quad + \|f * \rho_{N(\varepsilon)} - f_j * \rho_{N(\varepsilon)}\|_{p, B_r(0)}. \end{aligned}$$
The last term may be treated as follows:
$$\begin{aligned} \|f * \rho_{N(\varepsilon)} - f_j * \rho_{N(\varepsilon)}\|_{p, B_r(0)} &= \left(\int_{B_r(0)} |f * \rho_{N(\varepsilon)}(x) - f_j * \rho_{N(\varepsilon)}(x)|^p\, dx\right)^{1/p} \\ &\le \varepsilon |B_r(0)|^{-1/p} |B_r(0)|^{1/p} = \varepsilon. \end{aligned}$$

Finally,

$$\|f - f_j\|_p \leq 5\varepsilon$$

and, hence, A is precompact in $L^p(\mathbb{R}^n)$. $\square$

1.4 Space of Distributions

Definition 1.4.1. *A linear functional T on $\mathcal{D}(\Omega)$ is said to be* continuous *if inverse image of open sets of $\mathbb{R}$ are open in $\mathcal{D}(\Omega)$. A linear functional T on $\mathcal{D}(\Omega)$ is said to be* sequential continuous *if $T\phi_m \to 0$ in $\mathbb{R}$ whenever $\phi_m \to 0$ in $\mathcal{D}(\Omega)$.*

Exercise 10. For any compact set $K \subset \Omega$, the restriction to $C^\infty(K)$ of any continuous map on $C_c^\infty(\Omega)$ is also continuous on $C^\infty(K)$. Similarly, for any compact set $K \subset \Omega$, the restriction to $C^\infty(K)$ of any sequentially continuous map on $C_c^\infty(\Omega)$ is also sequentially continuous on $C^\infty(K)$.

Hint. For each compact set K there is an positive integer i_0 such that $K \subset K_i$, for all $i \geq i_0$. $\square$

It is enough to define for zero convergent sequences because addition operation is continuous. For a first countable space the notion of continuity and sequential continuity are equivalent. A Hausdorff topological vector space is metrizable iff it is first countable. We know that $\mathcal{D}(\Omega)$ is not metrizable (cf. Exercise 8) and hence cannot be first countable.

Theorem 1.4.2. *Let $T : \mathcal{D}(\Omega) \to \mathbb{R}$ be a linear map. Then the following are equivalent:*

(i) T is continuous, i.e., inverse image of open set in $\mathbb{R}$, under T, is open in $\mathcal{D}(\Omega)$.

(ii) For every compact subset $K \subset \Omega$, there exists a constant $C_K > 0$ and an integer $N_K \geq 0$ (both depending on K) such that

$$|T(\phi)| \leq C_K \|\phi\|_{N_K}, \quad \forall \phi \in C^\infty(K).$$

(iii) T is sequentially continuous.

Proof. ((i) $\implies$ (ii)): Let T be continuous on $\mathcal{D}(\Omega)$. Then, for any compact subset $K \subset \Omega$, the restriction of T to $C^\infty(K)$ is continuous (cf. Exercise 10). The inverse image of $(-c, c)$ under T is open set in $C^\infty(K)$ containing origin. Since $C^\infty(K)$ is first countable (normed space), there is a local base at 0. Thus, there is a N_K and for all $\phi \in C^\infty(K)$ such that $\|\phi\|_{N_K} \leq 1/N_K$, we have $|T(\phi)| < c$. Thus, for $\phi \in C^\infty(K)$,

$$\left| T\left(\frac{\phi}{N_K \|\phi\|_{N_K}} \right) \right| < c$$

and hence $|T(\phi)| < N_K c \|\phi\|_{N_K}$.

((ii) $\implies$ (iii)): Let $\phi_m \to 0$ in $\mathcal{D}(\Omega)$. By definition of the test function convergence, there is a compact set K such that $\|\phi_m\|_j \to 0$ for all j. Using the fact that $|T(\phi)| \leq C_K \|\phi\|_{N_K}$, we get $|T(\phi_m)| \to 0$. Thus $T(\phi_m) \to 0$.

((iii) $\implies$ (i)): Let T be sequentially continuous. Then, by Exercise 10, the restriction of T on $C^\infty(K)$, for every compact subset $K \subset \Omega$, is also sequentially continuous. But $C^\infty(K)$ is metrizable and hence T on $C^\infty(K)$ is continuous, for all compact K of Ω. Thus, T is continuous on $\mathcal{D}(\Omega)$. □

Definition 1.4.3. *A linear functional T on $\mathcal{D}(\Omega)$ is said to be a* distribution *on Ω, if for every compact subset $K \subset \Omega$, there exists a constant $C_K > 0$ and an integer $N_K \geq 0$ (both depending on K) such that*

$$|T(\phi)| \leq C_K \|\phi\|_{N_K}, \quad \forall \phi \in C^\infty(K).$$

By Theorem 1.4.2, the above definition is saying that any continuous linear functional on $\mathcal{D}(\Omega)$ is a distribution. The space of all distributions in Ω is denoted by $\mathcal{D}'(\Omega)$.

Definition 1.4.4. *If the N_K is independent of K, i.e., the same N is enough for all compact sets K, then the smallest such N is called the* order *of T. If there exist no such N, we say T is of* infinite *order.*

Exercise 11. Which of the following are distributions? If your answer is affirmation, give the order of the distribution. If your answer is in negation, give reasons. For $\phi \in \mathcal{D}(\mathbb{R})$, $T(\phi)$ is defined as:

(i) $\phi'(1) - \phi''(-2)$.

(ii) $\sum_{k=0}^{\infty} \phi^{(k)}(\pi)$.

(iii) $\sum_{k=0}^{\infty} \phi^{(k)}(k)$.

(iv) $\sum_{k=1}^{\infty} \frac{1}{k}\phi^{(k)}(k)$.

(v) $\int_{\mathbb{R}} \phi^2(x)\,dx$.

Proof. (i) This is a distribution of order 2.

(ii) This is not a distribution because it is not defined for all $\phi \in \mathcal{D}(\mathbb{R})$. For instance, choose $\phi(x) = e^{x-\pi}\psi(x)$, where $\psi \in \mathcal{D}(\mathbb{R})$ such that $\psi(x) = 1$ in a neighbourhood of π. Note that $\phi^{(k)}(\pi) = 1$, for all k and the sum is infinite.

(iii) This is a distribution of infinite order.

(iv) This is a distribution of infinite order.

(v) This is not a distribution, because it is nonlinear.

□

Exercise 12. Show that if $T \in \mathcal{D}'(\Omega)$ and ω is an open subset of Ω, then $T \in \mathcal{D}'(\omega)$.

Definition 1.4.5. *Two distributions $S, T \in \mathcal{D}'(\Omega)$ are said to be* equal *if $S(\phi) = T(\phi)$ for all $\phi \in \mathcal{D}(\Omega)$.*

1.4.1 Functions as Distributions

Definition 1.4.6. *We say a function f is* locally integrable *in Ω, denoted as $f \in L^1_{loc}(\Omega)$, if f is measurable and $\int_K |f(x)|\,dx < +\infty$, for every compact set $K \subset \Omega$.*

In other words, $f \in L^1_{\text{loc}}(\Omega)$ if $f \in L^1(K)$ for all compact subsets K of Ω. In particular, $L^1(\Omega) \subset L^1_{\text{loc}}(\Omega)$. This inclusion is strict because, if Ω is not of finite measure, then constant functions do not belong to $L^1(\Omega)$ but they are in $L^1_{\text{loc}}(\Omega)$.

Example 1.6. The function $\ln|x| \in L^1_{\text{loc}}(\mathbb{R})$. To see this it is enough to show that $\lim_{\varepsilon\to 0^+} \int_\varepsilon^1 |\ln x|\,dx$ exists and is finite. Consider

$$\lim_{\varepsilon\to 0^+} \int_\varepsilon^1 |\ln x|\,dx = \lim_{\varepsilon\to 0^+} -(x\ln x - x)\,|_\varepsilon^1 = -[-1 - \lim_{\varepsilon\to 0^+}(\varepsilon\ln(\varepsilon) - \varepsilon)] = 1.$$

Because $\varepsilon \ln(\varepsilon) \to 0$ as $\varepsilon \to 0^+$ (use L'Hospital's rule after proper substitution).

Example 1.7. The function $1/x \notin L^1_{\text{loc}}(\mathbb{R})$ (after assigning a real value at $x = 0$). Similarly, $e^{1/x} \notin L^1_{\text{loc}}(\mathbb{R})$. In both these cases the integral is not finite on a compact set containing origin.

Example 1.8. For any integer $n \geq 1$ and $\alpha > 0$, the function $|x|^{-\alpha} \in L^1_{\text{loc}}(\mathbb{R}^n)$, for all $0 < \alpha < n$, because

$$\int_{B_\rho(0)} |x|^{-\alpha}\, dx = \int_{S_\rho(0)} \int_0^\rho r^{-\alpha+n-1}\, dr\, d\sigma.$$

Thus, for $-\alpha + n - 1 > -1$ or $\alpha < n$, the integral is finite and is equal to $\rho^{2n-\alpha-1}\frac{\omega_n}{n-\alpha}$, where ω_n is the surface measure of the unit ball. Note that the function $|x|^{-\alpha}$ has a blow-up near 0.

The notion of locally integrable functions can be extended to L^p spaces, for all $p \geq 1$.

Definition 1.4.7. *For $1 \leq p < \infty$, we say a function f is* locally p-integrable *in Ω, denoted as $f \in L^p_{loc}(\Omega)$, if f is measurable and $\int_K |f(x)|^p\, dx < +\infty$, for every compact set $K \subset \Omega$.*

We, in general, will not treat separately the notion of locally p-integrable functions because any locally p-integrable function is locally integrable. i.e., if $f \in L^p_{\text{loc}}(\Omega)$ for all $1 < p < \infty$, then $f \in L^1_{\text{loc}}(\Omega)$. Because the Hölder's inequality implies that for any compact subset K of Ω,

$$\int_K |f(x)|\, dx \leq \left(\int_K |f|^p\, dx\right)^{1/p} (\mu(K))^{1/q} < +\infty.$$

Thus, $f \in L^1_{\text{loc}}(\Omega)$. Since any $L^p(\Omega)$, for $1 \leq p < \infty$, is in $L^p_{\text{loc}}(\Omega)$ and hence are in $L^1_{\text{loc}}(\Omega)$.

Exercise 13. If f is continuous on Ω, then $f \in L^1_{\text{loc}}(\Omega)$.

We shall now observe that to every locally integrable function one can associate a distribution. For any $f \in L^1_{\text{loc}}(\Omega)$, we define the functional T_f on $\mathcal{D}(\Omega)$ defined as,

$$T_f(\phi) = \int_\Omega f(x)\phi(x)\, dx.$$

The functional T_f is continuous on $\mathcal{D}(\Omega)$ (and, hence, is in $\mathcal{D}'(\Omega)$) because for every compact set K in Ω,

$$|T_f(\phi)| \leq \left(\int_K |f|\,dx\right) \|\phi\|_0, \quad \forall \phi \in C^\infty(K).$$

The distribution T_f is of zero order. We usually identify the distribution T_f with the function f that induces it. Hence, every continuous function and every L^p function, for $p \geq 1$, induces the distribution described above.

Exercise 14. If $f : \Omega \to \mathbb{R}$ is such that

$$\int_\Omega |f\phi|\,dx < +\infty \quad \forall \phi \in \mathcal{D}(\Omega)$$

then $f \in L^1_{\text{loc}}(\Omega)$.

Proof. Let K be a compact subset of Ω. Then there is a $\phi_K \in \mathcal{D}(\Omega)$ such that $\phi_K \equiv 1$ on K and $0 \leq \phi_K \leq 1$. Hence,

$$\int_K |f(x)|\,dx = \int_\Omega |f(x)| \cdot 1_K\,dx \leq \int_\Omega |f(x)\phi_K(x)|\,dx < +\infty.$$

Thus, $f \in L^1_{\text{loc}}(\Omega)$. The function 1_K denotes the characteristic function which takes 1 on K and zero on the K^c. □

The constant function 0 on Ω induces the zero distribution (the zero functional on $\mathcal{D}(\Omega)$).

Exercise 15. If $f \in C(\Omega)$ such that $T_f \equiv 0$, i.e., for all $\phi \in \mathcal{D}(\Omega)$,

$$\int_\Omega f\phi\,dx = 0$$

then $f \equiv 0$ in Ω.

Proof. Let us assume f is continuous. If $f \neq 0$ in Ω, then there is a $x_0 \in \Omega$ such that $f(x_0) = \lambda \neq 0$. Define the function $g = f/\lambda$ which is also continuous on Ω such that $g(x_0) = 1$. Note the $\int g\phi = 0$ for all $\phi \in \mathcal{D}(\Omega)$. By the continuity of g, there is an $r > 0$ such that, for all $x \in B_r(x_0) \subset \Omega$, $g(x) > 1/2$. Choose a test function $\psi \in \mathcal{D}(\Omega)$ such that supp(ψ)$\subset B_r(x_0)$

and $\int_\Omega \psi(x)\,dx = 1$. For instance, choose ψ to be the mollifier function $\psi(x) = \rho_r(x - x_0)$. Then, in particular,

$$0 = \int_\Omega g\psi\,dx = \int_{B_r(x_0)} g\psi\,dx > \frac{1}{2}\int_{B_r(x_0)} \psi\,dx = \frac{1}{2} > 0,$$

a contradiction. Thus, $g \equiv 0$ on Ω and hence $f \equiv 0$ on Ω. □

Exercise 16. If $f \in L^1_{\text{loc}}(\Omega)$ is non-negative, $f \geq 0$, such that $T_f \equiv 0$, i.e., for all $\phi \in \mathcal{D}(\Omega)$,

$$\int_\Omega f\phi\,dx = 0$$

then $f = 0$ a.e. in Ω.

Proof. Consider a compact subset K of Ω. By outer regularity of Lebesgue measure, for every $\varepsilon > 0$, there is an open set $U_\varepsilon \supset K$ such that $\mu(U_\varepsilon \backslash K) < \varepsilon$. Let $\{\phi_\varepsilon\}$ be a sequence of functions in $\mathcal{D}(\Omega)$ such that $\phi_\varepsilon \equiv 1$ on K, $\phi_\varepsilon \equiv 0$ on U_ε^c and $0 \leq \phi_\varepsilon \leq 1$. Note that ϕ_ε converge point-wise to the characteristic function 1_K, as ε goes to 0. Therefore, by dominated convergence theorem,

$$0 = \lim_{\varepsilon \to 0} \int_\Omega f\phi_\varepsilon\,dx = \int_\Omega f \cdot 1_K\,dx = \int_K f\,dx.$$

Since $f \geq 0$, $\int_K f\,dx = \int_K |f|\,dx$. Thus, $\int_K f\,dx = \|f\|_{1,K} = 0$ for all compact subsets K of Ω. Hence $f = 0$ a.e. on all compact K of Ω and hence $f = 0$ a.e. in Ω. □

The non-negative hypthesis on f in the above result can be relaxed.

Exercise 17. If $f \in L^1_{\text{loc}}(\Omega)$, such that $T_f \equiv 0$, i.e., for all $\phi \in \mathcal{D}(\Omega)$,

$$\int_\Omega f\phi\,dx = 0$$

then $f = 0$ a.e. in Ω.

Proof. Note that it is enough to prove that $f = 0$ a.e. on all compact subsets K of Ω. For any compact subset K of Ω, choose $\phi_K \in \mathcal{D}(\Omega)$ such that $\phi_K \equiv 1$. Define, for each $x \in K$,

$$f_\varepsilon(x) := f\phi_K * \rho_\varepsilon(x) = \int_K f(y)\phi_K(y)\rho_\varepsilon(x - y)\,dy.$$

Since $y \mapsto \phi_K(y)\rho_\varepsilon(x-y)$ is in $\mathcal{D}(\Omega)$, we have $f_\varepsilon \equiv 0$ on K. Since $f_\varepsilon \to f$ in $L^1(K)$,

$$0 = \lim_{\varepsilon \to 0} \|f_\varepsilon - f\|_{1,K} = \lim_{\varepsilon \to 0} \int_K |f_\varepsilon - f|\, dx = \int_K |f|\, dx = \|f\|_{1,K}$$

and hence $f = 0$ a.e. in K. □

A consequence of the above observations is that the association $f \mapsto T_f$ is well-defined because if $f = g$ a.e. in Ω, then the distributions $T_f = T_g$. Thus, we have seen $L^1_{\text{loc}}(\Omega) \subset \mathcal{D}'(\Omega)$. In particular, $\mathcal{D}(\Omega) \subset \mathcal{D}'(\Omega)$. Any distribution that is induced by a locally integrable function is called a *regular* distribution. Otherwise, the distribution is called *singular*.

1.4.2 Measures as Distributions

To each Radon measure $\mu \in \mathcal{R}(\Omega)$, we associate a linear functional on $\mathcal{D}(\Omega)$ as follows:

$$T_\mu(\phi) = \int_\Omega \phi\, d\mu, \quad \forall \phi \in \mathcal{D}(\Omega).$$

The functional T_μ is continuous on $\mathcal{D}(\Omega)$ (and is in $\mathcal{D}'(\Omega)$) because for every compact set K in Ω,

$$|T_\mu(\phi)| \leq \left(\int_K d\mu \right) \|\phi\|_{\infty,K} = |\mu|(K)\|\phi\|_{\infty,\Omega}, \quad \forall \phi \in C^\infty(K),$$

where $|\mu|$ denotes the total variation of the measure. The distribution T_μ is of zero order.

Example 1.9. The Lebesgue measure on Ω restricted to the Borel σ-algebra of Ω is a Radon measure. The distribution induced by the Lebesgue measure is same as that induced by the locally integrable constant function 1 on Ω. Thus, the distribution induced by Lebesgue measure is a *regular* distribution.

Example 1.10 (Dirac Measure). The Dirac measure is a Radon measure that assigns a total mass of 1 at a point $a \in \Omega$. Thus, on the Borel measurable subsets of Ω,

$$\delta_a(E) = \begin{cases} 1 & \text{if } a \in E \\ 0 & \text{if } a \notin E. \end{cases}$$

If $a = 0$ then the corresponding measure $\delta_0 = \delta$ is the Dirac measure discussed in §1.1. The distribution induced by Dirac measure δ_a, called *Dirac distribution*, is:

$$\delta_a(\phi) = \int_\Omega \phi(x)\, d\delta_a.$$

Example 1.11. Let $\gamma \subset \Omega$ be a curve of length one ($|\gamma| = 1$). Given $\alpha \in \mathbb{R}$, we define on the Borel measurable subsets of Ω, $\mu_\gamma(E) = \alpha \times |\gamma \cap E|$. This measure is a generalization of the Dirac measure which signifies that α units of mass are uniformly distributed on a curve of length one.

Exercise 18. Show that, for any[9] function g on Ω,

$$\int_\Omega g(x)\, d\delta_a = g(a).$$

Proof. We shall give two proofs: The first proof is the usual measure theory technique, If $g = 1_E$, the characteristic function of E, then

$$\int_\Omega 1_E\, d\delta_a = \int_E d\delta_a = \delta_a(E) = 1_E(a).$$

For any simple function $g = \sum_{i=1}^k \alpha_i 1_{E_i}$, then

$$\int_\Omega g(x)\, d\delta_a = \sum_{i=1}^k \alpha_i \delta_a(E_i) = g(a).$$

For any non-negative g, by choosing an increasing sequence of simple functions g_i converging to g, one can show that

$$\int_\Omega g(x)\, d\delta_a = g(a).$$

Further any $g = g^+ - g^-$, which ends the first proof.

The second proof is quite elegant. Consider the constant function h on Ω defined as $h(x) = g(a)$. Let

$$E := \{x \in \Omega / g(x) \neq h(x)\}.$$

[9] Yes! we mean "any" function, because the class of Dirac measurable sets is the power set and hence any function is Dirac measurable. Though, in the above context we have restricted the Dirac measure to the class of Borel σ-algebra

Hence, $a \notin E$ and therefore $\delta_a(E) = 0$. This implies that $g = h$ almost everywhere w.r.t δ_a and, hence, their integrals coincide

$$\int_\Omega g(x)\, d\delta_a = \int_\Omega h(x) d\delta_a.$$

But

$$\int_\Omega h(x) d\delta_a = g(a)\delta_a(\Omega) = g(a).$$

□

With the above observations, we note that the Dirac distribution is just

$$\delta_a(\phi) = \phi(a) \quad \forall \phi \in \mathcal{D}(\Omega). \tag{1.4.1}$$

Recall that the distribution induced by Lebesgue measure is same as the one induced by the constant function $1 \in L^1_{\mathrm{loc}}(\Omega)$. If all measure induced distributions were also induced by some locally integrable functions, then we have not achieved anything different by considering measures as distribution. But it turns out that not all measure induced distribution are induced by locally integrable functions.

Proposition 1.4.8. *The Dirac distribution is a* singular *distribution, i.e., there is no $f \in L^1_{loc}(\Omega)$ such that $T_f = \delta_a$.*

Proof. Note that for any $a \in \Omega$, since Ω is open, there is a $\varepsilon_0 > 0$ such that $B(a, \varepsilon_0) \subset \Omega$. For each $0 < \varepsilon < \varepsilon_0$, we choose $\phi_\varepsilon \in \mathcal{D}(\Omega)$ with support in $B(a; \varepsilon)$, $0 \le \phi_\varepsilon \le 1$ and $\phi_\varepsilon = 1$ in $B(a; \frac{\varepsilon}{2})$. Thus,

$$\delta_a(\phi_\varepsilon) = \phi_\varepsilon(a) = 1 \quad \text{for all } 0 < \varepsilon < \varepsilon_0.$$

Suppose f is a locally integrable function such that $T_f = \delta_a$, then

$$\delta_a(\phi_\varepsilon) = T_f(\phi_\varepsilon) = \int_\Omega f\phi_\varepsilon\, dx = \int_{B(a;\varepsilon)} f\phi_\varepsilon\, dx \le \int_{B(a;\varepsilon)} |f|\, dx.$$

Therefore, $1 \le \int_{B(a;\varepsilon)} |f|\, dx$. Since $f \in L^1_{\mathrm{loc}}(\Omega)$, the quantity on RHS is finite. Hence, as $\varepsilon \to 0$, we get a contradiction $1 \le 0$. □

There are many more measures that induce singular distribution. In fact, one can identify the class of measures which induce regular distribution. They are, precisely, the absolutely continuous[10] measures w.r.t the Lebesgue measure. We shall not dwell on this topic, but a quick summary is as follows: Note that for each $f \in L^1_{\text{loc}}(\Omega)$, one can define a signed measure

$$\mu_f(E) := \int_E f\,dx \tag{1.4.2}$$

for all measurable subsets E of Ω. Conversely, the Radon-Nikodym theorem states that, for any absolutely continuous measure μ w.r.t the Lebesgue measure, there is a $f \in L^1_{\text{loc}}(\Omega)$ such that

$$\mu(E) = \int_E f\,dx$$

for all μ-measurable subsets of Ω. This is a generalisation of the Fundamental theorem of Calculus (FTC).

With results of this section we have the following inclusions

$$\mathcal{D}(\Omega) \subsetneq C_c(\Omega) \subseteq C_0(\Omega) \subsetneq L^1_{\text{loc}}(\Omega) \subsetneq \mathcal{R}(\Omega) \subset \mathcal{D}'(\Omega).$$

Note that so far we have only seen examples of distributions which are of zero order. This is because of the following result:

Theorem 1.4.9. *A distribution $T \in \mathcal{D}'(\Omega)$ is of zero order iff T is a distribution induced by a Radon measure.*

Proof. The implication that a Radon measure induced distribution is of zero order is already shown. Conversely, let T be a distribution of zero order. Then, for every compact set $K \subset \Omega$, there is a $C_K > 0$ such that

$$|T(\phi)| \leq C_K \|\phi\|_0 \quad \forall \phi \in C^\infty(K).$$

The idea is to continuously extend T to $C_c(\Omega)$, in a unique way, and then consider the Radon measure obtained from Theorem 1.3.15. Let $\phi \in C_c(\Omega)$

[10] A measure μ is absolutely continuous w.r.t another measure ν if for every element E of the σ-algebra $\mu(E) = 0$ whenever $\nu(A) = 0$, denoted as $\mu \ll \nu$. The Lebesgue measure is absolutely continuous w.r.t counting measure but the converse is not true.

and define $\phi_\varepsilon := \phi * \rho_\varepsilon$. By Theorem 1.3.28, ϕ_ε converges uniformly to ϕ. The sequence $\{T(\phi_\varepsilon)\}$ is Cauchy in $\mathbb{R}$ because

$$|T(\phi_\varepsilon) - T(\phi_\delta)| = |T(\phi_\varepsilon - \phi_\delta)| \leq C_K \|\phi_\varepsilon - \phi_\delta\|_0.$$

Hence the sequence $\{T(\phi_\varepsilon)\}$ converges, thus we extend T uniquely on $C_c(\Omega)$ as $T(\phi) = \lim_{\varepsilon \to 0} T(\phi_\varepsilon)$. Thus, Theorem **??**, there is a Radon measure associated to the extended T. $\square$

Definition 1.4.10. *A distribution $T \in \mathcal{D}'(\Omega)$ is said to be* positive *if $T(\phi) \geq 0$ for all $\phi \geq 0$ in $\mathcal{D}(\Omega)$.*

Exercise 19. Any positive functional on $\mathcal{D}(\Omega)$ is a distribution of order zero and hence corresponds to a positive Radon measure.

Proof. For any compact set $K \subset \Omega$, consider $\phi \in C^\infty(K)$ and $\psi \in \mathcal{D}(\Omega)$, a non-negative function, such that $\psi = 1$ on K. Set $\lambda = \|\phi\|_0$, then $\lambda - \phi(x) \geq 0$ on K. Thus, $\lambda\psi(x) - \phi(x) \geq 0$ is non-negative on Ω. Therefore,

$$\begin{aligned} 0 &\leq T(\lambda\psi(x) - \phi(x)) \\ T(\phi) &\leq \lambda T(\psi) = T(\psi)\|\phi\|_0. \end{aligned}$$

Hence, T is of order zero. Thus, there is a Radon measure associated to T. $\square$

1.4.3 Multipole Distributions

Recall the formulation of Dirac distribution as given in the equation (1.4.1). This motivates, for a fixed non-negative integer k and $a \in \mathbb{R}$, the linear functional on $\mathcal{D}(\mathbb{R})$

$$\delta_a^{(k)}(\phi) = \phi^{(k)}(a), \quad \forall \phi \in \mathcal{D}(\mathbb{R}),$$

where $\phi^{(k)}$ denotes the k-th derivative of ϕ. This functional is continuous on $\mathcal{D}(\mathbb{R})$ because, for all $\phi \in C^\infty(K)$,

$$|\delta_a^{(k)}(\phi)| = |\phi^{(k)}(a)| \leq \|\phi\|_k.$$

The order of this distribution is, at most, k. The situation $k = 0$ corresponds to the Dirac distribution. The situation $k = 1$ is called the *dipole or doublet distribution*.

Example 1.12. We show in this example that the order of the dipole distribution cannot be zero. Suppose that for all compact set $K \subset \mathbb{R}$ and $\phi \in C^\infty(K)$ we have

$$|\delta_0^{(1)}(\phi)| \leq C\|\phi\|_0.$$

Now, choose $\phi \in C^\infty(K)$ such that $0 \in \text{Int}(K)$ and $\phi'(0) \neq 0$. For each integer $m \geq 1$, set $\phi_m(x) := (1/m)\phi(mx)$. Note that $\phi'_m(x) = \phi'(mx)$ and hence

$$0 \neq |\phi'(0)| = |\phi'_m(0)| = |\delta_0^{(1)}(\phi_m)| \leq C\|\phi_m\|_0 = \frac{C}{m}\|\phi\|_0.$$

This is a contradiction because RHS converges to zero, as m increases, and LHS is strictly positive. □

Example 1.13. The example above can be tweaked to show that any k-th multipole distribution, $\delta_0^{(k)}$ cannot have order less than k and hence is of order k. Suppose that for all compact set $K \subset \mathbb{R}$ and $\phi \in C^\infty(K)$ we have

$$|\delta_0^{(k)}(\phi)| \leq C\|\phi\|_j \quad \text{for some } 0 \leq j < k.$$

Now, choose $\phi \in C^\infty(K)$ such that $0 \in \text{Int}(K)$ and $\phi^{(k)}(0) \neq 0$. For each integer $m \geq 1$, set $\phi_m(x) := m^{-k}\phi(mx)$. Note that $\phi_m^{(j)}(x) = m^{j-k}\phi^{(j)}(mx)$. Thus,

$$\begin{aligned} 0 &\neq |\phi^{(k)}(0)| = |\phi_m^{(k)}(0)| = |\delta_0^{(k)}(\phi_m)| \leq C\|\phi_m\|_j = C\sum_{i=0}^{j}\|\phi_m^{(i)}\|_0 \\ &\neq C\sum_{i=0}^{j} m^{i-k}\|\phi^{(i)}\|_0 \leq Cm^{j-k}\sum_{i=0}^{j}\|\phi^{(i)}\|_0 = Cm^{j-k}\|\phi\|_j. \end{aligned}$$

This is a contradiction because $j < k$ and RHS converges to zero, as m increases whereas LHS is strictly positive. □

One may generalise these notions appropriately to higher dimensions. For any open subset Ω of $\mathbb{R}^n$, we define the distribution

$$\delta_a^\alpha(\phi) = D^\alpha\phi(a), \quad \forall \phi \in \mathcal{D}(\Omega),$$

where α is the n tuple of non-negative integers. Just as Dirac distribution models charge density at a point charge, the dipole distribution models charge density for an electric dipole. In fact, any dipole behaviour, viz., the magnetic dipole layer on a surface. For instance, given a smooth surface Γ and f is a continuous function on Γ representing the density of the magnetic moment, one can define the distribution corresponding to the magnetic dipole layer on Γ as,

$$T(\phi) = \int_{\Gamma} f(x) \frac{d\phi(x)}{dn}\, ds(x)$$

where ds denotes the surface element of Γ and d/dn is the derivation in the direction of normal of Γ.

The dipole distribution is singular. In fact, the situation is much worse. The dipole distribution is not induced by any Radon measure.

Proposition 1.4.11. *The dipole distribution $\delta_a^{(1)} \in \mathcal{D}(\mathbb{R})$ is not induced by any Radon measure, i.e., there is no Radon measure μ such that $T_\mu = \delta_a^{(1)}$.*

Proof. Let $\Omega = \mathbb{R}$ and $a = 0$. Choose $\phi \in \mathcal{D}(\mathbb{R})$ with support in $[-1, 1]$ such that $0 \leq \phi \leq 1$ and $\phi \equiv 1$ on $[-1/2, 1/2]$. Set $\phi_m(x) = \sin(mx)\phi(x)$. Thus, $\operatorname{supp}(\phi_m) \subset [-1, 1]$ and $|\phi_m| \leq 1$. Note that the derivative of ϕ_m is $\phi_m'(x) = \sin(mx)\phi'(x) + m\cos(mx)\phi(x)$ and, hence, $\phi_m'(0) = m$ for all positive integer $m > 0$. Therefore, the dipole distribution takes the value,

$$\delta_0^{(1)}(\phi_m) = \phi_m'(0) = m, \quad \forall m \in \mathbb{N}.$$

Suppose there exists a Radon measure μ inducing the dipole distribution $\delta_0^{(1)}$, then

$$m = |\delta_0^{(1)}(\phi_m)| = |T_\mu(\phi_m)| = \left| \int_{-1}^{1} \phi(x)\sin(mx)\, d\mu \right| \leq |\mu|(B(0;1)).$$

The inequality $|\mu|(B(0;1)) \geq m$, for all $m \in \mathbb{N}$, implies that $|\mu|(B(0;1))$ is infinite which contradicts the fact that the Radon measure μ is finite on compact subsets of $\mathbb{R}$. $\square$

We shall see later that the multipole distribution δ_a^α is, in fact, the α-th distributional derivative of the Dirac measure δ_a up to a sign change. This motivation was behind the choice of notation of multipole distribution. In this section, we have noted that the space of Radon measures is properly contained in the space of distributions, i.e., $\mathcal{R}(\mathbb{R}) \subsetneq \mathcal{D}'(\mathbb{R})$.

Exercise 20. Do we have singular distributions which are neither induced by Radon measures nor distributional derivatives of Radon measures? (Hyperfunctions?)

1.4.4 Infinite Order Distributions

We have introduced distributions of finite order thus far. Does there exist distributions of infinite order? In this section, we given an example of an infinite order distribution.

Example 1.14. Define the functional T on $\mathcal{D}(\mathbb{R})$ as $T(\phi) := \sum_{k=0}^{\infty} \phi^{(k)}(k)$. Without loss of generality, consider the compact set $[0, m]$. If $\phi \in C^\infty[0, m]$, then

$$|T(\phi)| \leq \sum_{k=0}^{m} |\phi^{(k)}(k)| = \sum_{k=0}^{m} \|\phi^{(k)}\|_0 = \|\phi\|_m.$$

The larger the compact set becomes, the higher the derivatives of ϕ needs to be taken in. Thus, there is no fixed m for all compact subsets of $\mathbb{R}$.

Example 1.15. Define the functional $T \in \mathcal{D}'(0, \infty)$ as

$$T(\phi) = \sum_{k=1}^{\infty} \phi^{(k)}(1/k).$$

Let $\phi \in \mathcal{D}(0, \infty)$ be such that supp$(\phi) \subset [1/m, m]$ (exhaustion sets), then

$$T(\phi) = \sum_{k=1}^{m} \phi^{(k)}(1/k) \leq \sum_{k=1}^{m} \|\phi^{(k)}\|_0 \leq \sum_{k=0}^{m} \|\phi^{(k)}\|_0 = \|\phi\|_m.$$

Thus, $T \in \mathcal{D}'(0, \infty)$ is a distribution of infinite order.

1.4.5 Topology on Distributions

In this section, we give a suitable topology on the space of distributions. Recall that the space of distributions $\mathcal{D}'(\Omega)$ is the (topological) dual of $\mathcal{D}(\Omega)$. For every $\phi \in \mathcal{D}(\Omega)$, one can define the linear functional $\Lambda_\phi : \mathcal{D}'(\Omega) \to \mathbb{R}$ as follows, $\Lambda_\phi(T) = T(\phi)$. The linear functionals Λ_ϕ are included in the second (algebraic) dual of $\mathcal{D}(\Omega)$. We consider the coarsest[11] topology on $\mathcal{D}'(\Omega)$ such that all the linear maps $\Lambda_\phi : \mathcal{D}'(\Omega) \to \mathbb{R}$, corresponding to each $\phi \in \mathcal{D}(\Omega)$, are continuous. This is called the *weak-** topology on $\mathcal{D}'(\Omega)$. For every open subset of $V \subset \mathbb{R}$, consider the collection of subsets $\{\Lambda_\phi^{-1}(V)\}$ of $\mathcal{D}'(\Omega)$, for all $\phi \in \mathcal{D}(\Omega)$. The weak-* topology is the topology generated by this collection of subsets in $\mathcal{D}'(\Omega)$. The space $\mathcal{D}'(\Omega)$ is sequentially weak-* complete but not weak-* complete.

[11] weakest or smallest topology, one with fewer open sets

Definition 1.4.12. *We say a sequence of distributions $\{T_m\}$ is* Cauchy *in $\mathcal{D}'(\Omega)$, if $T_m(\phi)$ is Cauchy in $\mathbb{R}$, for all $\phi \in \mathcal{D}(\Omega)$. A sequence $\{T_m\} \subset \mathcal{D}'(\Omega)$ converges to T in weak-* topology of $\mathcal{D}'(\Omega)$ or in the distribution sense if $T_m(\phi) \to T(\phi)$ for all $\phi \in \mathcal{D}(\Omega)$, denoted as $T_m \rightharpoonup T$.*

Exercise 21. If S_m and T_m are sequences of distributions converging to S and T, respectively, then show that $\lambda S_m + \mu T_m$ converges to $\lambda S + \mu T$, for all $\lambda, \mu \in \mathbb{R}$.

Proposition 1.4.13 (Sequential Completeness)**.** *Let $\{T_m\}$ be a Cauchy sequence (of distributions) in $\mathcal{D}'(\Omega)$ and let $T(\phi) := \lim_{m\to\infty} T_m(\phi)$ (T is well-defined because $\mathbb{R}$ is complete). Then $T \in \mathcal{D}'(\Omega)$.*

Proof. To show $T \in \mathcal{D}'(\Omega)$, it is enough to show that $T : C^\infty(K) \to \mathbb{R}$ is continuous for all compact sets K in Ω. Fix a compact set K in Ω and $\phi \in \Omega$. Since $\{T_m(\phi)\}$ is convergent in $\mathbb{R}$ it is bounded and hence there is a real constant $C_\phi > 0$ (may depend on ϕ) such that $\sup_m |T_m(\phi)| \leq C_\phi$. The family $\{T_m\}$ is point-wise bounded on $C^\infty(K)$. Then, by Banach-Steinhaus theorem (uniform boundeness principle) of $C^\infty(K)$, T is uniformly bounded, i.e., there is a constant $C > 0$ (independent of ϕ) such that $\sup_m |T_m(\phi)| \leq C$ for all $\phi \in C^\infty(K)$. Therefore, for all $\phi \in C^\infty(K)$,

$$|T(\phi)| \leq \sup_m |T_m(\phi)| \leq C \quad \forall \phi \in C^\infty(K).$$

and hence T is continuous on $C^\infty(K)$ for all compact subsets K of Ω. Thus, T is continuous on $\mathcal{D}(\Omega)$ and hence $T \in \mathcal{D}'(\Omega)$. □

A natural question is the relation between distributional convergence and poin-wise, uniform or L^p convergence of regular distributions. The weak-* convergence in $\mathcal{D}'(\Omega)$ is weaker than any other topology that preserves the continuity of the functionals Λ_ϕ. For instance, this topology is weaker that L^p-topology or uniform norm topology etc., i.e., convergence in p-norm or uniform norm implies convergence in distribution sense of classical functions.

Example 1.16. Consider the sequence $f_m(x) = e^{imx}$ on $\mathbb{R}$ which converges point-wise to 0 iff $x \in 2\pi\mathbb{Z}$. The sequence of distributions corresponding to f_m converges to zero in the distributional sense because

$$\int e^{imx}\phi(x)\,dx = -\frac{1}{im}\int e^{imx}\phi'(x)\,dx.$$

We have used integration by parts to get the second integral and since the integrand is bounded in the second integral, it converges to zero as $m \to \infty$.

Example 1.17. Let $\rho \in C_c(\mathbb{R}^n)$ such that supp$(\rho)\subset \overline{B(0;1)}$ and $\int_{\mathbb{R}^n} \rho(y)\,dy = 1$. Now set $\rho_\varepsilon(x) := \varepsilon^{-n}\rho(\frac{x}{\varepsilon})$. This sequence is called the *Dirac Sequence*. In particular, the sequence of mollifiers (cf. (1.3.3)) is one example. Another trivial example is

$$\rho(x) = \begin{cases} 2 & |x| < 1 \\ 0 & |x| \geq 1. \end{cases}$$

We claim that the sequence ρ_ε converges to the Dirac distribution δ_0, in the distribution sense. Let T_ε denote the distribution corresponding to ρ_ε. For any $\phi \in \mathcal{D}(\mathbb{R}^n)$, consider

$$T_\varepsilon(\phi) = \int_{\mathbb{R}^n} \rho_\varepsilon(x)\phi(x)\,dx = \int_{\mathbb{R}^n} \varepsilon^{-n}\rho\left(\frac{x}{\varepsilon}\right)\phi(x)\,dx = \int_{\mathbb{R}^n} \rho(y)\phi(\varepsilon y)\,dy.$$

Taking limit $\varepsilon \to 0$ both sides we get

$$\lim_{\varepsilon\to 0} T_\varepsilon(\phi) = \int_{\mathbb{R}^n} \rho(y)\phi(0)\,dy = \phi(0) = \delta_0(\phi).$$

The interchange of limit is possible due to uniform continuity of ϕ which induces the uniform convergence of $\phi(\varepsilon y) \to \phi(0)$. Thus, $T_\varepsilon \to \delta_0$ in the distribution sense, whereas the point-wise limit at $x = 0$ did not exist for ρ_ε.

Example 1.18. Consider the sequence of functions

$$f_m x = \begin{cases} m^2 x & 0 \leq x < \frac{1}{m} \\ m^2\left(\frac{2}{n} - x\right) & \frac{1}{m} < x \leq \frac{2}{m} \\ 0 & \text{otherwise.} \end{cases}$$

Note that they converge point-wise to zero for all $x \in \mathbb{R}$. Let T_m denote the distribution corresponding to f_m and hence

$$T_m(\phi) = m^2 \int_0^{1/m} x\phi(x)\,dx + m^2 \int_{1/m}^{2/m} \left(\frac{2}{n} - x\right)\phi(x)\,dx.$$

Both the integral above converges to $\phi(0)/2$ and hence the sequence of distributions converges to δ_0. Let us give the proof for the first integral. Since

$$m^2 \int_0^{1/m} x\,dx = \frac{1}{2}$$

we consider

$$\begin{aligned}\left| m^2 \int_0^{1/m} x\phi(x)\,dx - \frac{1}{2}\phi(0) \right| &= \left| m^2 \int_0^{1/m} x[\phi(x) - \phi(0)]\,dx \right| \\ &\leq m^2 \int_0^{1/m} |x||\phi(x) - \phi(0)|\,dx.\end{aligned}$$

Since ϕ is continuous at $x = 0$, for every $\varepsilon > 0$, there is a $\delta > 0$ such that $|\phi(x) - \phi(0)| < \varepsilon$ whenever $|x| < \delta$. Thus, for all m such that $1/m < \delta$, we have

$$\left| m^2 \int_0^{1/m} x\phi(x)\,dx - \frac{1}{2}\phi(0) \right| \leq m^2 \frac{\varepsilon}{m} \int_0^{1/m} dx < \varepsilon.$$

Hence,

$$\lim_{m\to\infty} m^2 \int_0^{1/m} x\phi(x)\,dx = \frac{1}{2}\phi(0).$$

Similarly, one can show that the second integral is also $(1/2)\phi(0)$ and hence $\int f_m(x)\phi(x)\,dx = \phi(0)$.

The following theorem gives the condition that is violated by the above example for the point-wise limit to coincide with the distributional limit.

Theorem 1.4.14 (Weak Dominated Convergence Theorem)**.** *Let $\{f_m\} \subset L^1_{loc}(\Omega)$ such that $f_m(x) \to f(x)$ point-wise for a.e. $x \in \Omega$ and there is a $g \in L^1(\Omega)$ such that for every compact set $K \subset \Omega$ $|f_m| \leq g$ for all m. Then $f \in L^1_{loc}(\Omega)$ and $f_m \rightharpoonup f$ in $\mathcal{D}'(\Omega)$.*

Proof. Let T_m denote the distribution corresponding to f_m and fix $\phi \in \mathcal{D}(\Omega)$. Set $h_m(x) = f_m(x)\phi(x)$ and $h(x) = f(x)\phi(x)$. Note that $h_m(x) \to h(x)$ point-wise a.e. in Ω (in fact in the supp(ϕ)). Also, $|h_m| \leq g|\phi|$ and $g|\phi| \in L^1(\Omega)$. By classical Lebesgue's dominated convergence theorem, $h \in L^1(\Omega)$ and

$$\int_\Omega h(x)\,dx = \lim_{m\to\infty} \int_\Omega h_m(x)\,dx = \lim_{m\to\infty} \int_\Omega f_m(x)\phi(x)\,dx.$$

In particular, choosing $\phi \equiv 1$ on a given compact set K of Ω and since for this ϕ, $h \in L^1(\Omega)$, we have $f \in L^1_{\text{loc}}(\Omega)$ because

$$\infty > \int_\Omega |h|\,dx = \int_\Omega |f\phi|\,dx = \int_K |f|\,dx.$$

Let T_f be the distribution corresponding to f. Moreover,

$$\lim_{m\to\infty} T_m(\phi) = \lim_{m\to\infty} \int_\Omega f_m \phi \, dx = \int_\Omega f\phi \, dx = T_f(\phi).$$

$\square$

Corollary 1.4.15. *Let $\{f_m\} \subset C(\Omega)$ be a sequence of continuous functions that converges uniformly on compact subsets of Ω to f. Then $f_m \rightharpoonup f$ in $\mathcal{D}'(\Omega)$.*

By the vector space structure of $\mathcal{D}'(\Omega)$, we already have the finite sum of distributions. If $\{T_i\}_1^k \subset \mathcal{D}'(\Omega)$ then $T := \sum_{i=1}^k T_i \in \mathcal{D}'(\Omega)$ defined as $T(\phi) = \sum_{i=1}^k T_i(\phi)$. The topology on $\mathcal{D}'(\Omega)$ can be used to give the notion of series of distributions.

Definition 1.4.16. *For any countable collection of distributions $\{T_i\}_1^\infty \subset \mathcal{D}'(\Omega)$ the* series *$\sum_{i=1}^\infty T_i$ is said to* converge *to $S \in \mathcal{D}'(\Omega)$ if the sequence of partial sums $T_m := \sum_{i=1}^m T_i$ converges to S in $\mathcal{D}'(\Omega)$.*

1.4.6 Principal Value Distribution

In practice, we encounter functions whose integral diverges. This motivates other ways of studying divergent integrals, like principal part, finite part of divergent integral.

Definition 1.4.17 (Singularity at infinity)**.** *We say that the integral of a function f (with singularity at ∞) exists in the* generalised sense *if the limit*

$$\lim_{a,b\to+\infty} \int_{-a}^{b} f(x)\, dx$$

exists. We say that the principal value *of the integral exists if the limit (with $b = a$)*

$$\lim_{a\to+\infty} \int_{-a}^{b} f(x)\, dx$$

exists.

Definition 1.4.18 (Singularity at point)**.** *We say that the integral of a compact supported function f (with singularity at a point 0) exists in the* generalised sense *if the limit*

$$\lim_{a,b\to 0^+} \int_{-\infty}^{-a} f(x)\,dx + \int_{b}^{\infty} f(x)\,dx$$

exists. We say that the principal value *of the integral exists if the limit (with $b = a$)*

$$\lim_{a\to 0^+} \int_{|x|>a} f(x)\,dx$$

exists.

We have already seen that every locally integrable function can be identified with a distribution. We have already noted, in §1.4.1, that $1/x \notin L^1_{\text{loc}}(\mathbb{R})$. Consider the function $f_\varepsilon : \mathbb{R} \to \mathbb{C}$, for each $\varepsilon > 0$, defined as

$$f_\varepsilon(x) = \frac{1}{x + i\varepsilon}.$$

What is the limit of f_ε, as $\varepsilon \to 0$? Of course, when $x \neq 0$, $f_\varepsilon(x) \to 1/x$. What is the limit when $x = 0$? Classically we cannot make sense of this limit. Similarly, note that the integral $\int_{-1}^{1}(1/x)\,dx$ does not converge. However,

$$\lim_{\varepsilon\to 0} \int_{-1}^{1} f_\varepsilon(x)\,dx = \lim_{\varepsilon\to 0} \ln(x + i\varepsilon)\,|_{-1}^{1} = \ln(1) - \ln(-1) = 0 - (i\pi) = -i\pi.$$

Thus, we expect $\int_{-1}^{1}(1/x)\,dx = -i\pi$. This cannot be made sense classically[12]. Using the theory of distributions one can give a meaning to this observation.

We do know that $1/x \in L^1_{\text{loc}}(\mathbb{R} \setminus \{0\})$ and hence induces a distribution in $\mathcal{D}(\mathbb{R} \setminus \{0\})$. One can extend this distribution to yield a distribution corresponding $1/x$ on $\mathbb{R}$. We define the linear functional $\text{PV}\left(\frac{1}{x}\right)$ on $\mathcal{D}(\mathbb{R})$ as

$$\text{PV}\left(\frac{1}{x}\right)(\phi) = \lim_{\varepsilon\to 0} \int_{|x|\geq\varepsilon} \frac{1}{x}\phi(x)\,dx. \tag{1.4.3}$$

As defined before, the limit on RHS is called the *Cauchy's Principal Value* and hence the use of 'PV' in the notation. Note that $PV(1/x)$ is defined as

[12] recall (1.1.1)

a distributional limit of the sequence of distributions corresponding to the locally integrable functions

$$f_\varepsilon(x) := \begin{cases} \frac{1}{x} & |x| > \varepsilon, \\ 0 & |x| \leq \varepsilon, \end{cases}$$

i.e., $f_\varepsilon \rightharpoonup PV\left(\frac{1}{x}\right)$ in $\mathcal{D}'(\Omega)$.

Lemma 1.4.19. *The principal value of* $1/x$ *(limit on the RHS) exists.*

Proof. Let $\phi \in \mathcal{D}(\mathbb{R})$. Without loss of generality, let supp$(\phi) \subset [-a, a]$ for some real $a > 0$. For $\varepsilon > 0$ small enough, consider

$$\begin{aligned} \int_{|x| \geq \varepsilon} \frac{\phi(x)}{x}\, dx &= \int_{-a}^{-\varepsilon} \frac{\phi(x)}{x}\, dx + \int_{\varepsilon}^{a} \frac{\phi(x)}{x}\, dx \\ &= \int_{\varepsilon}^{a} \frac{\phi(x)}{x}\, dx - \int_{\varepsilon}^{a} \frac{\phi(-x)}{x}\, dx = \int_{\varepsilon}^{a} \frac{\phi(x) - \phi(-x)}{x}\, dx. \end{aligned}$$

Therefore, taking limit both sides as $\varepsilon \to 0$,

$$\begin{aligned} \mathrm{PV}\left(\frac{1}{x}\right)(\phi) &= \int_0^a \frac{\phi(x) - \phi(-x)}{x}\, dx \\ &= \int_0^a \frac{1}{x}\left(\int_{-x}^{x} \phi'(s)\, ds\right) dx \quad \text{(Using FTC)} \\ &= \int_0^a \frac{1}{x}\left(\int_{-1}^{1} x\phi'(xt)\, dt\right) dx. \end{aligned}$$

Set $\psi(x) = \int_{-1}^{1} \phi'(xt)\, dt$, then

$$\left|\mathrm{PV}\left(\frac{1}{x}\right)(\phi)\right| = \left|\int_0^a \frac{1}{x} x\psi(x)\, dx\right| = \left|\int_0^a \psi(x)\, dx\right| \leq \int_0^a |\psi(x)|\, dx.$$

But $|\psi(x)| \leq 2\|\phi'\|_0 \leq 2\|\phi\|_1$. Hence,

$$\left|\mathrm{PV}\left(\frac{1}{x}\right)(\phi)\right| \leq 2a\|\phi\|_1.$$

Since RHS is finite, the limit is finite. $\square$

We have shown that the principal value functional is a distribution, called the *Principal Value (PV) Distribution* corresponding to $1/x$. The $1/x$ in (1.1.1) derived by Dirac will make sense as a principal value distribution. The proof also highlights that, at most, the PV distribution can have order one. We show in the example below that the PV distribution cannot be of order zero.

Example 1.19. The PV distribution cannot be of zero order and hence its order is one. Observe the following characterisation of principal value distribution in the above argument when supp(ϕ)$\subset [-a,a]$,

$$\mathrm{PV}\left(\frac{1}{x}\right)(\phi) = \int_0^a \frac{\phi(x) - \phi(-x)}{x}\,dx.$$

Choose ϕ such that supp(ϕ)$\subset [0,a]$, $0 \le \phi \le 1$ and $\phi \equiv 1$ on $[c,d] \subset [0,a]$ with $c \neq 0$ and $d \neq a$. Then,

$$\begin{aligned}
PV(1/x)(\phi) &= \int_0^c \frac{\phi(x)}{x}\,dx + \int_c^d \frac{\phi(x)}{x}\,dx + \int_d^a \frac{\phi(x)}{x}\,dx \\
&> \int_c^d \frac{1}{x}\,dx \quad \text{first and third integral being non-negative} \\
&> \frac{1}{a}\int_c^d dx \quad \text{since } x < a \\
&= \frac{1}{a}(d-c) \ge \frac{d-c}{a} \sup_{x\in[0,a]} |\phi| = \frac{d-c}{a}\|\phi\|_0.
\end{aligned}$$

Thus, the distribution cannot be of zero order. □

The principal value distribution is not the only choice for the function $1/x$. Consider the linear functional on $\mathcal{D}(\mathbb{R})$ corresponding to $1/x$,

$$T(\phi) = \lim_{\varepsilon\to 0} \int_{-\infty}^{-\varepsilon} \frac{\phi(x)}{x}\,dx + \int_{2\varepsilon}^{\infty} \frac{\phi(x)}{x}\,dx.$$

Exercise 22. Show that the T defined above is a distribution. Also, compare T and PV$(1/x)$.

Proof. Let $\phi \in \mathcal{D}(\mathbb{R})$. Without loss of generality, let supp(ϕ)$\subset [-a,a]$ for some real $a > 0$. For $\varepsilon > 0$ small enough, consider

$$\begin{aligned}\int_{-\infty}^{-\varepsilon} \frac{\phi(x)}{x}\,dx + \int_{2\varepsilon}^{\infty} \frac{\phi(x)}{x}\,dx &= \int_{-a}^{-2\varepsilon} \frac{\phi(x)}{x}\,dx + \int_{-2\varepsilon}^{-\varepsilon} \frac{\phi(x)}{x}\,dx \\ &\quad + \int_{2\varepsilon}^{a} \frac{\phi(x)}{x}\,dx \\ &= \int_{2\varepsilon}^{a} \frac{\phi(x) - \phi(-x)}{x}\,dx + \int_{-2\varepsilon}^{-\varepsilon} \frac{\phi(x)}{x}\,dx \\ &= \int_{2\varepsilon}^{a} \psi(x)\,dx - \int_{1}^{2} \frac{\phi(-\varepsilon x)}{x}\,dx.\end{aligned}$$

Therefore,

$$T(\phi) = \int_0^a \psi(x)\,dx - \int_1^2 \frac{\phi(0)}{x}\,dx = \int_0^a \psi(x)\,dx - \phi(0)\ln 2.$$

Thus, $|T(\phi)| \leq 2a\|\phi'\|_0 + \|\phi\|_0 \ln 2 \leq C\|\phi\|_1$ where $C = \max(2a, \ln 2)$. T is again a distribution of order one. Now, consider

$$(\mathrm{PV}(1/x) - T)(\phi) = \lim_{\varepsilon\to 0} \int_{\varepsilon}^{2\varepsilon} \frac{\phi(x)}{x}\,dx = \lim_{\varepsilon\to 0} \int_1^2 \frac{\phi(\varepsilon x)}{x}\,dx = \phi(0)\ln 2.$$

□

Remark 1.4.20. Yet another way to defining $PV(1/x)$ is as follows, this time from the complex plane. Consider the distributions T_ε corresponding to

$$f_\varepsilon(x) := \frac{1}{x + i\varepsilon}.$$

We shall show that $T_\varepsilon \rightharpoonup PV(1/x) - i\pi\delta_0$. Similarly, if S_ε are distributions corresponding to

$$f_\varepsilon(x) := \frac{1}{x - i\varepsilon}.$$

We shall show that $S_\varepsilon \rightharpoonup PV(1/x) + i\pi\delta_0$.

1.4.7 Functions, but not Distributions

It is not always possible to extend the notion of distributions to non-locally integrable functions as done above.

Example 1.20. The function $e^{1/x} \notin L^1_{\text{loc}}(\mathbb{R})$ but is in $L^1_{\text{loc}}(\mathbb{R} \setminus \{0\})$. Let $T \in \mathcal{D}'(0,\infty)$ be the distribution corresponding to $e^{1/x}$ on $(0,\infty)$. We shall show that there is no distribution $S \in \mathcal{D}'(\mathbb{R})$ corresponding to $e^{1/x}$ on $\mathbb{R}$ whose restriction coincides with the distribution T on $(0,\infty)$.

Example 1.21. Consider the distribution $T \in \mathcal{D}'(0,\infty)$ defined in Example 1.15, i.e.,

$$T(\phi) = \sum_{k=1}^{\infty} \phi^{(k)}(1/k).$$

We shall show that there is no distribution $S \in \mathcal{D}'(\mathbb{R})$ whose restriction to $(0,\infty)$ coincides with T. Suppose there exists a $S \in \mathcal{D}'(\mathbb{R})$ such that $T = S$ on $(0,\infty)$ then for the compact set $[-1,1]$ there exists a $C > 0$ and $N \in \mathbb{N} \cup \{0\}$ such that

$$|S(\phi)| \leq C\|\phi\|_N \quad \forall \phi \in C^\infty[-1,1].$$

Now, choose $m > N$ and ψ such that supp$(\psi) \subset (1/(m+1), 1/(m-1))$. Then $T(\psi) = \psi^{(m)}(1/m)$. Since T and S coincide for $\psi \in \mathcal{D}(0,\infty)$, we have

$$|\delta^{(m)}_{1/m}(\psi)| = |\psi^{(m)}(1/m)| = |T(\psi)| = |S(\psi)| \leq C\|\psi\|_N.$$

Thus, we have that the order of the multipole distribution $\delta^{(m)}_{1/m}$ is at most N, a quantity smaller than m. This is a contradiction. Hence, we can have no extension S of T to $\mathbb{R}$.

1.5 Operations With Distributions

The space of distributions $\mathcal{D}'(\Omega)$ is a vector space over $\mathbb{R}$ (or $\mathbb{C}$). Thus, we already have the operation of addition of distribution and multiplication by reals induced from the property of linear functionals. Hence, for any two $S, T \in \mathcal{D}'(\Omega)$, $(S+T)(\phi) = S(\phi) + T(\phi)$. Also, for any $\lambda \in \mathbb{R}$ and $T \in \mathcal{D}'(\Omega)$, the distribution λT is defined as $(\lambda T)(\phi) = \lambda T(\phi)$. We begin by introducing the concept of derivative of a distribution.

1.5.1 Differentiation

Recall the discussion leading to (1.3.1). For any $f \in C^\infty(\mathbb{R})$, $\phi \in \mathcal{D}(\mathbb{R})$ and all $k \in \mathbb{N}$, using integration by parts we have

$$\int f^{(k)}\phi \, dx = (-1)^k \int f\phi^{(k)} \, dx.$$

This motivates the following definition of derivative of a distribution.

Definition 1.5.1. *For any $T \in \mathcal{D}'(\Omega)$ and n-tuple α, the derivative $D^\alpha T$ of T is defined as*

$$(D^\alpha T)(\phi) = (-1)^{|\alpha|} T(D^\alpha \phi), \quad \phi \in \mathcal{D}(\Omega).$$

The derivative $D^\alpha T$ is also a distribution because for all compact subsets K of Ω there are $C_K > 0$ and non-negative integer $N_K = N$ such that $|T(\phi)| \leq C_K \|\phi\|_N$ for all $\phi \in C^\infty(K)$. Thus, for any compact subset K of Ω and $\phi \in C^\infty(K)$,

$$|(D^\alpha T)(\phi)| = |(-1)^{|\alpha|} T(D^\alpha \phi)| = |T(D^\alpha \phi)| \leq C_K \|D^\alpha \phi\|_N \leq C_K \|\phi\|_{N+|\alpha|}.$$

Thus, $D^\alpha T \in \mathcal{D}'(\Omega)$. For a finite order distribution T, the derivative $D^\alpha T$ has order more than T, i.e., differentiation operation increases the order of the distribution. Observe that every distribution is infinitely differentiable and the mixed derivatives are equal because the same holds for test functions. The following result highlights some of these properties of the differentiation operation:

Exercise 23. Let $S, T \in \mathcal{D}'(\Omega)$. Show that

(i) $D^\beta(D^\alpha T) = D^\alpha(D^\beta T) = D^{\alpha+\beta} T$ for all multi-indices α, β.

(ii) $D^\alpha(\lambda S + \mu T) = \lambda D^\alpha S + \mu D^\alpha T$, for each $\lambda, \mu \in \mathbb{R}$.

Proof. (i) We see that

$$\begin{aligned} D^\beta(D^\alpha T)(\phi) &= (-1)^{|\beta|} D^\alpha T(D^\beta \phi) = (-1)^{|\beta|+|\alpha|} T(D^\alpha D^\beta \phi) \\ &= (-1)^{|\beta|+|\alpha|} T(D^{\alpha+\beta} \phi) = (D^{\alpha+\beta} T)(\phi). \end{aligned}$$

Similarly for $D^\alpha(D^\beta T)$.

(ii) Consider

$$\begin{aligned} [D^\alpha(\lambda S + \mu T)](\phi) &= (-1)^{|\alpha|} (\lambda S + \mu T)(D^\alpha \phi) \\ &= (-1)^{|\alpha|} \left[\lambda S(D^\alpha \phi) + \mu T(D^\alpha \phi)\right] \\ &= (\lambda D^\alpha S + \mu D^\alpha T)(\phi). \end{aligned}$$

□

Exercise 24. Let $n = 1$, $\Omega = (0, 2)$ and

$$u(x) = \begin{cases} x & \text{if } 0 < x \leq 1 \\ 1 & \text{if } 1 \leq x < 2. \end{cases}$$

Compute the first distributional derivative of u.

Proof. For any $\phi \in C_c^\infty(\Omega)$, consider

$$\begin{aligned} Du(\phi) &= -\int_0^2 u(x)\phi'(x)\,dx \\ &= -\int_0^1 x\phi'\,dx - \int_1^2 \phi'\,dx \\ &= \int_0^1 \phi\,dx - \phi(1) + \phi(1) = \int_0^2 v(x)\phi(x)\,dx, \end{aligned}$$

where

$$v(x) = \begin{cases} 1 & \text{if } 0 < x \leq 1 \\ 0 & \text{if } 1 < x < 2. \end{cases}$$

Thus, $Du = v$. $\square$

Exercise 25. question Let $n = 1$, $\Omega = (0, 2)$ and

$$u(x) = \begin{cases} x & \text{if } 0 < x \leq 1 \\ 2 & \text{if } 1 < x < 2. \end{cases}$$

Compute the first distributional derivative of u.

Proof. Consider

$$\begin{aligned} Du(\phi) &= -\int_0^2 u\phi'\,dx = -\int_0^1 x\phi'\,dx - 2\int_1^2 \phi'\,dx \\ &= \int_0^1 \phi\,dx + \phi(1). \end{aligned}$$

Then, $Du = v + \delta_1$ in $\mathcal{D}'(\Omega)$ where

$$v = \begin{cases} 1 & \text{if } 0 < x \leq 1 \\ 0 & \text{if } 1 < x < 2. \end{cases}$$

$\square$

Definition 1.5.2. *A function $f \in L^1_{loc}(\Omega)$ is said to be* α-weakly differentiable *if $D^\alpha T_f$ is a regular distribution. In other words, for any given multi-index α, a function $f \in L^1_{loc}(\Omega)$ is said to be* α-weakly differentiable *if there exists a $g_\alpha \in L^1_{loc}(\Omega)$ such that*

$$\int_\Omega g_\alpha \phi \, dx = (-1)^{|\alpha|} \int_\Omega f D^\alpha \phi \, dx \quad \forall \phi \in \mathcal{D}(\Omega).$$

Example 1.22. Consider the continuous function $f : \mathbb{R} \to \mathbb{R}$ defined as $f(x) = |x|$. Classically, its derivative exists a.e. and is the function

$$f'(x) = \begin{cases} -1 & -\infty < x < 0 \\ 1 & 0 < x < \infty. \end{cases}$$

The distribution corresponding to f' is

$$T_{f'}(\phi) = -\int_{-\infty}^0 \phi(x) \, dx + \int_0^\infty \phi(x) \, dx$$

However, for $\phi \in \mathcal{D}(\mathbb{R})$, the distributional derivative of T_f is

$$\begin{aligned} DT_f(\phi) &= -T_f(\phi') = -\int_{-\infty}^0 (-x)\phi'(x) \, dx - \int_0^\infty x\phi'(x) \, dx \\ &= x\phi(x) \,|_{-\infty}^0 - \int_{-\infty}^0 \phi(x) \, dx + \int_0^\infty \phi(x) \, dx - x\phi(x) \,|_0^\infty \\ &= \int_0^\infty \phi(x) \, dx - \int_{-\infty}^0 \phi(x) \, dx \\ &= T_{f'}(\phi). \end{aligned}$$

Therefore, $DT_f = T_{f'}$. Moreover, $|x|$ is weakly differentiable.

Example 1.23. Consider the everywhere discontinuous function

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q}^c \\ 0 & x \in \mathbb{Q}. \end{cases}$$

The function f is not differentiable at any point. The Lebesgue measure of $\mathbb{Q}$ is zero, hence,

$$DT_f(\phi) = -T_f(\phi') = -\int_{\mathbb{R}} \phi'(x) \, dx = 0.$$

Thus, $DT_f = 0$ and f is weakly differentiable because $DT_f = T_0$.

Example 1.24. Consider the everywhere discontinuous function

$$f(x) = \begin{cases} \sin x & x \in \mathbb{Q}^c \\ 0 & x \in \mathbb{Q}. \end{cases}$$

The function f is not differentiable at any point. The Lebesgue measure of $\mathbb{Q}$ is zero, hence,

$$DT_f(\phi) = -T_f(\phi') = -\int_{\mathbb{R}} \sin x \phi'(x)\, dx = \int_{\mathbb{R}} \cos x \phi(x)\, dx.$$

Thus, $DT_f = T_{\cos x}$ and hence f is weakly differentiable.

Example 1.25. We shall now give an example of a function for which the classical and distributional derivative do not coincide. Consider the locally integrable function H_a, for every $a \in \mathbb{R}$,

$$H_a(x) = \begin{cases} 1 & \text{if } x > a \\ 0 & \text{if } x \leq a. \end{cases}$$

When $a = 0$, the function H_0 is the Heaviside function H as defined in (1.1.2). Classically, the function H_a is differentiable a.e. and $H_a'(x) = 0$ a.e. (except at a). Therefore, $T_{H_a'} = 0$. We shall now compute the distributional derivative of the regular distribution T_{H_a}, induced by H_a. For $\phi \in \mathbb{R}$, the derivative of T_{H_a} is computed as,

$$DT_{H_a}(\phi) = -T_{H_a}(\phi') = -\int_a^\infty \phi'(x)\, dx = \phi(a) = \delta_a(\phi).$$

Thus, the distributional derivative of the function H_a is the Dirac distribution, a singular distribtuion. Therefore, $DT_{H_a} = \delta_a(\phi) \neq 0 = T_{H_a'}$. In fact, H_a is not weakly differentiable.

Observe above that the function H_a had a jump discontinuity and Dirac measure appeared as its derivative at the point of "jump". This is the feature of Dirac measure. It appears as a derivative at points of jump. However, note that in Example 1.23 and 1.24, the jump at every point did not give rise to a Dirac measure because outside the set of jump points (in that case $\mathbb{Q}$) the function was continuous.

Example 1.26. Consider the locally integrable function $f : \mathbb{R} \to \mathbb{R}$ defined as

$$f(x) = \begin{cases} |x| & |x| < a \\ 0 & |x| > a. \end{cases}$$

Classically, its derivative exists a.e. and is the function

$$f'(x) = \begin{cases} -1 & -a < x < 0 \\ 1 & 0 < x < a \\ 0 & |x| > a. \end{cases}$$

The distribution corresponding to f' is

$$T_{f'}(\phi) = -\int_{-a}^{0} \phi(x)\,dx + \int_{0}^{a} \phi(x)\,dx$$

However, for $\phi \in \mathcal{D}(\mathbb{R})$, the distributional derivative of T_f is

$$\begin{aligned} DT_f(\phi) &= -T_f(\phi') = -\int_{-a}^{0} (-x)\phi'(x)\,dx - \int_{0}^{a} x\phi'(x)\,dx \\ &= x\phi(x)\mid_{-a}^{0} - \int_{-a}^{0} \phi(x)\,dx + \int_{0}^{a} \phi(x)\,dx - x\phi(x)\mid_{0}^{a} \\ &= a\phi(-a) - a\phi(a) + \int_{0}^{a} \phi(x)\,dx - \int_{-a}^{0} \phi(x)\,dx \\ &= a(\delta_{-a}(\phi) - \delta_a(\phi)) + T_{f'}(\phi). \end{aligned}$$

Therefore, $DT_f = a\delta_{-a} - a\delta_a + T_{f'}$. The function f is not weakly differentiable (compare this with Example 1.22).

Example 1.27. Consider the discontinuous function

$$f(x) = \begin{cases} x^2 + x & x < 1 \\ e^{-5x} & x > 1. \end{cases}$$

Classically, its derivative exists a.e. and is the function

$$f'(x) = \begin{cases} 2x + 1 & x < 1 \\ -5e^{-5x} & x > 1. \end{cases}$$

The distribution corresponding to f' is

$$T_{f'}(\phi) = \int_{-\infty}^{1} (2x+1)\phi(x)\,dx - 5\int_{1}^{\infty} e^{-5x}\phi(x)\,dx.$$

The distributional derivative of f (T_f) is

$$\begin{aligned}
DT_f(\phi) &= -T_f(\phi') = -\int_{-\infty}^{1} (x^2+x)\phi'(x)\,dx - \int_{1}^{\infty} e^{-5x}\phi'(x)\,dx \\
&= -(x^2+x)\phi(x)\,|_{-\infty}^{1} + \int_{-\infty}^{1} (2x+1)\phi(x)\,dx - 5\int_{1}^{\infty} e^{-5x}\phi(x)\,dx \\
&\quad - e^{-5x}\phi(x)\,|_{1}^{\infty} \\
&= (e^{-5}-2)\phi(1) + T_{f'}(\phi).
\end{aligned}$$

Therefore, $DT_f = (e^{-5}-2)\delta_1 + T_{f'}$. The function f is not weakly differentiable.

Example 1.28. Consider the Cantor function $f_C : [0,1] \to [0,1]$ (cf. Appendix A) extended continuously to $\mathbb{R}$ by setting

$$f_C(x) = \begin{cases} 0 & x \le 0 \\ 1 & x \ge 1. \end{cases}$$

Note that $f'_C = 0$ for all $x \le 0$ and $x \ge 1$. Since C is of Lebesgue measure zero and f_C is constant on each interval removed from $[0,1]$, we have $f'_C = 0$ a.e. in $\mathbb{R}$. Note that

$$\begin{aligned}
DT_{f_C}(\phi) &= -T_{f_C}(\phi') = -\int_0^1 f_C\phi'\,dx - \int_1^\infty \phi'\,dx = \phi(1) - \int_{[0,1]\setminus C} f_C\phi'\,dx \\
&= \phi(1) - \int_{\cup_{i=1}^\infty C_i^c} f_C\phi'\,dx \quad \text{(cf. Appendix A)} \\
&= \phi(1) - \sum_{i=1}^{\infty}\int_{C_i^c} f_C\phi'\,dx \quad \text{(since the union is disjoint)} \\
&= \phi(1) - \sum_{k=1}^{\infty} c_k \int_{a_k}^{b_k} \phi'\,dx \\
&\qquad (a_k, b_k \text{ are end-points of intervals removed}) \\
&= \phi(1) + \sum_{k=1}^{\infty} c_k(\phi(a_k) - \phi(b_k)) = \sum_{k=1}^{\infty} c_k(\delta_{a_k} - \delta_{b_k})(\phi) + \delta_1(\phi).
\end{aligned}$$

Note that f_C is not weakly differentiable. Moreover, the classical derivative and distributional derivative do not coincide.

Example 1.29. The k-th distributional derivative of the Dirac distribution is the k-th multipole distribution. For $\phi \in \mathcal{D}(\Omega)$,

$$\begin{aligned} D^\beta \delta_a^\alpha(\phi) &= (-1)^{|\beta|}\delta_a^\alpha(D^\beta\phi) = (-1)^{|\beta|}D^\alpha(D^\beta\phi(a)) \\ &= (-1)^{|\beta|}D^{\alpha+\beta}\phi(a)) = (-1)^{|\beta|}\delta_a^{\alpha+\beta}(\phi). \end{aligned}$$

In particular, the distributional derivative of the Dirac distribution is the dipole distribution, up to a factor of sign.

Example 1.30. Let us compute the distributional derivative of the distribution induced by $\ln|x|$. For supp$(\phi)\subset[-a,a]$,

$$\begin{aligned} DT_{\ln|x|}(\phi) &= -T_{\ln|x|}(\phi') = -\int_{\mathbb{R}} \ln|x|\phi'\,dx \\ &= -\lim_{\varepsilon\to 0^+}\left(\int_{-a}^{-\varepsilon}\ln|x|\phi'\,dx + \int_{\varepsilon}^{a}\ln|x|\phi'\,dx\right) \\ &= \lim_{\varepsilon\to 0^+}\left[(\phi(\varepsilon)-\phi(-\varepsilon))\ln\varepsilon + \int_{-a}^{-\varepsilon}\frac{\phi(x)}{x}\,dx + \int_{\varepsilon}^{a}\frac{\phi(x)}{x}\,dx\right] \\ &= \lim_{\varepsilon\to 0^+}\int_{\varepsilon}^{a}\frac{\phi(x)-\phi(-x)}{x}\,dx = \lim_{\varepsilon\to 0^+}\int_{|x|>\varepsilon}\frac{\phi(x)}{x}\,dx \\ &= PV\left(\frac{1}{x}\right)(\phi). \end{aligned}$$

Exercise 26. For any $\alpha\in\mathbb{R}$, we have already noted that the function $f(x) = |x|^{-\alpha}$ is in $L^1_{\text{loc}}(\mathbb{R}^n)$ for $\alpha < n$. However, the function is weakly differentiable for $\alpha+1<n$ and its weak partial derivative is $f_{x_i}(x) = -\frac{\alpha}{|x|^{\alpha+1}}\frac{x_i}{|x|}$.

Exercise 27. Let $\phi\in\mathcal{D}(\mathbb{R})$. Show that $\int_{\mathbb{R}}\phi(x)\,dx = 0$ iff there is a $\psi\in\mathcal{D}(\mathbb{R})$ such that $\phi(x) = \psi'(x)$.

Proof. Let there exist a $\psi\in\mathcal{D}(\mathbb{R})$ such that $\phi(x)=\psi'(x)$. Then,

$$\int_{\mathbb{R}}\phi(x)\,dx = \lim_{a\to\infty}\int_{-a}^{a}\psi'(x)\,dx = \lim_{a\to\infty}[\psi(a)-\psi(-a)] = 0.$$

Conversely, let $\int_{\mathbb{R}}\phi(x)\,dx = 0$. Then, set $\psi(x) := \int_{-\infty}^{x}\phi(t)\,dt$ and $\psi'=\phi$. It only remains to show that $\psi\in\mathcal{D}(\mathbb{R})$. $\psi\in C^\infty(\mathbb{R})$ since $\phi\in C^\infty(\mathbb{R})$.

Suppose that $\mathrm{supp}(\phi) \subset [-M, M]$. Then for all $x \in (\infty, -M)$, $\psi(x) = 0$ since $\phi(x) = 0$. For all $x \in (M, \infty)$, $\psi(x) = 0$ because $\int_{-M}^{M} \phi = 0$. Thus, $\mathrm{supp}(\psi) \subset [-M, M]$. □

Exercise 28. Let $T \in \mathcal{D}'(\mathbb{R})$ be such that $DT = 0$, then show that T is a regular distribution generated by a constant (a.e.) function.

Proof. Let $T \in \mathcal{D}'(\mathbb{R})$ be such that $DT(\phi) = 0$ for all $\phi \in \mathcal{D}(\mathbb{R})$. Therefore,

$$0 = DT(\phi) = -T(\phi')$$

for all $\phi \in \mathcal{D}(\mathbb{R})$. Consider the subset $E \subset \mathcal{D}(\mathbb{R})$, defined as

$$E := \{\psi \in \mathcal{D}(\mathbb{R}) \mid \psi = \phi' \text{ for some } \phi \in \mathcal{D}(\mathbb{R})\}.$$

We have $T(\psi) = 0$ for all $\psi \in E$. Using Exercise 27, we know that $\psi \in E$ iff $\int_{\mathbb{R}} \psi(x)\, dx = 0$. Choose any $\chi \in \mathcal{D}(\mathbb{R})$ such that $\int_{\mathbb{R}} \chi = 1$ and, for each $\phi \in \mathcal{D}(\mathbb{R})$, set $\psi := \phi - \alpha\chi$ where $\alpha = \int_{\mathbb{R}} \phi$. Then $\psi \in E$ and $T(\psi) = 0$. Therefore,

$$T(\phi) = \alpha T(\chi) = \int_{\mathbb{R}} T(\chi)\phi(x)\, dx.$$

Thus, $T = T_\lambda$ where $\lambda = T(\chi)$, a constant. □

An alternate proof of above exercise is given in Exercise 44. By "regular distribution generated by a constant function" we mean the function is constant except on a set of measure zero. (cf. Example 1.23). The same is not true with weak derivative. If the weak derivative is zero, then the function is constant on each connected component of Ω.

Exercise 29. Solve the equation $DT = S$, for any given distribution $S \in \mathcal{D}'(\mathbb{R})$.

Note that the distributional derivative of a locally integrable functions that are also differentiable in classical sense, admit two notions of differentiability. We have already seen in the examples above that the two notions of differentiability need not coincide. For smooth functions the two notions do coincide. Let $f \in C^\infty(\Omega)$ and T_f be its induced distribution. For any n-tuple α, $g := D^\alpha f \in C^\infty(\Omega)$ and let T_g be its induced distribution. Then

$$\begin{aligned} D^\alpha T_f(\phi) &= (-1)^{|\alpha|} T_f(D^\alpha \phi) = (-1)^{|\alpha|} \int f D^\alpha \phi\, dx \\ &= \int D^\alpha f \phi\, dx = \int g\phi = T_g(\phi). \end{aligned}$$

Thus, the two notions of derivatives coincide for all functions which respect integration by parts.

Proposition 1.5.3. *Let f, g be integrable functions on $[a, b]$ and c, d be some real constants. Set*

$$F(x) := c + \int_a^x f(t)\,dt; \quad \text{and } G(x) := d + \int_a^x g(t)\,dt.$$

Then,

$$\int_a^b [F(x)g(x) + G(x)f(x)]\,dx = F(b)G(b) - F(a)G(a).$$

Proof. Note that $c = F(a)$ and $d = G(a)$. Now, consider

$$\begin{aligned}
\int_a^b [Fg + Gf]\,dt &= \int_a^b g(t)\left(F(a) + \int_a^t f(s)\,ds\right)\,dt \\
&\quad + \int_a^b f(t)\left(G(a) + \int_a^t g(s)\,ds\right)\,dt \\
&= F(a)[G(b) - G(a)] + \int_a^b g(t)\left(\int_a^t f(s)\,ds\right)\,dt \\
&\quad + G(a)[F(b) - F(a)] + \int_a^b f(t)\left(\int_a^t g(s)\,ds\right)\,dt.
\end{aligned}$$

Set $s = x$, $t = y$ in the second integral and $s = y$, $t = x$ in the last integral. Then,

$$\begin{aligned}
\int_a^b [Fg + Gf]\,dt &= \int_a^b \int_a^y f(x)g(y)\,dx\,dy + \int_a^b \int_a^x f(x)g(y)\,dy\,dx \\
&\quad + F(a)G(b) + G(a)F(b) - 2F(a)G(a).
\end{aligned}$$

Consider the square $[a, b] \times [a, b]$ in $\mathbb{R}^2$ and its bisection by the line $y = x$ in $\mathbb{R}^2$. Then the first integral is evaluated in the region above the bisecting line and the second integral is evaluated below the bisecting line. Thus, combined together the region of computation is the square $[a,b] \times [a,b]$. Hence,

$$\begin{aligned}
\int_a^b [Fg + Gf]\, dt &= \int_a^b \int_a^b f(x)g(y)\, dx\, dy + F(a)G(b) + G(a)F(b) \\
&\quad - 2F(a)G(a) \\
&= \left(\int_a^b f(x)\, dx\right)\left(\int_a^b g(y)\, dy\right) + F(a)G(b) + G(a)F(b) \\
&\quad - 2F(a)G(a) \\
&= [F(b) - F(a)][G(b) - G(a)] + F(a)G(b) + G(a)F(b) \\
&\quad - 2F(a)G(a) \\
&= F(b)G(b) - F(a)G(a).
\end{aligned}$$

□

We know that F and G are bounded variation and $F' = f$ and $G' = g$. Thus, we have the following consequence:

Corollary 1.5.4. *If $f, g \in AC(\mathbb{R})$ then, for each subinterval $[a,b] \subset \mathbb{R}$,*

$$\int_a^b [f(x)g'(x) + f'(x)g(x)]\, dx = f(b)g(b) - f(a)g(a).$$

In particular, if $f \in AC(\mathbb{R})$ then, for all $\phi \in \mathcal{D}(\mathbb{R})$,

$$\int_{\mathbb{R}} f(x)\phi'(x) = -\int_{\mathbb{R}} f'(x)\phi(x)\, dx.$$

The above proposition shows that, in one dimensional case, the integration by parts formula can be extended, precisely, to the class of absolutely continuous functions. Every absolutely continuous function on $\mathbb{R}$ is of bounded variation and admits a derivative almost everywhere. We have the following inclusions:

$$AC(\mathbb{R}) \subsetneq C(\mathbb{R}) \cap BV(\mathbb{R}) \subsetneq BV(\mathbb{R}) \subsetneq L^1_{\text{loc}}(\mathbb{R}) \subsetneq R(\mathbb{R}) \subsetneq \mathcal{D}'(\mathbb{R}).$$

We know that $f \in AC(\mathbb{R})$ iff $f(b) - f(a) = \int_a^b f'(t)\, dt$ for all subintervals $[a,b] \subset \mathbb{R}$.

Proposition 1.5.5. *Conversely, if $f : \mathbb{R} \to \mathbb{R}$ is an integrable function such that*

$$\int_{\mathbb{R}} f(x)\phi'(x)\,dx = -\int_{\mathbb{R}} g(x)\phi(x)\,dx,$$

for some integrable function $g : \mathbb{R} \to \mathbb{R}$ and for all $\phi \in \mathcal{D}(\mathbb{R})$, then there is a $h \in AC(\mathbb{R})$ such that $h = f$ a.e. and $h' = g$.

Remark 1.5.6. Every $f \in \mathrm{AC}(\mathbb{R})$ can be associated with a measure μ_f, defined as $\mu_f(a,b) = f(b) - f(a)$, for every subinterval $(a,b) \subset \mathbb{R}$. Note that this measure is different from the measure $\mu_{f'}$ induced by $f' \in L^1_{\mathrm{loc}}(\mathbb{R})$ as in (1.4.2). Observe that $T_f \neq T_{\mu_f}$. In fact, $T_{\mu_f} = T_{f'}$.

We have already seen in Example 1.28 that for the Cantor function f_C, which is not absolutely continuous, the two derivatives do not coincide.

Exercise 30. Compute and compare the classical and distributional derivative of

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x^2} & x \neq 0 \\ 0 & x = 0. \end{cases}$$

Theorem 1.5.7. *The differential operator D^α, for each α, is a continuous map from $\mathcal{D}'(\Omega)$ to $\mathcal{D}'(\Omega)$.*

Proof. It is enough to show that when $T_m \to T$ in $\mathcal{D}'(\Omega)$ then $D^\alpha T_m \to D^\alpha T$ in $\mathcal{D}'(\Omega)$, for all multi-index α. Consider

$$\lim_{m\to\infty} D^\alpha T_m(\phi) = \lim_{m\to\infty} (-1)^\alpha T_m(D^\alpha \phi) = (-1)^\alpha T(D^\alpha \phi) = D^\alpha T(\phi).$$

□

Example 1.31. The above theorem is very special to the space of distributions $\mathcal{D}'(\Omega)$. For instance, the result is not true in $C^\infty(\Omega)$. Consider the sequence $f_m = (1/\sqrt{m}) \sin mx$ that converges uniformly to 0 and hence converges to 0 in the distribution sense too. However, it's derivative $f'_m(x) = \sqrt{m} \cos mx$ does not converge pointwise but converges to 0 in the distribution sense.

Corollary 1.5.8 (Term-by term differentiation of series)**.** *If $S := \sum_{i=1}^\infty T_i$, then $D^\alpha S = \sum_{i=1}^\infty D^\alpha T_i$.*

1.5.2 Product

Now that we have addition of distributions, a natural question is whether one can define product of any two distributions. The answer is in negation. The space of distribution cannot be made an algebra which extends the classical notion of point-wise multiplication. This is called the *Schwartz impossibility result*. L. Schwartz himself showed that it is not possible to define product of Dirac distributions, i.e., $\delta_0 \cdot \delta_0$. However, one may define product of a distribution with a C^∞ function which is a generalisation of the point-wise multiplication of two functions. To motivate this definition, we note that if $f \in L^1_{\text{loc}}(\Omega)$ and $\psi \in C^\infty(\Omega)$ then

$$\int_\Omega [\psi(x)f(x)]\phi(x)\,dx = \int_\Omega f(x)[\psi(x)\phi(x)]\,dx.$$

If T_f is the distribution induced by f, then above equality is same as saying $T_{\psi f}(\phi) = T_f(\psi\phi)$. But $T_f(\psi\phi)$ makes sense only when $\psi\phi \in \mathcal{D}(\Omega)$. This is the reason we demand $\psi \in C^\infty(\Omega)$ because then $\text{supp}(\psi\phi) \subseteq \text{supp}(\phi)$ and $\psi\phi$ has compact support. Also, by Leibniz' rule,

$$D^\alpha(\psi\phi) = \sum_{\beta \leq \alpha} \frac{\alpha!}{\beta!(\alpha-\beta)!} D^\beta\psi\, D^{\alpha-\beta}\phi$$

and hence $\psi\phi \in \mathcal{D}(\Omega)$.

Definition 1.5.9. *Let $\psi \in C^\infty(\Omega)$ and $T \in \mathcal{D}'(\Omega)$, then we define the functional $\psi T : \mathcal{D}(\Omega) \to \mathbb{R}$ as $\psi T(\phi) = T(\psi\phi)$.*

Proposition 1.5.10. *For every $\psi \in C^\infty(\Omega)$, $\psi T \in \mathcal{D}'(\Omega)$.*

Proof. We need to show that the functional ψT is continuous on $\mathcal{D}(\Omega)$. For any $\phi \in C^\infty(K)$,

$$\begin{aligned}
|\psi T(\phi)| &= |T(\psi\phi)| \leq C_K \|\psi\phi\|_{N_K} = C_K \sum_{|\alpha|=0}^{N_K} \|D^\alpha(\psi\phi)\|_0 \\
&\leq C_K \sum_{|\alpha|=0}^{N_K} \sum_{\beta\leq\alpha} \left|\frac{\alpha!}{\beta!(\alpha-\beta)!}\right| \|D^\beta\psi\|_0\, \|D^{\alpha-\beta}\phi\|_0 \\
&\leq C_K C_0 \|\phi\|_{N_K}.
\end{aligned}$$

□

If T is a distribution of order k, then it is enough to demand $\psi \in C^k(\Omega)$.

Example 1.32. Let $\psi(x) = x$ on $\mathbb{R}$ and $T = \delta_a \in \mathcal{D}'(\mathbb{R})$, then $x\delta_a \in \mathcal{D}'(\mathbb{R})$ and $x\delta_a(\phi) = \delta_a(x\phi) = a\phi(a)$.

Theorem 1.5.11 (Leibniz' formula)**.** *If $\psi \in C^\infty(\Omega)$, then*

$$D^\alpha(\psi T) = \sum_{\beta \leq \alpha} \frac{\alpha!}{\beta!(\alpha-\beta)!} D^\beta \psi \, D^{\alpha-\beta} T.$$

Exercise 31. Let $\phi \in \mathcal{D}(\mathbb{R})$. Show that $\phi^{(k)}(0) = 0$ for $k = 0, 1, 2, \dots, m-1$ iff there is a $\psi \in \mathcal{D}(\mathbb{R})$ such that $\phi(x) = x^m \psi(x)$.

Proof. Let $\psi \in \mathcal{D}(\mathbb{R})$ be such that $\phi(x) = x^m \psi(x)$. Then

$$\phi^{(k)}(x) = \sum_{i=0}^{k} \frac{k!}{i!(k-i)!} \frac{d^i(x^m)}{dx^i} \psi^{(k-i)}$$

where $\frac{d^i(x^m)}{dx^i} = \frac{m!}{(m-i)!} x^{m-i}$. Thus, for all $i = 0, 1, \dots, m-1$, the value of $\frac{d^i(x^m)}{dx^i}$ at $x = 0$ is 0 and $\frac{d^m(x^m)}{dx^m} = m! \neq 0$. Therefore, $\phi^{(k)}(0) = 0$ for all $k = 0, 1, 2, \dots, m-1$, irrespective of the value of $\psi(0)$. Conversely, let $\phi^{(k)}(0) = 0$ for all $k = 0, 1, 2, \dots, m-1$. Let $k = 0$, then

$$\phi(x) = \phi(0) + \int_0^x \phi^{(1)}(t)\, dt.$$

Since $\phi(0) = 0$ and setting $t = xs$, we get

$$\phi(x) = \int_0^x \phi^{(1)}(t)\, dt = x \int_0^1 \phi^{(1)}(xs)\, ds = x\psi_0(x),$$

where $\psi_0(x) := \int_0^1 \phi^{(1)}(xs)\, ds$ is in $\mathcal{D}(\mathbb{R})$. Since $\phi^{(1)}(x) = x\psi_0^{(1)}(x) + \psi_0(x)$, we have $\psi(0) = 0$ and hence $\psi_0(x) = x\psi_1(x)$ for some $\psi_1 \in \mathcal{D}(\mathbb{R})$. Thus, $\phi(x) = x^2\psi_1(x)$. Using the fact that $\phi^{(2)}(0) = 0$, we get $\psi_1(0) = 0$ and hence $\phi(x) = x^3\psi_2(x)$ for some $\psi_2 \in \mathcal{D}(\mathbb{R})$. Proceeding this way, using $\phi^{(m-1)}(0) = 0$, we get $\psi_{m-1} \in \mathcal{D}(\mathbb{R})$ such that $\phi(x) = x^m\psi_{m-1}(x)$. □

Exercise 32. (i) Find all real valued functions f on $\mathbb{R}$ which is a solution to the equation $xf(x) = 0$ for all $x \in \mathbb{R}$.

(ii) Find all distributions $T \in \mathcal{D}'(\mathbb{R})$ that solve the equation $xT = 0$ in the distribution sense.

(iii) Find all distributions $T \in \mathcal{D}'(\mathbb{R})$ that solve the equation $x^{2013}T = 0$ in the distribution sense.

Proof. (i) Consider the function $f : \mathbb{R} \to \mathbb{R}$ such that

$$f(x) = \begin{cases} 0 & x \neq 0 \\ 1 & x = 0. \end{cases}$$

Then f solves $xf(x) = 0$ for all $x \in \mathbb{R}$. Moreover, for any scalar $\lambda \in \mathbb{R}$, the function λf is also a solution of $xf(x) = 0$.

(ii) For any $\phi \in \mathcal{D}(\mathbb{R})$, since $0 = xT(\phi) = T(x\phi)$. By Exercise 31, any solution $T \in \mathcal{D}'(\mathbb{R})$ satisfies $T(\psi) = 0$ for all $\psi \in \mathcal{D}(\mathbb{R})$ such that $\psi(0) = 0$. Obviously, $T \equiv \delta_0$ is a solution. In fact, for any scalar $\lambda\delta_0 \in \mathcal{D}'(\mathbb{R})$ is also a solution. We need to show that any solution is a scalar multiple of δ_0. Let $T \in \mathcal{D}'(\mathbb{R})$ be a solution of $xT = 0$. Hence $T(\psi) = 0$ for all test functions on $\mathbb{R}$ such that $\psi(0) = 0$. Let $\chi \in \mathcal{D}(\mathbb{R})$ such that $\chi(0) = 1$, then for any $\phi \in \mathcal{D}(\mathbb{R})$, $\psi(x) := \phi(x) - \phi(0)\chi(x)$ is a test function such that $\psi(0) = 0$. Therefore,

$$0 = T(\psi) = T(\phi) - T(\chi)\phi(0) = T(\phi) - T(\chi)\delta_0(\phi).$$

Thus, $T(\phi) = \lambda\delta_0(\phi)$ where the choice of λ depends on the choice of χ.

(iii) By Exercise 31, any solution $T \in \mathcal{D}'(\mathbb{R})$ satisfies $T(\psi) = 0$, for all $\psi \in \mathcal{D}(\mathbb{R})$, such that $\psi^{(k)}(0) = 0$ for all $k = 0, 1, 2, \ldots, 2012$. Therefore, $\delta_0^{(k)}$ is a solution for all $k = 0, 1, 2, \ldots, 2012$. Thus,

$$T \equiv \sum_{k=0}^{2012} \lambda_k \delta_0^{(k)}$$

is a solution for every possible choice of $\lambda_k \in \mathbb{R}$.

$\square$

Example 1.33. Let $\psi(x) = x$ on $\mathbb{R}$ and $T = PV(1/x) \in \mathcal{D}'(\mathbb{R})$, then $xPV(1/x) \in \mathcal{D}'(\mathbb{R})$ and

$$\begin{aligned} xPV\left(\frac{1}{x}\right)(\phi) &= PV\left(\frac{1}{x}\right)(x\phi) = \lim_{\varepsilon\to 0^+}\int_{|x|\geq\varepsilon}\frac{1}{x}x\phi(x)\,dx \\ &= \int_{\mathbb{R}}\phi(x)\,dx = T_1(\phi) \end{aligned}$$

where T_1 is the distribution corresponding to the constant function 1.

Exercise 33. Find all distributions $T \in \mathcal{D}'(\mathbb{R})$ that solve the equation $xT = 1$ in the distribution sense.

Proof. The principal value distribution $T \equiv PV(1/x)$ is one particular solution of $xT = 1$. Also any solution S of $xS = 0$ when added to $PV(1/x)$ is also a solution. Thus, the general solution is $PV(1/x) + \lambda\delta_0$ for every $\lambda \in \mathbb{R}$. □

In light of the above example, we come back to the issue of not being able to have a product in the space of distributions. Suppose we give a notion of product of distributions, the product cannot be associative because $PV(1/x)\cdot(x\cdot\delta_0) = 0 \neq \delta_0 = (PV(1/x)\cdot x)\cdot\delta_0$. Further, recall that $f(x) = |x|^{-1/2} \in L^1_{\mathrm{loc}}(\mathbb{R})$, but $g(x) = |x| \notin L^1_{\mathrm{loc}}(\mathbb{R})$. How do we give a notion of product, in a suitable way, such that $(T_f)^2$ coincides with classical pointwise product. The failure to define a suitable notion of product is what makes the theory of distributions unsuitable for nonlinear differential equations.

Exercise 34. For any given $\psi \in C^\infty(\mathbb{R})$, find the solutions of the equation $DT + \psi T = 0$ in $\mathcal{D}'(\mathbb{R})$.

Proof. If $\psi \equiv 0$, we know from Exercise 28 that T is a regular distribution induced by a constant function. Therefore, we assume w.l.o.g that $\psi \not\equiv 0$. We copy the idea of integrating factor from ODE course. We shall now prove the existence of $\Psi \in C^\infty(\mathbb{R})$ (nowhere zero) such that $D(\Psi T) = \Psi(DT + \psi T)$. Multiply Ψ on both sides of the given equation to get $\Psi DT + \Psi\psi T = 0$. If we want Ψ such that $D(\Psi T) = \Psi(DT + \psi T)$, then $\Psi\psi = \Psi'$. Thus, $\Psi(x) = e^{\int_{-\infty}^x \psi(s)\,ds}$. By construction, Ψ is a non-zero function and the equation implies that $D(\Psi T) = 0$. By Exercise 28, for each scalar $\mu \in \mathbb{R}$, $\Psi T = T_\mu$ is a solution. Thus, $T = \Psi^{-1}T_\mu$, for each $\mu \in \mathbb{R}$ is the required solution of the equation. The solution makes sense because $\Psi^{-1} \in C^\infty(\mathbb{R})$. □

Exercise 35. For any given $\psi \in C^\infty(\mathbb{R})$ and $S \in \mathcal{D}'(\mathbb{R})$, find the solutions of the equation $DT + \psi T = S$ in $\mathcal{D}'(\mathbb{R})$.

Recall the topology on $C^\infty(\Omega)$ described in §1.3.4. The sequential description of the topology is the uniform convergence on compact sets. Thus, a sequence ψ_m is said to converge to 0 in $C^\infty(\Omega)$ if $D^\alpha \psi_m$, for all α, uniformly converges to 0 on all compact subsets of Ω.

Lemma 1.5.12. *If $\psi_m \to 0$ in $C^\infty(\Omega)$ then $\psi_m \phi \to 0$ in $\mathcal{D}(\Omega)$.*

Proof. Since $D^\alpha \psi_m \to 0$ uniformly on all compact subsets of Ω, in particular on $K = \text{supp}(\phi)$. Thus, $D^\alpha \psi_m \phi \to 0$ uniformly on K, for all α, and $\text{supp}(\psi_m \phi) \subseteq K$, for all m. □

Theorem 1.5.13. *If $T_m \to T$ in $\mathcal{D}'(\Omega)$ and $\psi_m \to \psi$ in $C^\infty(\Omega)$ then $\psi_m T_m \to \psi T$ in $\mathcal{D}'(\Omega)$.*

1.5.3 Support

The fact that we *cannot* have a product on $\mathcal{D}'(\Omega)$, generalising point-wise multiplication, that makes $\mathcal{D}'(\Omega)$ an algebra leads us to look for other ways to make $\mathcal{D}'(\Omega)$ an algebra. One such choice is the convolution operation and whether this can be extended to distributions. We shall see later than one can define a notion of convolution for distributions provided one of them has compact support! This motivates us to understand the support of a distribution which coincides for classical functions. Classically, support of a function is complement of the largest open set on which the function vanishes.

Definition 1.5.14 (Localisation)**.** *The* restriction *of $T \in \mathcal{D}'(\Omega)$ to an open subset $\omega \subset \Omega$, denoted as $T\mid_\omega$, is defined as*

$$T\mid_\omega(\phi) = T(\phi) \quad \forall \phi \in \mathcal{D}(\omega) \subset \mathcal{D}(\Omega).$$

We say a distribution T vanishes *on an open set $\omega \subset \Omega$ if $T\mid_\omega = 0$. More generally, two distributions S and T coincide on a open set ω if $S\mid_\omega = T\mid_\omega$.*

The inclusion $\mathcal{D}(\omega) \subset \mathcal{D}(\Omega)$ is continuous because any $\phi \in \mathcal{D}(\omega)$ can be extended by zero outside ω which belongs to $\mathcal{D}(\Omega)$. Since $T\mid_\omega \in \mathcal{D}'(\omega)$, T vanishing on ω is same as the zero distribution in $\mathcal{D}'(\omega)$.

Example 1.34. The restriction of the Dirac distribution δ_a to any open set $\omega \subset \Omega \setminus \{a\}$ is zero.

Proposition 1.5.15. *If $T \in \mathcal{D}'(\Omega)$ vanishes on the open subsets ω_i of Ω, then T vanishes on the union $\cup_{i\in I}\omega_i$.*

Proof. Let $\omega = \cup_{i\in I}\omega_i$. We choose the C^∞ locally finite partition of unity $\{\phi_i\}$ subordinate to $\{\omega_i\}$, i.e., $\text{supp}(\phi_i) \subset \omega_i$ and $\sum_i \phi = 1$ (the summation is finite for each $x \in \omega$). Fix $\phi \in \mathcal{D}(\omega)$ and $K = \text{supp}(\phi)$ then there exists finitely many $i_1, i_2, \dots, i_k$ such that $K \cap \phi_{i_m} \neq \emptyset$. Thus, $\phi = \sum_{m=1}^{k} \phi\phi_{i_m}$ and, by linearity of T, $T(\phi) = \sum_{m=1}^{k} T(\phi\phi_{i_m})$. But $\text{supp}(\phi\phi_{i_m}) \subset \omega_{i_m}$ and hence $T(\phi\phi_{i_m}) = 0$. Therefore, $T(\phi) = 0$ for all $\phi \in \mathcal{D}(\omega)$ and hence $T\mid_\omega = 0$. □

Definition 1.5.16. *The* support *of a distribution T is the complement of the largest open set ω such that $T\mid_\omega = 0$. Equivalently, the support of a distribution is the relative complement in Ω of the union of all open sets ω of Ω such that $T\mid_\omega = 0$.*

If E is the support of a distribution T, then $T(\phi) = 0$ for all $\phi \in \mathcal{D}(\Omega)$ such that $E \cap \text{supp}(\phi) = \emptyset$ because all such ϕ is in $\mathcal{D}(E^c)$. Obviously, the support of the zero distribution is the empty set.

Exercise 36. Show that for any continuous function f, $\text{supp}(T_f)$=$\text{supp}(f)$. More generally, for any $f \in L^1_{\text{loc}}(\Omega)$, there exists a measurable function g such that $g = f$ a.e. and $\text{supp}(T_f) = \text{supp}(g)$.

Example 1.35. The support of the Dirac distribution δ_a is the singleton set $\{a\}$. Similarly, the support of the derivatives of Dirac distribution δ_a^α is also the singleton set $\{a\}$. The open set $\Omega \setminus \{a\} = \cup_i B_i$, is the union of punctured open balls of rational radius r_i with centre $\{a\}$ removed. δ_a^α vanishes on B_i for each i and hence vanishes on $\Omega \setminus \{a\}$.

Exercise 37. Show that the support of $DT \in \mathcal{D}(\mathbb{R})$ is contained in the support of $T \in \mathcal{D}(\mathbb{R})$.

Proof. Let E and F denote the support of T and DT, respectively. We need to show that $F \subseteq E$, or equivalently, $E^c \subseteq F^c$. For any open set $\omega \subset E^c$, then $T\mid_\omega = 0$, i.e., for all $\phi \in \mathcal{D}(\omega)$ we have $T(\phi) = 0$. Consider $DT(\phi) = -T(\phi') = 0$. The second equality is due to the fact that if $\phi \in \mathcal{D}(\omega)$ then $\phi' \in \mathcal{D}(\omega)$ with $\text{supp}(\phi') \subseteq \text{supp}(\phi)$. □

A distribution is said to have *compact support*, if its support is a compact subset of Ω. A nice characterisation of distributions with compact support, which is a subclass of $\mathcal{D}'(\Omega)$, is that it can be identified as a dual of $C^\infty(\Omega)$ endowed with the topology of uniform convergence on compact sets. This observation is actually quite natural. Recall the motivation for the choice of $C_c^\infty(\Omega)$ as a test function. The compact support of test function was necessary to kill the boundary evaluation while applying integration by parts. If the function f itself is compact, to begin with, then we can expand the space of test functions to $C^\infty(\Omega)$. For convenience sake (for ease of notation), we set $C^\infty(\Omega) = \mathcal{E}(\Omega)$, the space of infinitely differentiable functions on Ω with the topology of uniform convergence on compact subsets. The inclusion $\mathcal{D}(\Omega) \subset \mathcal{E}(\Omega)$ is continuous w.r.t the respective topology. This is because if $\phi_m \to 0$ in $\mathcal{D}(\Omega)$ then $\phi_m \to 0$ in $\mathcal{E}(\Omega)$, as well. The following theorem asserts that any distribution $T : \mathcal{D}(\Omega) \to \mathbb{R}$ can be continuously extended as a functional to $\mathcal{E}(\Omega)$. Thus $T \in \mathcal{E}'(\Omega)$, for compact support T.

Theorem 1.5.17. *The dual space $\mathcal{E}'(\Omega)$ is the collection of all distributions with compact support.*

Proof. Let $T \in \mathcal{D}'(\Omega)$ such that support of T is compact. Set $K = \operatorname{supp}(T)$. Thus, $T(\phi) = 0$ for all $\phi \in \mathcal{D}(\Omega)$ such that $\operatorname{supp}(\phi) \subset K^c$. Choose $\chi \in \mathcal{D}(\Omega)$ such that $\chi \equiv 1$ on K. For all $\phi \in \mathcal{D}(\Omega)$, consider

$$T(\phi) - \chi T(\phi) = T(\phi - \chi\phi) = T((1 - \chi)\phi).$$

Since $1 - \chi = 0$ on K, $\operatorname{supp}(1 - \chi)\phi \subset K^c$ and hence $T = \chi T$ for any χ such that $\chi \equiv 1$ on K. This fact helps us to extend T as a functional on $\mathcal{E}(\Omega)$, because for all $\psi \in \mathcal{E}(\Omega)$, $\psi\chi \in \mathcal{D}(\Omega)$. We extend T to $\mathcal{E}(\Omega)$ by setting $T(\psi) = T(\chi\psi)$. The assignment is well-defined (it is independent of choice of χ). Suppose $\chi_1 \in \mathcal{D}(\Omega)$ is such that $\chi_1 \neq \chi$ and $\chi_1 \equiv 1$ on K, then $(\chi - \chi_1)\psi = 0$ on K. Therefore, $\operatorname{supp}((\chi - \chi_1)\psi) \subset K^c$ and $T((\chi - \chi_1)\psi) = 0$, By linearity of T, $T(\chi\psi) = T(\chi_1\psi)$. The conitnuity of T on $\mathcal{E}(\Omega)$ is a consequence of Lemma 1.5.12. Thus, $T \in \mathcal{E}'(\Omega)$.

Conversely, let $T \in \mathcal{E}'(\Omega)$. We need to show that $T \in \mathcal{D}'(\Omega)$ and has a compact support. The fact that $T \in \mathcal{D}'(\Omega)$ is obvious due to the continuous inclusion of $\mathcal{D}(\Omega)$ in $\mathcal{E}(\Omega)$. Let $\{K_i\}$ be the collection of compact subsets of Ω such that $K_i \subset K_{i+1}$ and $\Omega = \cup_{i=1}^\infty K_i$ (exhaustion subsets). Suppose that the support of T is not compact in Ω, then the support of T intersects $\Omega \setminus K_i$, for all i. Thus, for each i, there is a $\phi_i \in \mathcal{D}(\Omega)$ such that $\operatorname{supp}(\phi_i) \subset \Omega \setminus K_i$ and $T(\phi_i) \neq 0$. Say $T(\phi_i) = \alpha_i \neq 0$. Then $T(\alpha_i^{-1}\phi_i) = 1$. Set $\psi_i = \alpha_i^{-1}\phi_i$. Thus, $\operatorname{supp}(\psi_i) \subset \Omega \setminus K_i$ and $T(\psi_i) = 1$. Hence, the sequence $\psi_i \to 0$ in $\mathcal{E}(\Omega)$ because their support tend to empty set. Now, by continuity of T on $\mathcal{E}(\Omega)$, $T(\psi_i) \to 0$ which contradicts the fact that $T(\psi_i) \to 1$. Hence, T has a compact support. $\square$

We have proved that $\mathcal{E}'(\Omega) \subset \mathcal{D}'(\Omega)$ is the class of all distributions whose support is compact. Distributions with compact support are always of finite order.

Proposition 1.5.18. *If $T \in \mathcal{E}'(\Omega)$ then order of T is finite.*

Proof. Since $T \in \mathcal{E}'(\Omega)$, T has compact support, say K. Let $\chi \in \mathcal{D}(\Omega)$ be such that $\chi \equiv 1$ on K, then we know that (cf. proof of last theorem) $T = \chi T$. Let K' be the support of χ and hence $K \subset K'$. Now, for any compact subset $K \subset \Omega$ and $\phi \in C^\infty(K)$, $\operatorname{supp}(\phi\chi) \subset K'$. Consider,

$$|T(\phi)| = |\chi T(\phi)| = |T(\chi\phi)| \leq C_{K'} \|\chi\phi\|_{N_{K'}} \leq C_0 \|\phi\|_{N_{K'}}.$$

The $N_{K'}$ is independent of all compact subsets K. Thus, T has finite order. □

1.5.4 Singular Support

The singular support of a distribution is the measure of the set on which the distribution fails to be smooth or C^∞.

Definition 1.5.19. *A distribution $T \in \mathcal{D}'(\Omega)$ is said to be C^∞ on $\omega \subset \Omega$ if there is a $f \in C^\infty(\omega)$ such that $T = T_f$ on $\mathcal{D}(\omega)$.*

Example 1.36. The Dirac distribution, δ_a, is C^∞ in $\Omega \setminus \{a\}$ because δ_a restricted to the open set $\Omega \setminus \{a\}$ is the zero function.

Local behaviour of “compatible” distributions can be patched up to get a global description of the distribution. This concept will facilitate the definition of singular support.

Theorem 1.5.20. *Let $T_i \in \mathcal{D}'(\omega_i)$, an arbitrary collection of distributions, such that $T_i \mid_{\omega_i \cap \omega_j} = T_j \mid_{\omega_i \cap \omega_j}$ then there is a unique distribution $T \in \mathcal{D}'(\Omega)$, where $\Omega = \cup_i \omega_i$, such that $T \mid_{\omega_i} = T_i$ for all i.*

Proof. Let $\{\phi_i\} \subset \mathcal{D}(\Omega)$ be a C^∞ locally finite partition of unity subordinate to $\{\omega_i\}$. Thus, $\operatorname{supp}(\phi_i) \subset \omega_i$ and $1 = \sum_i \phi_i$. We define the functional $T : \mathcal{D}(\Omega) \to \mathbb{R}$ as $T(\phi) = \sum_i T_i(\phi\phi_i)$. We first check the continuity of T. One way is to show the sequential continuity. Alternately, For any $\phi \in$

$C^\infty(K)$ (compact subset of Ω), there exist finitely many $i_1, i_2, \ldots, i_k$ such that $K \cap \mathrm{supp}(\phi_{i_m}) \neq \emptyset$. Then

$$|T(\phi)| \leq \sum_{m=1}^{k} |T_{i_m}(\phi\phi_{i_m})| \leq \sum_m C_k \|\phi\phi_{i_m}\|_{N_k} \leq C_0\|\phi\|_{N_0}.$$

It only remains to show that the restriction of T to ω_k is T_k. Let $\phi \in \mathcal{D}(\omega_k)$, then $\phi\phi_i \in \mathcal{D}(\omega_i \cap \omega_k)$ for all i. Thus,

$$T(\phi) = \sum_i T_i(\phi\phi_i) = \sum_i T_k(\phi\phi_i) = T_k(\sum_i \phi\phi_i) = T_k(\phi).$$

The uniqueness of T follows from Proposition 1.5.15. □

Corollary 1.5.21. *If $T \in \mathcal{D}'(\Omega)$ is C^∞ on an arbitrary collection of open subsets ω_i, then T is C^∞ on the union $\cup_{i\in I}\omega_i$.*

A consequence of the above corollary is the following definition.

Definition 1.5.22. *The* singular support *of $T \in \mathcal{D}'(\Omega)$ is the complement of the largest open set on which T is C^∞, denoted as sing.supp(T).*

A simple observation is that sing.supp$(T) \subset$ supp(T).

1.5.5 Shifting and Scaling

Given any function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and a fixed $a \in \mathbb{R}^n$, we introduce the shift operator $\tau_a f(x) := f(x-a)$. In particular, $\tau_0 f = f$ for all f. Observe that $\tau_a\tau_b f(x) = \tau_a f(x-b) = f(x-b-a) = f(x-(b+a)) = \tau_{a+b}f(x)$.

Note that if $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ then $\tau_a f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. We have the following relation between T_f and $T_{\tau_a f}$. For any $\phi \in \mathcal{D}(\mathbb{R}^n)$, consider

$$\begin{aligned} T_{\tau_a f}(\phi) &= \int_{\mathbb{R}^n} \tau_a f(x)\phi(x)\,dx = \int_{\mathbb{R}^n} f(x-a)\phi(x)\,dx = \int_{\mathbb{R}^n} f(y)\phi(y+a)\,dy \\ &= \int_{\mathbb{R}^n} f(y)\tau_{-a}\phi(y)\,dy = T_f(\tau_{-a}\phi). \end{aligned}$$

The last equality is valid because $\tau_{-a}\phi \in \mathcal{D}(\mathbb{R}^n)$ whenever $\phi \in \mathcal{D}(\mathbb{R}^n)$. We wish to extend the notion of shift to distributions such that for $\tau_a T_f = T_{\tau_a f}$. This motivates the following definition of a shift of a distribution by $a \in \mathbb{R}^n$.

Definition 1.5.23. *For any $T \in \mathcal{D}'(\mathbb{R}^n)$, we define its* shift by $a \in \mathbb{R}^n$ *as* $\tau_a T(\phi) = T(\tau_{-a}\phi)$.

Given any function $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ and a fixed scalar $\lambda \in (0, \infty)$, we introduce the scaling function $f_\lambda(x) := f(\lambda x)$. If $f = g + h$ then $f_\lambda(x) = f(\lambda x) = g(\lambda x) + h(\lambda x) = g_\lambda(x) + h_\lambda(x)$. Also, $\tau_a(f + g) = \tau_a f + \tau_a g$. Note that if $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ then $f_\lambda \in L^1_{\text{loc}}(\mathbb{R}^n)$. We have the following relation between T_f and T_{f_λ}. For any $\phi \in \mathcal{D}(\mathbb{R}^n)$, consider

$$\begin{aligned} T_{f_\lambda}(\phi) &= \int_{\mathbb{R}^n} f(\lambda x)\phi(x)\, dx = \lambda^{-n} \int_{\mathbb{R}^n} f(y)\phi\left(\frac{y}{\lambda}\right)\, dy \\ &= \lambda^{-n} \int_{\mathbb{R}^n} f(y)\phi_{1/\lambda}(y)\, dy = \lambda^{-n} T_f(\phi_{1/\lambda}). \end{aligned}$$

The last equality is valid because $\phi_{1/\lambda} \in \mathcal{D}(\mathbb{R}^n)$ whenever $\phi \in \mathcal{D}(\mathbb{R}^n)$. Similarly, one can argue for $\lambda \in (-\infty, 0)$ and deduce that

$$T_{f_\lambda}(\phi) = (-1)^n \lambda^{-n} T_f(\phi_{1/\lambda}).$$

Thus for any $\lambda \in \mathbb{R} \setminus \{0\}$, we have $T_{f_\lambda}(\phi) = |\lambda|^{-n} T_f(\phi_{1/\lambda})$. We wish to extend the notion of scaling to distributions such that for $(T_f)_\lambda = T_{f_\lambda}$. This motivates the following definition of a scaling of a distribution by $\lambda \in \mathbb{R} \setminus \{0\}$.

Definition 1.5.24. *For any $T \in \mathcal{D}'(\mathbb{R}^n)$, we define its* scaling by $\lambda \in \mathbb{R} \setminus \{0\}$ *as $T_\lambda(\phi) = |\lambda|^{-n} T(\phi_{1/\lambda})$. In particular, when $\lambda = -1$, we denote T_{-1} as $\check{T}$ and hence $\check{T}(\phi) = T(\check{\phi})$. We say $T \in \mathcal{D}'(\mathbb{R}^n)$ is even if $\check{T} = T$.*

Exercise 38. Show that $\check{\check{T}} = T$.

1.5.6 Convolution

Recall the definition of convolution of functions given in Definition 1.3.20. Thus, for any $x \in \mathbb{R}^n$,

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

In this section, we wish to extend the convolution operation to distributions such that for L^1 distributions the notion coincide. An obvious extension is visible when the convolution integral is rewritten using the shift and scaling operator. Note that, for a fixed $x \in \mathbb{R}^n$, $f(x-y) = f(-(y-x)) = \check{f}(y-x) = \tau_x \check{f}(y)$. With this notation the convolution is rewritten as

$$(f * g)(x) = \int_{\mathbb{R}^n} \tau_x \check{f}(y) g(y)\, dy = \int_{\mathbb{R}^n} f(y) \tau_x \check{g}(y)\, dy.$$

The second equality is due to the commutativity of convolution operation. The reformulation above motivates the following definition of convolution between a distribution and a test function.

Definition 1.5.25. *Let $T \in \mathcal{D}'(\mathbb{R}^n)$ and $\phi \in \mathcal{D}(\mathbb{R}^n)$. The* convolution $T * \phi : \mathbb{R}^n \to \mathbb{R}$ *is defined as*

$$(T * \phi)(x) = T(\tau_x \check{\phi}).$$

The above definition is meaningful because $\tau_x \check{\phi} \in \mathcal{D}(\mathbb{R}^n)$, for all $x \in \mathbb{R}^n$, whenever $\phi \in \mathcal{D}(\mathbb{R}^n)$. In particular, $(T * \check{\phi})(x) = T(\tau_x \phi)$ since $\check{\check{\phi}} = \phi$. Thus, $T(\phi) = (T * \check{\phi})(0)$. As a consequence, if $T * \phi = 0$ for all $\phi \in \mathcal{D}(\mathbb{R}^n)$ then $T = 0$ in $\mathcal{D}'(\mathbb{R}^n)$.

Exercise 39. (i) Show that $T * (\phi_1 + \phi_2) = T * \phi_1 + T * \phi_2$.

(ii) Show that $(S + T) * \phi = S * \phi + T * \phi$.

Proof. (i) Consider

$$T*(\phi_1+\phi_2)(x) = T(\tau_x(\phi_1+\phi_2)\check{}) = T(\tau_x\check{\phi}_1+\tau_x\check{\phi}_2) = (T*\phi_1+T*\phi_2)(x).$$

(ii) Consider

$$(S + T) * \phi = (S + T)(\tau_x \check{\phi}) = S * \phi + T * \phi.$$

□

Exercise 40. Let $T \in \mathcal{D}'(\mathbb{R}^n)$ and $\phi \in \mathcal{D}(\mathbb{R}^n)$, then

(i) for any $a \in \mathbb{R}^n$, show that $\tau_a(T * \phi) = \tau_a T * \phi = T * \tau_a \phi$.

(ii) show that, for any multi-index α, $D^\alpha(T * \phi) = D^\alpha T * \phi = T * D^\alpha \phi$. In particular, $T * \phi \in \mathcal{E}(\mathbb{R}^n)$.

(iii) for $\chi \in \mathcal{D}(\mathbb{R}^n)$, $T * (\phi * \chi) = (T * \phi) * \chi$.

For any $\psi \in \mathcal{E}(\mathbb{R}^n)$, $\tau_x\check{\psi} \in \mathcal{E}(\mathbb{R}^n)$. Thus, the convolution notion can be extended to any $\psi \in \mathcal{E}(\mathbb{R}^n)$ and $T \in \mathcal{E}'(\mathbb{R}^n)$.

Definition 1.5.26. *Let $T \in \mathcal{E}'(\mathbb{R}^n)$ and $\psi \in \mathcal{E}(\mathbb{R}^n)$. The* convolution $T * \psi : \mathbb{R}^n \to \mathbb{R}$ *is defined as*

$$(T * \psi)(x) = T(\tau_x\check{\psi}).$$

Exercise 41. Let $T \in \mathcal{E}'(\mathbb{R}^n)$ and $\psi \in \mathcal{E}(\mathbb{R}^n)$, then

(i) for any $a \in \mathbb{R}^n$, show that $\tau_a(T * \psi) = \tau_a T * \psi = T * \tau_a\psi$.

(ii) show that, for any multi-index α, $D^\alpha(T * \psi) = D^\alpha T * \psi = T * D^\alpha\psi$. In particular, $T * \phi \in \mathcal{E}(\mathbb{R}^n)$. Further, if $\phi \in \mathcal{D}(\mathbb{R}^n)$ then $T * \phi \in \mathcal{D}(\mathbb{R}^n)$ and

$$T * (\psi * \phi) = (T * \psi) * \phi = (T * \phi) * \psi.$$

Definition 1.5.27. *Let $S, T \in \mathcal{D}'(\mathbb{R}^n)$ such that one of them has compact support, i.e. either T or S is in $\mathcal{E}'(\mathbb{R}^n)$. We define the* convolution $S * T$ *as,*

$$(S * T)(\phi) = (S * (T * \check{\phi}))(0), \quad \forall \phi \in \mathcal{D}(\mathbb{R}^n).$$

Equivalently, one can also define the convolution $S * T$ to be

$$(S * T) * \phi = S * (T * \phi) \quad \forall \phi \in \mathcal{D}(\mathbb{R}^n).$$

Exercise 42. Let $S, T \in \mathcal{D}'(\mathbb{R}^n)$ such that one of them is in $\mathcal{E}'(\mathbb{R}^n)$.

(i) Show that $T * S = S * T$.

(ii) $\operatorname{supp}(S * T) \subset \operatorname{supp}(S) + \operatorname{supp}(T)$.

(iii) For any multi-index α, $D^\alpha(T * S) = D^\alpha T * S = T * D^\alpha S$.

Exercise 43. For any three distributions $T_1, T_2, T_3 \in \mathcal{D}'(\mathbb{R}^n)$ such that at least two of them have compact support then

$$T_1 * (T_2 * T_3) = (T_1 * T_2) * T_3.$$

Theorem 1.5.28. *Let $T \in \mathcal{D}'(\mathbb{R}^n)$, then $T = T * \delta_0 = \delta_0 * T$. Also, for any multi-index α, $D^\alpha T = D^\alpha\delta_0 * T$.*

Proof. Since δ_0 has compact support $\{0\}$, the convolution makes sense. By commutativity, $T * \delta_0 = \delta_0 * T$. For $\phi \in \mathcal{D}(\mathbb{R}^n)$, consider $(T * \delta_0)(\phi) = (T * (\delta_0 * \check{\phi}))(0)$. Set $\psi := \delta_0 * \check{\phi}$. Then $(T * \delta_0)(\phi) = T(\check{\psi})$. Note that

$$\check{\psi}(x) = \psi(-x) = (\delta_0 * \check{\phi})(-x) = \delta_0(\tau_{-x}\phi) = \tau_{-x}\phi(0) = \phi(x).$$

Hence, $(T * \delta_0)(\phi) = T(\phi)$. □

Theorem 1.5.29. *$\mathcal{D}(\Omega)$ is dense in $\mathcal{D}'(\Omega)$.*

The following exercise is same as Exercise 28. However, using convolution one can give a simpler proof to Exercise 28.

Exercise 44. Let $T \in \mathcal{D}'(\mathbb{R})$ be such that $DT = 0$, then show that T is a regular distribution generated by a constant function.

Proof. For a given sequence of mollifiers, let $T_\varepsilon(x) = (T * \rho_\varepsilon)(x)$. Thus, $T_\varepsilon \in \mathcal{E}(\mathbb{R})$ and $DT_\varepsilon = DT * \rho_\varepsilon = 0$. Hence, $T_\varepsilon = \lambda_\varepsilon$, where λ_ε is a constant function for each ε. Also, $\lambda_\varepsilon \rightharpoonup T$ in $\mathcal{D}'(\mathbb{R})$. In particular, choose $\phi \in \mathcal{D}(\mathbb{R})$ such that $\int \phi = 1$, then the sequence of real numbers λ_ε converges to some λ. Thus, $T = T_\lambda$. □

Exercise 45. Let T denote the distribution corresponding to the constant function 1, $\delta_0^{(1)}$ be the dipole distribution at 0 and H be the Heaviside function

$$H(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0. \end{cases}$$

Show that $(T * \delta_0^{(1)}) * H \neq T * (\delta_0^{(1)} * H)$.

Proof. Consider

$$(1 * \delta_0^{(1)})(\phi) = 1 * (\delta_0^{(1)} * \check{\phi})(0) = (1 * \psi)(0) = T_1(\check{\psi}) = \int_{\mathbb{R}} \check{\psi}(x)\, dx.$$

But $\check{\psi}(x) = \delta_0^{(1)} * \check{\phi}(-x) = \delta_0^{(1)}(\tau_{-x}\phi) = -\tau_{-x}\phi'(0) = -\phi'(x)$. Hence $(1 * \delta_0^{(1)})(\phi) = -\int_{\mathbb{R}} \phi'(x)\, dx = 0$. Further, $0 * H = 0$. Now, consider

$$(\delta_0^{(1)} * H)(\phi) = \delta_0^{(1)} * (H * \check{\phi})(0) = (\delta_0^{(1)} * \psi)(0) = \delta_0^{(1)}(\check{\psi}) = (\check{\psi})'(0) = (\check{\psi})'(0).$$

But $(\check{\psi})'(0) = (H' * \check{\phi})(0) = (\delta_0 * \check{\phi})(0) = \delta_0(\phi) = \phi(0)$. Since δ_0 is the identity distribution under convolution operation, $1 * \delta_0 = 1$.

Alternately, notice that by derivative of convolution $(1 * \delta_0^{(1)}) * H = (D1 * \delta_0) * H = D1 * H = 0 * H = 0$. On the other hand, $1 * (\delta_0^{(1)} * H) = 1 * (\delta_0 * DH) = 1 * (\delta_0 * \delta_0) = 1 * \delta_0 = 1$. □

1.6 Tempered Distributions

Recall the definition of Schwartz space from Definition **??**. By definition, $\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$. The inclusion is also dense because, for any $f \in \mathcal{S}(\mathbb{R}^n)$, $\phi_m f \to f$ in $\mathcal{S}$ where $\phi_m(x) = \phi(x/m)$ and $\phi \in \mathcal{D}(\mathbb{R}^n)$ is such that $\phi \equiv 1$ on the unit ball in $\mathbb{R}^n$. Hence the dual of $\mathcal{S}$, denoted $\mathcal{S}'$, can be identified with a subspace of $\mathcal{D}'(\mathbb{R}^n)$.

Definition 1.6.1. *The space of distributions $\mathcal{S}'$ is called the space of* tempered distributions.

Example 1.37. Any distribution with compact support is tempered, i.e., $\mathcal{E}' \subset \mathcal{S}'$ because $\mathcal{D} \subset \mathcal{S} \subset \mathcal{E}$ and all inclusions are dense.

Example 1.38. Any slowly increasing measure induces a tempered distribution. A measure μ on $\mathbb{R}^n$ is said to be *slowly increasing* if, for some integer $k \geq 0$,

$$\int_{\mathbb{R}^n} \frac{d\mu(x)}{(1+|x|^2)^k} < +\infty.$$

For instance, any bounded measure is slowly increasing. Set

$$T_\mu(f) = \int_{\mathbb{R}^n} f(x)\, d\mu(x), \quad f \in \mathcal{S}(\mathbb{R}^n).$$

Then T_μ is a linear functional on $\mathcal{S}(\mathbb{R}^n)$ and

$$|T_\mu(f)| \leq \left(\sup_{x \in \mathbb{R}^n} |f(x)(1+|x|^2)^k| \right) \int_{\mathbb{R}^n} \frac{d\mu(x)}{(1+|x|^2)^k}.$$

Further, it follows that if $f_m \to 0$ in $\mathcal{S}(\mathbb{R}^n)$ then $T_\mu(f_m) \to 0$. Thus, $T_\mu \in \mathcal{S}'(\mathbb{R}^n)$.

Example 1.39. If $1 \leq p \leq \infty$ then $L^p(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$. If $f \in L^p(\mathbb{R}^n)$, define

$$T_f(\phi) = \int_{\mathbb{R}^n} f\phi\, dx \quad \forall \phi \in \mathcal{S}(\mathbb{R}^n).$$

If $1 < p < \infty$, let q be the conjugate exponent of p given as $(1/p) + (1/q) = 1$. Then for $k > \frac{n}{2q}$, $g(x) = (1+|x|^2)^{-k} \in L^q(\mathbb{R}^n)$. Hence, by Hölder's inequality,

$$|T_f(\phi)| \leq \left(\sup_{x \in \mathbb{R}^n} |(1+|x|^2)^k \phi(x)| \right) \|f\|_{L^p(\mathbb{R}^n)} \|g\|_{L^q(\mathbb{R}^n)}.$$

Thus, T_f is tempered, and $f \mapsto T_f$ is continuous on $L^p(\mathbb{R}^n)$. If $p = 1$,

$$|T_f(\phi)| \leq \|\phi\|_{\infty,\mathbb{R}^n}\|f\|_{1,\mathbb{R}^n}$$

and if $p = \infty$,

$$|T_f(\phi)| < \sup_{x\in\mathbb{R}^n} \left|(1+|x|^2)^k\phi(x)\right| \|f\|_{\infty\,\mathbb{R}^n}\|g\|_{1,(\mathbb{R}^n)}$$

where g is as above with $k > \frac{n}{2}$.

Definition 1.6.2. *Let $T \in \mathcal{S}'(\mathbb{R}^n)$. The* Fourier transform *of T, denoted as $\hat{T}$, is defined as $\hat{T}(f) = T(\hat{f})$ for all $f \in \mathcal{S}(\mathbb{R}^n)$.*

Since $f \mapsto \hat{f}$ is continuous on $\mathcal{S}$, it follows that $\hat{T} \in \mathcal{S}'$.

Remark 1.6.3. Since $\mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$. It seems that, a priori, there are two definitions of the Fourier transform on $\mathcal{S}(\mathbb{R}^n)$, each inherited from $L^1(\mathbb{R}^n)$ and $\mathcal{S}'(\mathbb{R}^n)$, respectively. But they are one and the same because if $f \in \mathcal{S}(\mathbb{R}^n)$ then, for any $g \in \mathcal{S}(\mathbb{R}^n)$, by Parseval relation,

$$\hat{T}_f(g) = \int_{\mathbb{R}^n} f\hat{g} = \int_{\mathbb{R}^n} \hat{f}g = T_{\hat{f}}(g).$$

Hence $\hat{T}_f = T_{\hat{f}}$ and, hence, both definitions of the Fourier transform coincide on $\mathcal{S}(\mathbb{R}^n)$.

Recall that $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$, defined as $\mathcal{F}(f) = \hat{f}$, is an isomorphism on $\mathcal{S}(\mathbb{R}^n)$, it extends as a bijection to $\mathcal{S}'(\mathbb{R}^n)$, as well.

Remark 1.6.4. Since $\mathcal{S}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$, then there are two definitions of Fourier transform on $L^2(\mathbb{R}^n)$ inherited from $\mathcal{S}'(\mathbb{R}^n)$ and the other by extending from $\mathcal{S}(\mathbb{R}^n)$. It turns out that both are same. Consider the weak-* convergence of tempered distributions, i.e., $T_m \to T$ weak-* sense if, for every $\phi \in \mathcal{S}(\mathbb{R}^n)$, $T_m(\phi) \to T(\phi)$. If $T_m \to T$ in $\mathcal{S}'(\mathbb{R}^n)$ then $\hat{T_m} \to \hat{T}$ in $\mathcal{S}'$ because

$$\hat{T}_m(\phi) = T_m(\hat{\phi}) \to T(\hat{\phi}) = \hat{T}(\phi).$$

Let $f \in L^2(\mathbb{R}^n)$ and let $f_k \in \mathcal{S}(\mathbb{R}^n)$ such that $f_k \to f$ in $L^2(\mathbb{R}^n)$. Since the inclusion of $L^2(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$ is continuous, $f_k \to f$ in $\mathcal{S}'(\mathbb{R}^n)$, as well. Hence $\hat{f}_k \to \hat{f}$ in $\mathcal{S}'(\mathbb{R}^n)$. On the other hand, $\mathcal{F}(f_k) \to \mathcal{F}(f)$ in $L^2(\mathbb{R}^n)$ and hence in $\mathcal{S}'(\mathbb{R}^n)$. But as $f_k \in \mathcal{S}(\mathbb{R}^n)$, we know that $\mathcal{F}(f_k) = \hat{f}_k$. Now, by uniqueness of the weak-* limit $\hat{f} = \mathcal{F}(f)$.

By Proposition **??**, for any polynomial P and mulit-index α, if $T \in \mathcal{S}'(\mathbb{R}^n)$ then $PT \in \mathcal{S}'(\mathbb{R}^n)$ and $D^\alpha T \in \mathcal{S}'(\mathbb{R}^n)$.

We now extend the result of Theorem **??** to tempered distribution.

Theorem 1.6.5. *Let $T \in \mathcal{S}'(\mathbb{R}^n)$ and let α be a multi-index. Then*

$$D^\alpha \hat{T} = (-\imath)^{|\alpha|}\widehat{(x^\alpha T)}$$

and

$$\widehat{(D^\alpha T)} = (\imath)^{|\alpha|}\xi^\alpha \hat{T}.$$

Proof. Let $f \in \mathcal{S}$. Then

$$\begin{aligned}
\widehat{(x^\alpha T)}(f) &= (x^\alpha T)(\hat{f}) = T(x^\alpha \hat{f}(x)) \\
&= \frac{1}{(\imath)^{|\alpha|}} T(\widehat{D^\alpha f}) = \frac{1}{(\imath)^{|\alpha|}}\hat{T}(D^\alpha f) \\
&= \frac{1}{(-\imath)^{|\alpha|}} D^\alpha \hat{T}(f).
\end{aligned}$$

Thus, the first relation is proved. The second relation can be proved similarly. $\square$

Example 1.40. $\hat{\delta}_0 = 1$, $\widehat{\left(\frac{\partial \delta_0}{\partial x_k}\right)}(\xi) = \imath \xi_k$ and $\hat{1} = \delta_0$. Let $\phi \in \mathcal{S}(\mathbb{R}^n)$. Then

$$\hat{\delta_0}(\phi) = \delta_0(\hat{\phi}) = \hat{\phi}(0) = \int_{\mathbb{R}^n} \phi(x)\, dx.$$

Thus, $\hat{\delta_0} = 1$, i.e., the distribution induced by the constant function 1. Further,

$$\widehat{\left(\frac{\partial \delta_0}{\partial x_k}\right)}(\xi) = \imath \xi_k \hat{\delta_0} = \imath \xi_k.$$

The $\hat{1} = \delta_0$ follows by the Fourier inversion formula.

1.7 Fundamental Solution

Let $L = \sum_{|\alpha| \leq n} c_\alpha D^\alpha$ be a n-th order linear differential operator with constant coefficients. Thus, any differential equation is of the form $LT = S$, in the distribution sense for any given distribution S and T is the distribution solution of the differential equation.

Definition 1.7.1. *A distribution K is said to be a* fundamental solution *of the operator L if $L(K) = \delta_0$, where δ_0 is the Dirac distribution at $\{0\}$.*

The existence of fundamental solution of a linear differential operator for constant coefficients is the famous result of Malgrange and Ehrenpreis (cf. [Rud91]).

Suppose we wish to solve $LT = S$ for a given distribution S, we first find the fundamental solution $LK = \delta_0$. Then $K * S$ (as long as the convolution makes sense) is a solution of the equation $L(\cdot) = S$, because $L(K * S) = LK * S = \delta_0 * S = S$. The fundamental solution is not unique because any solution U of the homogeneous equation $L(U) = 0$ can be added to the fundamental solution K to obtain other fundamental solutions, since $L(U + K) = L(U) + L(K) = \delta_0$.

Definition 1.7.2. *A differential operator $L = \sum_{|\alpha| \leq n} c_\alpha D^\alpha$ with $c_\alpha \in C^\infty$ is said to be* hypoelliptic *whenever $Lu \in C^\infty$ implies that u is C^∞.*

Theorem 1.7.3. *For a differential operator with constant coefficients, the following are equivalent:*

(i) L is hypoelliptic.

(ii) Every fundamental solution of L is C^∞ on $\mathbb{R}^n \setminus \{0\}$.

(iii) Some fundamental solution of L is C^∞ on $\mathbb{R}^n \setminus \{0\}$.

1.8 Further Reading

Most of the concepts discussed in this chapter can be found in detail in [Kes89, Rud91, Hor70, Tre70].

Chapter 2

Sobolev Spaces

In this chapter we develop a class of functional spaces which form the right setting to study partial differential equations. Recall that $\Omega \subset \mathbb{R}^n$ is open and not necessarily bounded. The situations where Ω needs to be bounded will be specified clearly. Before we describe Sobolev spaces in detail, we introduce a classical functional space very useful in the theory of Sobolev spaces.

2.1 Hölder Spaces

In this section, we introduce some class of function spaces in $C^k(\Omega)$, for all integers $k \geq 0$, which can viewed in some sense as spaces of "fractional" derivatives. For any $\gamma \in (0, \infty)$ and $x_0 \in \Omega$, a function $u : \Omega \to \mathbb{R}$ satisfies at x_0, for all $x \in \Omega$, the estimate

$$|u(x) - u(x_0)| = \frac{|u(x) - u(x_0)|}{|x - x_0|^\gamma}|x - x_0|^\gamma \leq \sup_{\substack{x \in \Omega \\ x \neq x_0}} \left\{ \frac{|u(x) - u(x_0)|}{|x - x_0|^\gamma} \right\} |x - x_0|^\gamma.$$

However, it is not always necessary that the supremum is finite. Note that the modulus in the numerator and denominator are in $\mathbb{R}$ and $\mathbb{R}^n$, respectively.

Definition 2.1.1. *Let $\gamma \in (0, 1]$. We say a function $u : \Omega \to \mathbb{R}$ is* Hölder continuous *of exponent γ at $x_0 \in \Omega$, if*

$$p_\gamma(u)(x_0) := \sup_{\substack{x \in \Omega \\ x \neq x_0}} \left\{ \frac{|u(x) - u(x_0)|}{|x - x_0|^\gamma} \right\} < +\infty.$$

Note that any Hölder continuous function at $x_0 \in \Omega$ satisfies the estimate

$$|u(x) - u(x_0)| \leq p_\gamma(u)(x_0)|x - x_0|^\gamma, \quad \text{for } x \in \Omega. \tag{2.1.1}$$

The constant $p_\gamma(u)(x_0)$ may depend on u, Ω, γ and x_0. It follows from (2.1.1) that any Hölder continuous function at x_0, for any exponent γ, is also continuous at x_0. Because, for every $\varepsilon > 0$, we can choose $\delta = \left[\frac{\varepsilon}{p_\gamma(u)(x_0)}\right]^{1/\gamma}$. The case when $\gamma = 1$ corresponds to the Lipshitz continuity of u at x_0.

Definition 2.1.2. *Let $\gamma \in (0, 1]$. We say a function $u : \Omega \to \mathbb{R}$ is* uniformly Hölder continuous *of exponent γ, if*

$$p_\gamma(u) := \sup_{\substack{x,y\in\Omega,\\ x\neq y}} \left\{ \frac{|u(x) - u(y)|}{|x - y|^\gamma} \right\} < +\infty \tag{2.1.2}$$

and is denoted by $C^{0,\gamma}(\Omega)$. The quantity $p_\gamma(u)$ is called the γ-th Hölder coefficient *of u. If the Hölder coefficient is finite on every compact subsets of Ω, then u is said to be* locally *Hölder continuous with exponent γ, denoted as $C^{0,\gamma}_{loc}(\Omega)$.*

Note that $p_\gamma(u) = \sup_{x_0\in\Omega} p_\gamma(u)(x_0)$. In other words, uniformly Hölder continuous function do not include all those that are Hölder continuous at all points of Ω. However, bounded locally Hölder continuous functions are Hölder continuous at all points of Ω.

Example 2.1. Note that, for all $\delta < 0$, $u(x) = |x|^\delta$ on the open ball $B_1(0) \subset \mathbb{R}^n$ do not belong to $C(B_1(0))$. For $\delta \in [0, \infty)$, $|x|^\delta \in C(B_1(0))$ and, for $\delta \in [2, \infty) \cup \{0\}$, $|x|^\delta \in C^1(B_1(0))$. For $\delta \in (0, 1]$, $|x|^\delta$ is Hölder continuous, with exponent γ, for all $0 < \gamma \leq \delta$, but is not Hölder continuous for $\gamma > \delta$. In particular, $|x|^\delta \in C^{0,\delta}(B_1(0))$ for each $\delta \in [0, 1]$. Thus, we have one example from each of the space $C^{0,\delta}(B_1(0))$. We first check the Hölder continuity at $x_0 = 0$. The γ-th Hölder coefficient is

$$\sup_{x\in B_1(0)} \frac{|x|^\delta}{|x|^\gamma} = |x|^{\delta-\gamma} \leq 1 \quad \text{for } \gamma \leq \delta.$$

For $\gamma > \delta$ the γ-th Hölder coefficient blows up. More generally, for $x \neq 0$ and $|x| > |y|$ (wlog),

$$\frac{\left||x|^\delta - |y|^\delta\right|}{|x - y|^\gamma} = \frac{|x|^{\delta-\gamma}\left|1 - \left(\frac{|y|}{|x|}\right)^\delta\right|}{\left|\frac{x}{|x|} - \frac{y}{|x|}\right|^\gamma}.$$

Since $0 < 1 - \frac{|y|}{|x|} < 1$ and $0 < \delta \leq 1$, $\left|1 - \left(\frac{|y|}{|x|}\right)^\delta\right| \leq \left|1 - \frac{|y|}{|x|}\right|$. Thus,

$$\frac{|x|^{\delta-\gamma}\left|1 - \left(\frac{|y|}{|x|}\right)^\delta\right|}{\left|\frac{x}{|x|} - \frac{y}{|x|}\right|^\gamma} \leq \frac{|x|^{\delta-\gamma}\left|1 - \frac{|y|}{|x|}\right|}{\left|\frac{x}{|x|} - \frac{y}{|x|}\right|^\gamma} \leq \frac{||x| - |y||}{|x|^{1-\gamma}|x-y|^\gamma}.$$

The last inequality is true for $\gamma \leq \delta$ and we have used $|x|^{\delta-\gamma} \leq 1$. For one dimension, the last quantity is equal to 1 because

$$\frac{||x| - |y||}{|x|^{1-\gamma}|x-y|^\gamma} = \frac{|x|\left|1 - \frac{|y|}{|x|}\right|}{|x||1 - \frac{y}{x}|^\gamma} \leq \frac{\left|1 - \frac{y}{x}\right|}{\left|1 - \frac{y}{x}\right|} = 1.$$

For higher dimensions, INCOMPLETE!!!

Example 2.2. The Cantor function $f_C \in C^{0,\gamma}([0,1])$, i.e., is Hölder continuous with exponent $0 < \gamma \leq \log_3 2$.[1] Geometrically, the graph of Hölder continuous function have fractal appearance which increases with smaller γ.

The case $\gamma = 0$ corresponds to bounded functions on Ω and hence is ignored as an possible exponent. In fact, *continuous* functions in $C^{0,0}(\Omega)$ can be identified with the space $C_b(\Omega)$ of bounded continuous functions on Ω. The space $C(\Omega)$ can be identified with continuous functions in $C^{0,0}_{\text{loc}}(\Omega)$, the space of all locally Hölder continuous with exponent $\gamma = 0$. The case $\gamma > 1$ is also ignored because only constant functions, on each connected component of Ω, can satisfy Hölder estimate with exponent greater than one.

Exercise 46. If $u \in C^{0,\gamma}(\Omega)$ with exponent $\gamma > 1$, then u is constant in each of the connected component of Ω.

Proof. Let $u \in C^{0,\gamma}(\Omega)$. For any $x \in \Omega$,

$$|u_{x_i}(x)| = \lim_{t\to 0}\frac{1}{|t|}|u(x+te_i)-u(x)| \leq \lim_{t\to 0}\frac{p_\gamma(u)(x)}{|t|}|t|^\gamma = \lim_{t\to 0}p_\gamma(u)(x)|t|^{\gamma-1} = 0.$$

Therefore, the gradient $Du(x) = 0$ for all $x \in \Omega$ and hence u is constant in each connected components of Ω. □

[1] $\log_3 2 \approx 0.6309\ldots$

The case $\gamma = 1$ corresponds to u being *Lipschitz* continuous. The space of all Lipschitz functions is denoted as $\mathrm{Lip}(\Omega) = C^{0,1}(\Omega)$.

Exercise 47. If Ω has finite diameter, then for any $0 < \gamma < \delta \leq 1$ we have $C^{0,1}(\Omega) \subseteq C^{0,\delta}(\Omega) \subsetneq C^{0,\gamma}(\Omega) \subsetneq C(\Omega)$.

Proof. The last inclusion is trivial because every $u \in C^{0,\gamma}(\Omega)$ is uniformly continuous by choosing $\delta = \left[\frac{\varepsilon}{p_\gamma(u)}\right]^{1/\gamma}$, for any given $\varepsilon > 0$. Thus, $C^{0,\gamma}(\Omega) \subsetneq C(\Omega)$ for all $0 < \gamma \leq 1$. On the other hand, let $u \in C^{0,\delta}(\Omega)$. Consider

$$\begin{aligned} p_\gamma(u) &= \sup_{\substack{x,y\in\Omega \\ x\neq y}} \left\{ \frac{|u(x) - u(y)|}{|x-y|^\gamma} \right\} \\ &= \sup_{\substack{x,y\in\Omega \\ x\neq y}} \left\{ \frac{|u(x) - u(y)|}{|x-y|^\delta} |x-y|^{\delta-\gamma} \right\} \\ &\leq \mathrm{diam}(\Omega)^{\delta-\gamma} p_\delta(u). \end{aligned}$$

Thus, $p_\gamma(u) < +\infty$ and $u \in C^{0,\gamma}(\Omega)$. The inclusions are strict by Example 2.1. If $\gamma < \delta$, then $u(x) = |x|^\gamma$ is in $C^{0,\gamma}(-1,1)$ but not in $C^{0,\delta}(-1,1)$. In particular, $\sqrt{x} \in C^{0,1/2}(-1,1)$ but is not Lipschitz, i.e. do not belong to $C^{0,1}(-1,1)$. □

A natural question at this juncture is: what is the relation between the spaces $C^1(\Omega)$ and $C^{0,1}(\Omega)$. Of course $|x| \in C^{0,1}(-1,1)$ and not in $C(-1,1)$. Given the inclusion relation in above exercise, we would tend to believe that $C^1(\Omega)$ is contained in $C^{0,1}(\Omega)$, but this is not true. In fact, only a subclass of $C^1(\Omega)$ belong to Lipschitz class.

Theorem 2.1.3. *Let Ω be a convex domain. If $u \in C^1(\Omega)$ and $D^\alpha u$ is bounded on Ω, for each $|\alpha| = 1$, then $u \in C^{0,1}(\Omega)$.*

Proof. Let $u \in C^1(\Omega)$ such that there is a $C_0 > 0$ such that $|\nabla u(x)| \leq C_0$ for all $x \in \Omega$. For any two given points $x, y \in \Omega$, we define the function $F : [0,1] \to \mathbb{R}$ as $F(t) = u((1-t)x + ty)$. F is well defined due to the convexity of Ω. Since u is differentiable, F is differentiable in $(0,1)$. Thus, for all $t \in (0,1)$, $F'(t) = \nabla u[(1-t)x + ty] \cdot (y-x)$. Therefore, $|F'(t)| \leq C_0|y-x|$. By mean value theorem, there is a $\xi \in (0,1)$ such that $F'(\xi) = F(1) - F(0)$. Thus,

$$|u(y) - u(x)| = |F(1) - F(0)| = |F'(\xi)| \leq C_0|y-x|.$$

Hence, u is Lipschitz continuous. $\square$

What fails in the converse of the above theorem is the fact that a Lipschitz function can fail to be differentiable. If we assume, in addition to Lipschitz continuity, that u is differentiable then $u \in C^1$ and derivative is bounded.

Exercise 48. Prove the above theorem when Ω is path connected.

Theorem 2.1.4. *For any $\gamma \in (0,1]$, the space $C^{0,\gamma}(\Omega)$ can be identified with the space $C^{0,\gamma}(\overline{\Omega})$.*

Proof. The space $C^{0,\gamma}(\overline{\Omega})$ can be identified with a subset of $C^{0,\gamma}(\Omega)$, by identifying any $u \in C^{0,\gamma}(\overline{\Omega})$ with $u\mid_\Omega \in C^{0,\gamma}(\Omega)$. On the other hand, for any $u \in C^{0,\gamma}(\Omega)$ we wish to extend it uniquely to a function $\tilde{u} \in C^{0,\gamma}(\overline{\Omega})$ such that $\tilde{u}\mid_\Omega = u$. Let $u \in C^{0,\gamma}(\Omega)$. Set $\tilde{u}(x) = u(x)$ for all $x \in \Omega$. Now, for any $x_0 \in \partial\Omega$, choose a sequence $\{x_m\} \subset \Omega$ such that $\lim_m x_m = x_0$. Then, we have
$$|u(x_k) - u(x_l)| \leq C|x_k - x_l|^\gamma \to 0 \text{ as } k, l \to \infty.$$
Thus, $\{u(x_m)\}$ is a Cauchy sequence and converges in $\mathbb{R}$. Set $\tilde{u}(x_0) = \lim_m u(x_m)$. We now show that the definition of $\tilde{u}(x_0)$ is independent of the choice of the sequence. If $\{y_m\} \subset \Omega$ is any other sequence converging to x_0, then
$$|u(x_m) - u(y_m)| \leq C|x_m - y_m|^\gamma \to 0 \text{ as } m \to \infty.$$
Thus, $\tilde{u}(x_0)$ is well-defined for all $x_0 \in \partial\Omega$. Moreover, for all $x, y \in \overline{\Omega}$, we have

$$|\tilde{u}(x) - \tilde{u}(y)| = \lim_m |u(x_m - u(y_m)| \leq C \lim_m |x_m - y_m|^\gamma = C|x-y|^\gamma.$$

Hence, $\tilde{u}$ satisfies the Hölder estimate with exponent γ and is in $C^{0,\gamma}(\overline{\Omega})$. $\square$

The above result is very special to uniformly continuous spaces and, hence, Hölder spaces. For instance, a similar result is not true between $C(\Omega)$ and $C(\overline{\Omega})$. If $\overline{\Omega}$ is compact then every function in $C(\overline{\Omega})$ is bounded, where as $C(\Omega)$ might have unbounded functions. An interesting point in the proof of above result is that the Hölder coefficient are same for the extension, i.e., $p_\gamma(u) = p_\gamma(\tilde{u})$. An advantage of above theorem is that, without loss of generality, we can use the space $C^{0,\gamma}(\Omega)$ and $C^{0,\gamma}(\overline{\Omega})$ interchangeably.

Exercise 49. Show that the Hölder coefficient, p_γ, defines a semi-norm on $C^{0,\gamma}(\Omega)$.

Proof. For any constant function u, $p_\gamma(u) = 0$. Also,

$$\begin{aligned} p_\gamma(u+v) &= \sup_{\substack{x,y\in\overline{\Omega}\\ x\neq y}} \left\{ \frac{|u(x)+v(x)-u(y)-v(y)|}{|x-y|^\gamma} \right\} \\ &= \sup_{\substack{x,y\in\overline{\Omega}\\ x\neq y}} \left\{ \frac{|u(x)-u(y)+v(x)-v(y)|}{|x-y|^\gamma} \right\} \\ &\leq \sup_{\substack{x,y\in\overline{\Omega}\\ x\neq y}} \left\{ \frac{|u(x)-u(y)|+|v(x)-v(y)|}{|x-y|^\gamma} \right\} \\ &\leq p_\gamma(u) + p_\gamma(v). \end{aligned}$$

Further $p_\gamma(\lambda u) = |\lambda| p_\gamma(u)$. □

Note that for bounded open subsets Ω, the space $C^{0,\gamma}(\Omega)$ inherits the uniform topology from $C(\Omega)$.

Exercise 50. For any bounded open set $\Omega \subset \mathbb{R}^n$, show that $C^{0,\gamma}(\Omega)$ is dense in $C(\Omega)$.

To make the space $C^{0,\gamma}(\Omega)$ complete on bounded Ω, we define the γ-Hölder norm on $C^{0,\gamma}(\overline{\Omega})$ as

$$\|u\|_{C^{0,\gamma}(\Omega)} := \|u\|_\infty + p_\gamma(u),$$

where $\|u\|_\infty := \sup_{x\in\overline{\Omega}} |u(x)|$.

Exercise 51. For any bounded open set $\Omega \subset \mathbb{R}^n$, show that $\|\cdot\|_{C^{0,\gamma}(\Omega)}$ is a norm on $C^{0,\gamma}(\Omega)$.

Theorem 2.1.5. *For any bounded open set $\Omega \subset \mathbb{R}^n$, the space $C^{0,\gamma}(\Omega)$ is a Banach space with norm $\|\cdot\|_{C^{0,\gamma}(\Omega)}$.*

Proof. We need to prove the completeness of the space $C^{0,\gamma}(\Omega)$ w.r.t the norm $\|\cdot\|_{C^{0,\gamma}(\Omega)}$. Let $\{u_m\}$ be a Cauchy sequence in $C^{0,\gamma}(\Omega)$, then $\{u_m\} \subset C(\Omega)$ is Cauchy w.r.t the supremum norm. Thus, there is a $u \in C(\Omega)$ such that $\|u_m - u\|_\infty \to 0$, as $m \to \infty$. We first show that $u \in C^{0,\gamma}(\Omega)$. For $x, y \in \Omega$ with $x \neq y$, consider

$$\frac{|u(x)-u(y)|}{|x-y|^\gamma} = \lim_m \frac{|u_m(x)-u_m(y)|}{|x-y|^\gamma} \leq \limsup_m p_\gamma(u_m) \leq \lim_m \|u_m\|_{C^{0,\gamma}(\Omega)}.$$

Since $\{u_m\}$ is Cauchy, $\lim_m \|u_m\|_{C^{0,\gamma}(\Omega)} < \infty$. Hence $u \in C^{0,\gamma}(\Omega)$. Finally, we show that the sequence $\{u_m\}$ converges to u in $C^{0,\gamma}(\Omega)$ w.r.t the γ-Hölder norm. Consider,

$$\begin{aligned}\frac{|u_m(x) - u(x) - u_m(y) + u(y)|}{|x-y|^\gamma} &= \lim_k \frac{|u_m(x) - u_k(x) - u_m(y) + u_k(y)|}{|x-y|^\gamma} \\ &\leq \limsup_k p_\gamma(u_m - u_k).\end{aligned}$$

Therefore, $\lim_m p_\gamma(u_m - u) \leq \lim_m \limsup_k p_\gamma(u_m - u_k)$ and the RHS converges to 0 since the sequence is Cauchy. Hence, $\|u_m - u\|_{C^{0,\gamma}(\Omega)} \to 0$. □

Exercise 52. The space $C^{0,\gamma}(\Omega)$ norm is not separable.[2]

For an unbounded open set Ω, we consider a sequence of exhaustion compact sets, as described in §1.3.4, and make the space $C^{0,\gamma}(\Omega)$ a Fréchet space.

Definition 2.1.6. *Let X and Y be Banach spaces. A continuous (bounded) linear operator $L : X \to Y$ is said to be* compact *if L maps every bounded subset of X to precompact (closure compact) subsets of Y. Equivalently, L maps bounded sequences of X to sequences in Y that admit convergent subsequences.*

Definition 2.1.7. *Let X and Y be Banach spaces such that $X \subset Y$. We say X is* continuously imbedded *in Y (denoted as $X \hookrightarrow Y$), if there is a constant such that*

$$\|x\|_Y \leq C\|x\|_X \quad \forall x \in X.$$

Further, we say X is compactly imbedded *in Y (denoted as $X \subset\subset Y$) if in addition to being continuously imbedded in Y, every bounded set in X (w.r.t the norm in X) is compact in Y (w.r.t the norm in Y).*

Theorem 2.1.8. *Let Ω be a bounded open subset of $\mathbb{R}^n$. For any $0 < \gamma < \delta \leq 1$, the inclusion map $I : C^{0,\delta}(\Omega) \to C^{0,\gamma}(\Omega)$ is continuous and compact. Further, the inclusion map $I : C^{0,\gamma}(\Omega) \to C(\Omega)$ is compact, for all $0 < \gamma \leq 1$.*

Proof. The continuity of I follows from the inequality proved in Exercise 47 and hence I is a bounded linear operator. We need to show I is compact. Let $\{u_m\}$ be a sequence bounded in $C^{0,\delta}(\Omega)$. Without loss of generality, we

[2] The uniform norm in $C(\Omega)$ is separable

can assume that $\|u_m\|_{C^{0,\delta}(\Omega)} \leq 1$. Therefore, $p_\delta(u_m) \leq 1$ for all m which implies that $\{u_m\}$ is an equicontinuous sequence in $C(\Omega)$. By Arzelà-Ascoli theorem, there exists a subsequence $\{u_{m_k}\}$ of $\{u_m\}$ and a $u \in C(\Omega)$ such that $\|u_{m_k} - u\|_\infty \to 0$, as $k \to \infty$. The $u \in C(\Omega)$ is, in fact, in $C^{0,\delta}(\Omega) \subset C^{0,\gamma}(\Omega)$ because

$$\frac{|u(x) - u(y)|}{|x - y|^\delta} = \lim_k \frac{|u_{m_k}(x) - u_{m_k}(y)|}{|x - y|^\delta} \leq 1$$

and, hence, $p_\delta(u) \leq 1$. We now show that the u_{m_k} converges to u in $C^{0,\gamma}(\Omega)$. For simplicity, set $v_k := u_{m_k} - u$ and we will show that $\{v_k\}$ converges to 0 in $C^{0,\gamma}(\Omega)$. Obviously, $\|v_k\|_\infty \to 0$ thus it is enough to show that $p_\gamma(v_k) \to 0$. Note that, for every given $\varepsilon > 0$, we have $p_\gamma(v_k) \leq S_k^\varepsilon + T_k^\varepsilon$, where

$$S_k^\varepsilon = \sup_{\substack{x,y \in \Omega \\ x \neq y; |x-y| \leq \varepsilon}} \left\{ \frac{|v_k(x) - v_k(y)|}{|x - y|^\gamma} \right\}$$

and

$$T_k^\varepsilon = \sup_{\substack{x,y \in \Omega \\ |x-y| > \varepsilon}} \left\{ \frac{|v_k(x) - v_k(y)|}{|x - y|^\gamma} \right\}.$$

Consider

$$S_k^\varepsilon = \sup_{\substack{x,y \in \Omega \\ x \neq y; |x-y| \leq \varepsilon}} \left\{ \frac{|v_k(x) - v_k(y)|}{|x - y|^\delta} |x - y|^{\delta - \gamma} \right\} \leq \varepsilon^{\delta - \gamma} p_\delta(v_k) \leq 2\varepsilon^{\delta - \gamma}.$$

Similarly,

$$T_k^\varepsilon \leq 2\varepsilon^{-\gamma} \|v_k\|_\infty.$$

Hence, $\limsup_k p_\gamma(v_k) \leq 2\varepsilon^{\delta-\gamma} + 2\varepsilon^{-\gamma} \limsup_k \|v_k\|_\infty = 2\varepsilon^{\delta-\gamma} + 0$. Since ε can be made as small as possible, we have $p_\gamma(v_k) \to 0$. $\square$

We extend the Hölder continuous functions in to all differentiable class of functions. We denote by $C^{k,\gamma}(\Omega)$ the space of all $C^k(\Omega)$, k-times continuously differentiable, functions such that $D^\alpha u$ is Hölder continuous with exponent γ, i.e., $D^\alpha u \in C^{0,\gamma}(\Omega)$ for all $|\alpha| = k$. For a bounded open set Ω, we give the γ-th Hölder norm on $C^{k,\gamma}(\Omega)$ as

$$\|u\|_{C^{k,\gamma}(\Omega)} := \sum_{|\alpha|=0}^{k} \|D^\alpha u\|_\infty + \sum_{|\alpha|=k} p_\gamma(D^\alpha u). \tag{2.1.3}$$

It is enough to consider the Hölder coefficient of only the k-th derivative because this is enough to complete the space.

Exercise 53. For any bounded open set $\Omega \subset \mathbb{R}^n$, show that $\|\cdot\|_{C^{k,\gamma}(\Omega)}$ is a norm on $C^{k,\gamma}(\Omega)$. and the space $C^{k,\gamma}(\Omega)$ is a Banach space with norm $\|\cdot\|_{C^{k,\gamma}(\Omega)}$.

Proof. Let $\{u_m\}$ be a Cauchy sequence in $C^{k,\gamma}(\Omega)$. Then, $\{u_m\} \subset C^k(\Omega)$ is Cauchy in the supremum norm. Thus, there is a $u \in C^k(\Omega)$ such that $\|u_m - u\|_{C^k(\Omega)} \to 0$, as $m \to \infty$. The fact that $u \in C^{k,\gamma}(\Omega)$ and that u is a limit, in the γ-Hölder norm, of the Cauchy sequence is similar to case $k = 0$ proved in Theorem 2.1.5. □

As seen before, for an unbounded open set Ω, we consider a sequence of exhaustion compact sets, as described in §1.3.4, and make the space $C^{k,\gamma}(\Omega)$ a Fréchet space. We end this section with a final result on the inclusion between different order Hölder spaces.

Theorem 2.1.9. *Let Ω be a bounded, convex open subset of $\mathbb{R}^n$. For any $0 < \gamma, \delta \leq 1$ and $k, \ell \in \mathbb{N} \cup \{0\}$ such that $k + \gamma < \ell + \delta$, then the inclusion map $I : C^{\ell,\delta}(\Omega) \to C^{k,\gamma}(\Omega)$ is continuous and compact.*

Proof. If $\gamma = \delta$ and $k = \ell$ then there is nothing to prove. Alternately, without loss of generality we assume $0 < \gamma < \delta \leq 1$ and $k < \ell$. Then, we know from Theorem 2.1.8 that $C^{\ell,\delta}(\Omega) \subset C^{\ell,\gamma}(\Omega)$ and the inclusion is continuous and compact. Let $u \in C^{\ell,\gamma}$. We need to show that $u \in C^{k,\gamma}(\Omega)$. i.e., $p_\gamma(D^k u) < \infty$. Note that it is enough to show that $p_1(D^k u) < \infty$. The fact that $u \in C^{\ell,\gamma}$ implies that $u \in C^\ell(\Omega)$ and $D^{\ell-1}u$ is bounded. Thus, by Theorem 2.1.3, we have $u \in C^{\ell-1,1}(\Omega)$ and hence is in $C^{\ell-1,\gamma}(\Omega)$. By repeating the argument for each derivative of $D^{\ell-i}$ for each $i = 1, 2, \ldots, k$, we have the required result. The inclusion $C^{\ell,\gamma}(\Omega) \subset C^{k,\gamma}(\Omega)$ is continuous. The composition of a continuous and compact operator is compact. Thus, the inclusion I is compact. □

2.2 $W^{k,p}(\Omega)$ Spaces

For each $1 \leq p \leq \infty$, we define its *conjugate* exponent q to be,

$$q = \begin{cases} \frac{p}{p-1} & \text{if } 1 < p < +\infty \\ +\infty & \text{if } p = 1 \\ 1 & \text{if } p = +\infty. \end{cases}$$

Note that for, $1 < p < +\infty$, q is the number for which $1/p + 1/q = 1$. In this section, we shall define spaces that are analogous to $C^k(\Omega)$ using the generalised notion of derivative, viz. weak derivative (cf. Definition 1.5.2). We know that every $u \in L^p(\Omega)$ being locally integrable induces a distribution T_u. Further, the distribution T_u is differentiable for all multi-indices α. However, we have already seen that T_u need not be weakly differentiable. For a fixed multi-index α, if there exists a $v_\alpha \in L^p(\Omega)$ such that $T_{v_\alpha} = D^\alpha T_u$, then we denote the v_α as $D^\alpha u$. We know that such a v_α is unique up to a set of measure zero. As usual, Ω is an open subset of $\mathbb{R}^n$.

Definition 2.2.1. *Let $k \geq 0$ be an integer and $1 \leq p \leq \infty$. The* Sobolev space *$W^{k,p}(\Omega)$ is the subclass of all $u \in L^p(\Omega)$ such that there exists a $v_\alpha \in L^p(\Omega)$, for all $0 \leq |\alpha| \leq k$, such that*

$$\int_\Omega v_\alpha \phi \, dx = (-1)^{|\alpha|} \int_\Omega u D^\alpha \phi \, dx \quad \forall \phi \in \mathcal{D}(\Omega).$$

Equivalently,

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

With this convention, we have $W^{0,p}(\Omega) = L^p(\Omega)$. Note that, by definition, if $u \in W^{k,p}(\Omega)$ then every function in the equivalence class of u is also in $W^{k,p}(\Omega)$ and they all have their derivative as v_α upto a set of measure zero.

Example 2.3. Let $\Omega = (-a, a) \subset \mathbb{R}$, for a positive real number a. Recall that the function $u(x) = |x|$ is not in $C^1(-a, a)$. But $u \in W^{1,p}(-a, a)$. Consider the $v \in L^p(-a, a)$ defined as

$$v(x) = \begin{cases} 1 & x \in [0, a) \\ -1 & x \in (-a, 0). \end{cases}$$

We shall show that v is the weak derivative of u. Consider, for $\phi \in \mathcal{D}(-a, a)$,

$$\int_{-a}^{a} v\phi \, dx = -\int_{-a}^{0} \phi \, dx + \int_{0}^{a} \phi \, dx = \int_{-a}^{0} x\phi' \, dx - \int_{0}^{a} x\phi' \, dx = -\int_{-a}^{a} |x|\phi' \, dx.$$

More generally, any continuous function on $[-a, a]$ which is pieceswise differentiable, i.e., piecewise C^1 on $(-a, a)$ is in $W^{1,p}(-a, a)$.

Example 2.4. The function v defined in above example is not in $W^{1,p}(-a,a)$. The argument is similar to the function w defined as

$$w(x) = \begin{cases} 1 & x \in [0,a) \\ 0 & x \in (-a,0). \end{cases}$$

Note that $w \in L^p(-a,a)$, for all p. However, the distributional derivative of w (and v) is the Dirac measure at 0, δ_0, which is a singular distribution. Hence $w \notin W^{k,p}(\Omega)$, for all $k > 0$.

Example 2.5. The function u defined as

$$u(x) = \begin{cases} x & x \in (0,a) \\ 0 & x \in (-a,0] \end{cases}$$

is in $W^{1,p}(-a,a)$, for all $p \in [1,\infty]$, since both u and its distributional derivative $Du = w$ are in $L^p(-a,a)$ (cf. previous example). But $u \notin W^{k,p}(\Omega)$, for all $k \geq 2$.

Note that we have the inclusion $W^{k,p}(\Omega) \subsetneq W^{\ell,p}(\Omega)$ for all $\ell < k$. The inclusion is strict as seen from above examples.

Exercise 54. Find values of $\gamma \in \mathbb{R}$ such that $|x|^\gamma \in W^{1,p}(B_1(0))$ for a fixed $p \in [1,\infty)$ and $B_1(0)$ is the open unit ball in $\mathbb{R}^n$.

Proof. Let $u(x) = |x|^\gamma$. Note that $D^{e_i}u(x) = \gamma x_i|x|^{\gamma-2}$ and $|\nabla u(x)| = \gamma|x|^{\gamma-1}$. Consider

$$\|\nabla u\|_p^p = \gamma^p \int_\Omega |x|^{p(\gamma-1)}\,dx = \gamma^p\omega_n \int_0^1 r^{p\gamma-p+n-1}\,dr.$$

The last quantity is finite iff $p(\gamma-1)+n-1 > -1$, i.e., $\gamma > 1-\frac{n}{p}$ (also $\gamma = 0$, if not already included in the inequality condition) and $|x|^\gamma \in W^{1,p}(\Omega)$. $\square$

Using the above exercise, we have an example of a function in $L^p(B_1(0))$ but not in $W^{1,p}(B_1(0))$. For instance, if $\Omega = B_1(0)$ in $\mathbb{R}^n$ and $u(x) = |x|^\gamma$, for non-zero γ, such that $-\frac{n}{p} < \gamma \leq 1-\frac{n}{p}$, is not in $W^{1,p}(B_1(0))$ but is in $L^p(B_1(0))$.

Exercise 55. Find values of $\gamma \in \mathbb{R}$ such that $|x|^\gamma \in W^{2,p}(B_1(0))$, for a fixed $p \in [1,\infty)$.

Proof. Let $u(x) = |x|^\gamma$. Note that

$$|u_{x_i x_j}| \leq |\gamma(\gamma - 2)||x|^{\gamma-2} + |\gamma||x|^{\gamma-2}.$$

Hence, $u_{x_i x_j} \in L^p(B_1(0))$ if $\gamma - 2 > \frac{-n}{p}$, or $\gamma > 2 - \frac{n}{p}$. Thus $u \in W^{2,p}(B_1(0))$ if $\gamma > 2 - \frac{n}{p}$ and $\gamma = 0$. □

Exercise 56. If $u \in L^p(\Omega)$ is such that $D^\alpha u \in L^p(\Omega)$ for $|\alpha| = k$, then is it true that $D^\alpha u \in L^p(\Omega)$ for all $|\alpha| = 1, 2, \ldots, k-1$?

Exercise 57. Show that the spaces $W^{k,p}(\Omega)$ are all (real) vector spaces.

Exercise 58. Show that $u \mapsto \sum_{|\alpha|=k} \|D^\alpha u\|_p$ defines a semi-norm on $W^{k,p}(\Omega)$ for $k > 0$. We denote the semi-norm by $|u|_{k,p,\Omega}$.

For $1 \leq p < \infty$, we endow the space $W^{k,p}(\Omega)$ with the natural norm,

$$\|u\|_{k,p,\Omega} := \sum_{|\alpha|=0}^{k} \|D^\alpha u\|_p = \sum_{|\alpha|=0}^{k} \left(\int_\Omega |D^\alpha u|^p \right)^{1/p}.$$

For $p = \infty$, we define the norm on $W^{k,\infty}(\Omega)$ to be,

$$\|u\|_{k,\infty,\Omega} = \sum_{|\alpha|=0}^{k} \operatorname{ess\,sup}_\Omega |D^\alpha u|.$$

Observe that $\|u\|_{0,p,\Omega} = |u|_{0,p,\Omega} = \|u\|_p$, the usual L^p-norm.

Exercise 59. Show that the norm $\|u\|_{k,p,\Omega}$ defined above is equivalent to the norm

$$\left(\sum_{|\alpha|=0}^{k} \|D^\alpha u\|_p^p \right)^{1/p} = \left(\sum_{|\alpha|=0}^{k} \int_\Omega |D^\alpha u|^p \right)^{1/p}, \quad \text{for } 1 < p < \infty.$$

For $p = 1$, the norms are same.

Frequently, we may skip the domain subscript in the norm where there is no confusion on the domain of function. Also, we shall tend to use the symbol $\|\cdot\|_p$ for $\|\cdot\|_{0,p}$ in L^p-spaces. We set $H^k(\Omega) = W^{k,2}(\Omega)$. For $u \in H^k(\Omega)$, we shall denote the norm as $\|\cdot\|_{k,\Omega}$.

Exercise 60. Show that

$$\langle u, v\rangle_{k,\Omega} = \sum_{|\alpha|=0}^{k} \int_{\Omega} D^{\alpha} u D^{\alpha} v, \quad \forall u, v \in H^k(\Omega)$$

defines an inner-product in $H^k(\Omega)$.

Theorem 2.2.2. *For every $1 \leq p \leq \infty$, the space $W^{k,p}(\Omega)$ is a Banach space. If $1 < p < \infty$, it is reflexive and if $1 \leq p < \infty$, it is separable. In particular, $H^k(\Omega)$ is a separable Hilbert space.*

Proof. Let $\{u_m\}$ be a Cauchy sequence in $W^{k,p}(\Omega)$. Thus, by the definition of the norm on $W^{k,p}(\Omega)$, $\{u_m\}$ and $\{D^{\alpha}u_m\}$, for $1 \leq |\alpha| \leq k$ are Cauchy in $L^p(\Omega)$. Since $L^p(\Omega)$ is complete, there exist function $u, v_{\alpha} \in L^p(\Omega)$ such that

$$u_m \to u \quad \text{and } D^{\alpha}u_m \to v_{\alpha} \quad \forall 1 \leq |\alpha| \leq k.$$

To show that $W^{k,p}(\Omega)$ is complete, it is enough to show that $D^{\alpha}u = v_{\alpha}$, for $1 \leq |\alpha| \leq k$. Let $\phi \in \mathcal{D}(\Omega)$, then for each α such that $1 \leq |\alpha| \leq k$, we have

$$\begin{aligned}
\int_{\Omega} u D^{\alpha}\phi \, dx &= \lim_{m\to\infty} \int_{\Omega} u_m D^{\alpha}\phi \, dx \\
&= \lim_{m\to\infty} (-1)^{|\alpha|} \int_{\Omega} D^{\alpha}u_m \phi \, dx \\
&= (-1)^{|\alpha|} \int_{\Omega} v_{\alpha}\phi \, dx
\end{aligned}$$

Thus, $D^{\alpha}u = v_{\alpha}$, i.e., v_{α} is the α weak derivative of u.

Note that we have an isometry from $W^{k,p}(\Omega)$ to $(L^p(\Omega))^{\beta}$, where $\beta := \sum_{i=0}^{k} n^i$ and the norm endowed on $(L^p(\Omega))^{\beta}$ is $\|u\| := \left(\sum_{i=1}^{\beta} \|u_i\|_p^p\right)^{1/p}$ with $u = (u_1, u_2, \ldots, u_{\beta})$. The image of $W^{k,p}(\Omega)$ under this isometry is a closed subspace of $(L^p(\Omega))^{\beta}$. The reflexivity and separability of $W^{k,p}(\Omega)$ is inherited from reflexivity and separability of $(L^p(\Omega))^{\beta}$, for $1 < p < \infty$ and $1 \leq p < \infty$, respectively. $\square$

Remark 2.2.3. The proof above uses an important fact which is worth noting. If $u_m \to u$ in $L^p(\Omega)$ and $D^{\alpha}u_m$, for all $|\alpha| = 1$, is bounded in $L^p(\Omega)$ then $u \in W^{1,p}(\Omega)$.

2.3 Smooth Approximations

Recall that in Theorem 1.3.30, we proved that $C_c^\infty(\Omega)$ is dense in $L^p(\Omega)$ for any open subset $\Omega \subseteq \mathbb{R}^n$. We are now interested in knowing if the density of $C_c^\infty(\Omega)$ holds true in the Sobolev space $W^{k,p}(\Omega)$, for any $k \geq 1$[3].

Theorem 2.3.1. *For $1 \leq p < \infty$, $C_c^\infty(\mathbb{R}^n)$ is dense in $W^{k,p}(\mathbb{R}^n)$, for all $k \geq 1$.*

Proof. Let $k \geq 1$ and $u \in W^{k,p}(\mathbb{R}^n)$. The $D^\alpha u \in L^p(\mathbb{R}^n)$ for all $0 \leq |\alpha| \leq k$. By Theorem 1.3.29, we know that $\{\rho_m * u\} \subset C^\infty(\mathbb{R}^n)$ converges to u in the $W^{k,p}(\Omega)$-norm, for all $k \geq 0$. Choose a function $\phi \in C_c^\infty(\mathbb{R}^n)$ such that $\phi \equiv 1$ on $B(0;1)$, $\phi \equiv 0$ on $\mathbb{R}^n \setminus B(0;2)$ and $0 \leq \phi \leq 1$ on $B(0;2) \setminus B(0;1)$. Note that $|D^\alpha \phi|$ is bounded. In fact, $|D^\alpha \phi(x)| \leq 1$ for all $0 \leq |\alpha| \leq k$. Now set $\phi_m = \phi(x/m)$. This is one choice of cut-off functions as introduced in Theorem 1.3.30. Set $u_m := \phi_m(\rho_m * u)$ in $\mathbb{R}^n$, then

$$
\begin{aligned}
\|D^\alpha u_m - D^\alpha u\|_p &\leq \sum_{\beta<\alpha} \left\|C_\alpha (D^{\alpha-\beta}\phi_m)(\rho_m * D^\beta u)\right\|_p \\
&\quad + \sum_{\beta=\alpha} \|\phi_m(\rho_m * D^\alpha u) - D^\alpha u\|_p \\
&\leq \sum_{\beta<\alpha} |C_\alpha| \|D^{\alpha-\beta}\phi_m\|_\infty \|\rho_m * D^\beta u\|_p \\
&\quad + \sum_{\beta=\alpha} \|\phi_m(\rho_m * D^\alpha u - D^\alpha u)\|_p \\
&\quad + \sum_{\beta=\alpha} \|\phi_m D^\alpha u - D^\alpha u\|_p \\
&\leq \frac{M}{m^{|\alpha-\beta|}} \sum_{\beta<\alpha} |C_\alpha| \|D^\beta u\|_p \\
&\quad + \sum_{\beta=\alpha} \|\rho_m * D^\alpha u - D^\alpha u\|_p \\
&\quad + \sum_{\beta=\alpha} \|\phi_m D^\alpha u - D^\alpha u\|_p .
\end{aligned}
$$

The first term with $1/m$ goes to zero, the second term tends to zero by Theorem 1.3.29 and the last term goes to zero by Dominated convergence

[3]The case $k = 0$ is precisely the result of Theorem 1.3.30

theorem. Since

$$\begin{aligned}\|u_m - u\|_{k,p} &= \sum_{|\alpha|=0}^{k} \|D^\alpha u_m - D^\alpha u\|_p \\ &= \sum_{|\alpha|=0}^{k} \left\| \sum_{\beta \le \alpha} C_\alpha (D^{\alpha-\beta}\phi_m)(\rho_m * D^\beta u) - D^\alpha u \right\|_p\end{aligned}$$

and each term in the sum goes to zero as m increases, we have u_m converges to u in $W^{k,p}(\mathbb{R}^n)$ norm. □

The real trouble in approximating $W^{k,p}(\Omega)$ by $C_c^\infty(\Omega)$ functions is that the approximation approach boundary continuously, if Ω has a boundary.

Theorem 2.3.2. *Let $1 \le p < \infty$ and $\Omega \subset \mathbb{R}^n$. Then $C_c^\infty(\mathbb{R}^n)$ is dense in $W^{k,p}_{loc}(\Omega)$, i.e., for each $u \in W^{k,p}(\Omega)$ there is a sequence $v_m \in C_c^\infty(\mathbb{R}^n)$ such that $\|v_m - u\|_{k,p,\omega}$, for all $k \ge 1$ and for all $\omega \subset\subset \Omega$ (relatively compact in Ω).*

Proof. Let $u \in W^{k,p}(\Omega)$. Fix ω relatively compact subset of Ω and a α such that $1 \le |\alpha| \le k$. Let $\tilde{u}$ denote the extension of u by zero in Ω^c. Choose a $\phi \in C_c^\infty(\Omega)$ such that $\phi \equiv 1$ on ω and $0 \le \phi \le 1$ in Ω. Then $\operatorname{supp}(\phi u) \subset \Omega$ and we extend ϕu to all of $\mathbb{R}^n$ by zero in Ω^c and denote it as $\widetilde{\phi u}$. Set $v = \widetilde{\phi u}$ on $\mathbb{R}^n$. Note that $v \in W^{k,p}(\mathbb{R}^n)$. By Theorem 2.3.1, we have the sequence $v_m := \phi_m(\rho_m * v) \in C_c^\infty(\mathbb{R}^n)$ converging to v in α Sobolev norm. Since on ω, $v = u$, we have

$$\|v_m - u\|_{k,p,\omega} \le \|v_m - v\|_{k,p,\mathbb{R}^n} \to 0.$$

This is true for all relatively compact subset ω of Ω. □

Note that the restriction 'relatively compact set' is only for $k \ge 1$. For the case $k = 0$, the restriction to Ω works, as seen in Theorem 1.3.30. The density of $C_c^\infty(\Omega)$ in $W^{k,p}(\Omega)$ may fail to generalise for an arbitrary *proper* subset $\Omega \subset \mathbb{R}^n$, because a "bad" derivative may be introduced at the boundary while extending by zero outside Ω.

Example 2.6. Let $\Omega = (0,1) \subset \mathbb{R}$ and $u \equiv 1$ on Ω. Then $u \in W^{1,p}(0,1)$. Setting $\tilde{u} = 0$ in $\mathbb{R} \setminus (0,1)$, we see that $\tilde{u} \in L^p(\mathbb{R})$ but not in $W^{1,p}(\mathbb{R})$. Because $D\tilde{u} = \delta_0 - \delta_1$ is not in $L^1_{\text{loc}}(\mathbb{R})$.

Recall the inclusion $C_c^\infty(\Omega) \subset C^\infty(\overline{\Omega}) \subset C^\infty(\Omega)$. The $C^\infty(\overline{\Omega})$ denotes all functions in $C^\infty(\Omega)$ such that all its derivatives can be extended continuously to $\overline{\Omega}$. For an arbitrary subset $\Omega \subseteq \mathbb{R}^n$ the best one can do is the following density result.

Theorem 2.3.3 (Meyers-Serrin). *Let $1 \leq p < \infty$ and $\Omega \subseteq \mathbb{R}^n$. For $u \in W^{k,p}(\Omega)$ and any $\varepsilon > 0$, there is a $\phi \in C^\infty(\Omega)$ such that $\|\phi\|_{k,p,\Omega} < \infty$ and $\|u - \phi\|_{k,p,\Omega} < \varepsilon$.*

Proof. For each $m \in \mathbb{N}$, consider the sets

$$\omega_m := \{x \in \Omega \mid |x| < m \text{ and } \operatorname{dist}(x, \partial\Omega) > \frac{1}{m}\}$$

and set $\omega_0 = \emptyset$. Define the collection of open sets $\{U_m\}$ as $U_m := \omega_{m+1} \cap (\overline{\omega}_{m-1})^c$. Note that $\Omega = \cup_m U_m$ is an open covering of Ω. Thus, we choose the C^∞ locally finite partition of unity $\{\phi_m\} \subset C_c^\infty(\Omega)$ such that $\operatorname{supp}(\phi_m) \subset U_m$, $0 \leq \phi_m \leq 1$ and $\sum_m \phi_m = 1$. For the given $u \in W^{k,p}(\Omega)$, note that $\phi_m u \in W^{k,p}(U_m)$ with $\operatorname{supp}(\phi_m u) \subset U_m$. We extend $\phi_m u$ to all of $\mathbb{R}^n$ by zero outside U_m, i.e.,

$$\widetilde{\phi_m u}(x) = \begin{cases} \phi_m u(x) & x \in U_m \\ 0 & x \in \mathbb{R}^n \setminus U_m. \end{cases}$$

Observe that $\widetilde{\phi_m u} \in W^{k,p}(\mathbb{R}^n)$. Let ρ_δ be the sequence of mollifiers and consider the sequence $\rho_\delta * \widetilde{\phi_m u}$ in $C^\infty(\mathbb{R}^n)$. Support of $\rho_\delta * \widetilde{\phi_m u} \subset U_m + B(0;\delta)$. Note that for all $x \in U_m$, $1/(m+1) < \operatorname{dist}(x, \partial\Omega) < 1/(m-1)$. Thus, for all $0 < \delta < 1/(m+1)(m+2)$[4], $\operatorname{supp}(\rho_\delta * \widetilde{\phi_m u}) \subset \omega_{m+2} \cap (\overline{\omega}_{m-2})^c$, which is compactly contained in Ω. Since $\phi_m u \in W^{k,p}(\Omega)$, we can choose a subsequence $\{\delta_m\}$ going to zero in $(0, 1/(m+1)(m+2)$ such that

$$\|\rho_{\delta_m} * \phi_m u - \phi_m u\|_{k,p,\Omega} = \|\rho_{\delta_m} * \widetilde{\phi_m u} - \phi_m u\|_{k,p,\mathbb{R}^n} < \frac{\varepsilon}{2^m}.$$

Set $\phi = \sum_m \rho_{\delta_m} * \phi_m u$. Note that $\phi \in C^\infty(\Omega)$ and $\|\phi\|_{k,p,\Omega} < \infty$. Since every U_m intersects U_{m-1} and U_{m+1}, we at most have three non-zero terms in the sum for each $x \in U_m$, i.e.,

$$\phi(x) = \sum_{i=-1}^{1} (\rho_{\delta_{m+i}} * \phi_{m+i})(x) \quad x \in U_m.$$

[4] The choice of this range for δ is motivated from the fact that $\left(\frac{1}{m+1} - \frac{1}{(m+1)(m+2)}\right) = \frac{1}{(m+2)}$

Therefore,

$$\|u-\phi\|_{k,p,\Omega} = \|\sum_m (u\phi_m - \rho_{\delta_m} * \phi_m u)\|_{k,p,\Omega} \leq \sum_m \|\phi_m u - \rho_{\delta_m} * \phi_m u\|_{k,p,\Omega} < \varepsilon.$$

This proves the density of $C^\infty(\Omega)$ in $W^{k,p}(\Omega)$. □

Let $C^{k,p}(\Omega)$ denote the closure of $E := \{\phi \in C^k(\Omega) \mid \|\phi\|_{k,p,\Omega} < \infty\}$ w.r.t the Sobolev norm $\|\cdot\|_{k,p,\Omega}$. The space $C^{k,p}(\Omega)$ is a subspace of $W^{k,p}(\Omega)$ because classical derivative (continuous) and distributional derivative coincide (because integraion by parts is valid). Since $W^{k,p}(\Omega)$ is a Banach space, we have $C^{k,p}(\Omega)$ is a closed subspace of $W^{k,p}(\Omega)$. In fact, a consequence of above result is that the Sobolev space $W^{k,p}(\Omega) = C^{k,p}(\Omega)$, a result due to Meyers and Serrin (cf. [MS64]) proved in 1964.

Corollary 2.3.4 (Meyers-Serrin)**.** *Let $1 \leq p < \infty$ and $\Omega \subseteq \mathbb{R}^n$, then $C^{k,p}(\Omega) = W^{k,p}(\Omega)$.*

Proof. It is enough to show that E is dense in $W^{k,p}(\Omega)$ because, as a consequence, $W^{k,p}(\Omega) = \overline{E} = C^{k,p}(\Omega)$. For each given $\varepsilon > 0$ and $u \in W^{k,p}(\Omega)$, we need to show the existence of $\phi \in E$ such that $\|u-\phi\|_{k,p,\Omega} < \varepsilon$. By Theorem 2.3.3 there is a $\phi \in C^\infty(\Omega) \cap E$ such that $\|u-\phi\|_{k,p,\Omega} < \varepsilon$. Hence proved. □

The density of $C^\infty(\overline{\Omega})$ is not true, in general, and fails for some "bad" domains as seen in examples below. This, in turn, means that $C_c^\infty(\Omega)$ cannot, in general, be dense in $W^{k,p}(\Omega)$.

Example 2.7. Let $\Omega := \{(x,y) \in \mathbb{R}^2 \mid 0 < |x| < 1 \text{ and } 0 < y < 1\}$. Consider the function $u : \Omega \to \mathbb{R}$ defined as

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

Then for any fixed $\varepsilon > 0$ there exists no $\phi \in C^1(\overline{\Omega})$ such that $\|u-\phi\|_{1,p} < \varepsilon$.

Example 2.8. Let $\Omega := \{(r,\theta) \in \mathbb{R}^2 \mid 1 < r^2 < 2 \text{ and } \theta \neq 0\}$. Consider $u(r,\theta) = \theta$. Then there exists no $\phi \in C^1(\overline{\Omega})$ such that $\|u-\phi\|_{1,1} < 2\pi$.

The trouble with domains in above examples is that they lie on both sides of the boundary $\partial\Omega$ which becomes the main handicap while trying to approximate $W^{k,p}(\Omega)$ by $C^\infty(\overline{\Omega})$ functions.

Definition 2.3.5. *A subset $\Omega \subset \mathbb{R}^n$ is said to satisfy the* segment property *if for every $x \in \partial\Omega$, there is a open set B_x, containing x, and a non-zero unit vector e_x such that $z + te_x \in \Omega$, for all $0 < t < 1$ whenever $z \in \overline{\Omega} \cap B_x$.*

Observe that a domain with segment property cannot lie on both sides of its boundary.

Theorem 2.3.6. *Let $1 \leq p < \infty$ and $\Omega \subseteq \mathbb{R}^n$ satisfy the segment property. Then, for every $u \in W^{k,p}(\Omega)$ there exists a sequence $\{v_m\} \subset C_c^\infty(\mathbb{R}^n)$ such that $\|v_m - u\|_{k,p,\Omega} \to 0$, for all $k \geq 1$, i.e., $C_c^\infty(\mathbb{R}^n)$ is dense in $W^{k,p}(\Omega)$.*

Theorem 2.3.1 is a particular case of the above result, because $\mathbb{R}^n$ trivially satisfies the segment property as it has 'no boundary'. A stronger version of above result is Theorem 2.5.2 proved in next section.

Corollary 2.3.7. *Let $1 \leq p < \infty$ and $\Omega \subseteq \mathbb{R}^n$ satisfy the segment property. Then $C^\infty(\overline{\Omega})$ is dense in $W^{k,p}(\Omega)$.*

In general, the closure of $C_c^\infty(\Omega)$ is a proper subspace of $W^{k,p}(\Omega)$ w.r.t the $\|\cdot\|_{k,p,\Omega}$.

Definition 2.3.8. *Let $W_0^{k,p}(\Omega)$ be the closure of $C_c^\infty(\Omega)$ in $W^{k,p}(\Omega)$. For $p = 2$, we denote $W_0^{k,p}(\Omega)$ by $H_0^k(\Omega)$.*

Exercise 61. Show that $W_0^{k,p}(\Omega)$ is a closed subspace of $W^{k,p}(\Omega)$.

In general, $W_0^{k,p}(\Omega)$ is a strict subspace of $W^{k,p}(\Omega)$. However, as observed in Theorem 2.3.1, for $\Omega = \mathbb{R}^n$, $W^{k,p}(\mathbb{R}^n) = W_0^{k,p}(\mathbb{R}^n)$ for all $1 \leq p < \infty$ and $k \geq 0$. In fact, $W^{k,p}(\Omega) = W_0^{k,p}(\Omega)$ iff $\mathrm{cap}_p(\mathbb{R}^n \setminus \Omega) = 0$[5].

The density results discussed in this section is not true for $p = \infty$ case as seen from the example below.

Example 2.9. Let $\Omega = (-1, 1)$ and

$$u(x) = \begin{cases} 0 & x \leq 0 \\ x & x \geq 0. \end{cases}$$

Then its distributional derivative is $u'(x) = 1_{(0,\infty)}$. Let $\phi \in C^\infty(\Omega)$ such that $\|\phi' - u'\|_\infty < \varepsilon$. Thus, if $x < 0$, $|\phi'(x)| < \varepsilon$ and if $x > 0$, $|\phi'(x) - 1| < \varepsilon$. In particular, $\phi'(x) > 1 - \varepsilon$. By continuity, $\phi'(0) < \varepsilon$ and $\phi'(0) > 1 - \varepsilon$ which is impossible if $\varepsilon < 1/2$. Hence, u cannot be approximated by smooth functions in $W^{1,\infty}(\Omega)$ norm.

[5] (k,p) polar sets

2.4 Characterisation of $W_0^{k,p}(\Omega)$

Theorem 2.4.1 (Chain rule)**.** *Let $G \in C^1(\mathbb{R})$ such that $G(0) = 0$ and $|G'(x)| \leq M$ for all $x \in \mathbb{R}$. Also, let $u \in W^{1,p}(\Omega)$. Then $G \circ u \in W^{1,p}(\Omega)$ and*

$$\frac{\partial}{\partial x_i}(G \circ u) = (G' \circ u)\frac{\partial u}{\partial x_i}, \quad 1 \leq i \leq n.$$

Recall the definition of $W_0^{k,p}(\Omega)$ from §2.2. We have already characterised $W_0^{k,p}(\Omega)$ when $\Omega = \mathbb{R}^n$ (cf. Theorem 2.3.1).

Theorem 2.4.2. *Let $1 \leq p < \infty$ and let $u \in W^{1,p}(\Omega)$ such that u vanishes outside a compact set contained in Ω. Then $u \in W_0^{1,p}(\Omega)$.*

Theorem 2.4.3. *Let $1 \leq p \leq \infty$ and $u \in W^{1,p}(\Omega) \cap C(\overline{\Omega})$. If $u = 0$ on $\partial\Omega$ then $u \in W_0^{1,p}(\Omega)$.*

Theorem 2.4.4 (Stampacchia)**.** *Let $G : \mathbb{R} \to \mathbb{R}$ be a Lipschitz continuous function such that $G(0) = 0$. For $1 < p < \infty$, if Ω is bounded and $u \in W_0^{1,p}(\Omega)$, then we have $G \circ u \in W_0^{1,p}(\Omega)$.*

Exercise 62. Let $G : \mathbb{R} \to \mathbb{R}$ be a Lipschitz continuous function such that $G(0) = 0$. Let Ω be a bounded open subset of $\mathbb{R}^n$ which admits an extension operator. Then, for $1 < p < \infty$, $u \in W^{1,p}(\Omega)$ implies that $G \circ u \in W^{1,p}(\Omega)$.

Corollary 2.4.5. *Let Ω be a bounded open set of $\mathbb{R}^n$. If $u \in H_0^1(\Omega)$ then $|u|, u^+$ and u^- belong to $H_0^1(\Omega)$, where*

$$u^+(x) = \max\{u(x), 0\} \tag{2.4.1}$$
$$u^-(x) = \max\{-u(x), 0\} = \min\{u(x), 0\}. \tag{2.4.2}$$

Lemma 2.4.6. *Let $\Omega \subset \mathbb{R}^n$ be an open set and let $u \in W^{1,p}(\Omega)$. If $K \subset \Omega$ is a closed set and u vanishes outside K, then the function*

$$\tilde{u}(x) = \begin{cases} u(x) & \text{if } x \in \Omega \\ 0 & \text{if } x \in \mathbb{R}^n \setminus \Omega \end{cases}$$

belongs to $W^{1,p}(\mathbb{R}^n)$.

Exercise 63. In the above Lemma, if K is compact then $u \in W_0^{1,p}(\Omega)$.

Theorem 2.4.7. *Let $1 < p < \infty$ and Ω be an open subset of $\mathbb{R}^n$. If $u \in W_0^{1,p}(\Omega)$ then $\tilde{u} \in W^{1,p}(\mathbb{R}^n)$. Further, for any $1 \leq i \leq n$, $\frac{\partial \tilde{u}}{\partial x_i} = \widetilde{\frac{\partial u}{\partial x_i}}$.*

2.5 Extension Operators

In general, many properties of $W^{k,p}(\Omega)$ can be inherited from $W^{k,p}(\mathbb{R}^n)$ provided the domain is "nice". For instance, we will see that if $W^{k,p}(\mathbb{R}^n)$ is continuously imbedded in $L^q(\mathbb{R}^n)$ then, for nice domains, $W^{k,p}(\Omega)$ is also continuously imbedded in $L^q(\Omega)$. In this section we classify such classes of "nice" domain.

Definition 2.5.1. *Let Ω be an open subset of $\mathbb{R}^n$. We say P is an (k,p)-*extension *operator for Ω, if $P : W^{k,p}(\Omega) \to W^{k,p}(\mathbb{R}^n)$ is a bounded linear operator, i.e., there is a constant $C > 0$ (depending on Ω, k and p) such that*

$$\|Pu\|_{k,p,\mathbb{R}^n} \leq C\|u\|_{k,p,\Omega} \quad \forall u \in W^{k,p}(\Omega)$$

and $Pu\mid_\Omega = u$ a.e. for every $u \in W^{k,p}(\Omega)$. If P is same for all $1 \leq p < \infty$ and $0 \leq m \leq k$, then P is called strong k-extension *operator. If P is a strong k-extension operator for all k then P is called* total extension *operator.*

Example 2.10. There is a natural extension operator $P : W_0^{k,p}(\Omega) \to W^{k,p}(\mathbb{R}^n)$ which is the extension by zero. Define

$$Pu := \tilde{u} = \begin{cases} u(x) & x \in \Omega \\ 0 & x \in \mathbb{R}^n \setminus \Omega. \end{cases}$$

Obviously, $\tilde{u}\mid_\Omega = u$ and $\|\tilde{u}\|_{0,p,\mathbb{R}^n} = \|u\|_{0,p,\Omega}$. We shall show that $\|\tilde{u}\|_{k,p,\mathbb{R}^n} = \|u\|_{k,p,\Omega}$. Since $u \in W_0^{k,p}(\Omega)$, there is a sequence $\{\phi_m\} \subset C_c^\infty(\Omega)$ converging to u in $\|\cdot\|_{k,p,\Omega}$. For any $\phi \in C_c^\infty(\mathbb{R}^n)$ and $|\alpha| \leq k$, we have

$$\begin{aligned} D^\alpha \tilde{u}(\phi) &= (-1)^{|\alpha|} \int_{\mathbb{R}^n} \tilde{u} D^\alpha \phi \, dx = (-1)^{|\alpha|} \int_\Omega u D^\alpha \phi \, dx \\ &= (-1)^{|\alpha|} \lim_{m\to\infty} \int_\Omega \phi_m D^\alpha \phi \, dx = \lim_{m\to\infty} \int_\Omega D^\alpha \phi_m \phi \, dx \\ &= \int_\Omega D^\alpha u \phi \, dx = \widetilde{D^\alpha u}(\phi). \end{aligned}$$

Thus, $D^\alpha \tilde{u} = \widetilde{D^\alpha u}$ and therefore

$$\|\tilde{u}\|_{k,p,\mathbb{R}^n} = \sum_{|\alpha|=0}^k \|D^\alpha \tilde{u}\|_{p,\mathbb{R}^n} = \sum_{|\alpha|=0}^k \|\widetilde{D^\alpha u}\|_{p,\mathbb{R}^n} = \sum_{|\alpha|=0}^k \|D^\alpha u\|_{p,\Omega} = \|u\|_{k,p,\Omega}.$$

Theorem 2.5.2. *Let $1 \leq p < \infty$. If Ω is an open subset of $\mathbb{R}^n$ such that there is a (k,p)-extension operator $P : W^{k,p}(\Omega) \to W^{k,p}(\mathbb{R}^n)$, then $C_c^\infty(\mathbb{R}^n)$ is dense in $W^{k,p}(\Omega)$. In particular, $C^\infty(\overline{\Omega})$ is dense in $W^{k,p}(\Omega)$ w.r.t $\|\cdot\|_{k,p,\Omega}$.*

Proof. For each $u \in W^{k,p}(\Omega)$, we choose the sequence $\phi_m(\rho_m * Pu)$ in $C_c^\infty(\mathbb{R}^n)$ which converges to Pu in $W^{k,p}(\mathbb{R}^n)$ and their restriction to Ω is in $C^\infty(\overline{\Omega})$ and converges in $W^{k,p}(\Omega)$. □

Above result is a particular case of Theorem 2.3.6. A natural question provoked by Theorem 2.5.2 is: For what classes of open sets Ω can one expect an extension operator P.

Definition 2.5.3. *For an open set $\Omega \subset \mathbb{R}^n$ we say that its boundary $\partial\Omega$ is C^k $(k \geq 1)$, if for every point $x \in \partial\Omega$, there is a $r > 0$ and a C^k diffeomorphism $\gamma : B_r(x) \to B_1(0)$ (i.e. γ^{-1} exists and both γ and γ^{-1} are k-times continuously differentiable) such that*

1. *$\gamma(\partial\Omega \cap B_r(x)) \subset B_1(0) \cap \{x \in \mathbb{R}^n \mid x_n = 0\}$ and*
2. *$\gamma(\Omega \cap B_r(x)) \subset B_1(0) \cap \{x \in \mathbb{R}^n \mid x_n > 0\}$*

We say $\partial\Omega$ is C^∞ if $\partial\Omega$ is C^k for all $k = 1, 2, \dots$ and $\partial\Omega$ is analytic if γ is analytic.

Equivalently, a workable definition of C^k boundary would be the following: if for every point $x \in \partial\Omega$, there exists a neighbourhood B_x of x and a C^k function $\gamma : \mathbb{R}^{n-1} \to \mathbb{R}$ such that

$$\Omega \cap B_x = \{x \in B_x \mid x_n > \gamma(x_1, x_2, \dots, x_{n-1})\}.$$

To keep the illustration simple, we shall restrict ourselves to $k = 1$. For more general results we refer to [Ada75]. We begin by constructing an extension operator for the half-space and then use it along with partition of unity to construct an extension operator for C^1 boundary domains.

Theorem 2.5.4. *Let $\mathbb{R}^n_+ := \{x = (x', x_n) \in \mathbb{R}^n \mid x_n > 0\}$ where $x' = (x_1, x_2, \dots, x_{n-1})$. Given $u \in W^{1,p}(\mathbb{R}^n_+)$, we define the extension to $\mathbb{R}^n$ as*

$$Pu = u^*(x) := \begin{cases} u(x', x_n) & x_n > 0 \\ u(x', -x_n) & x_n < 0. \end{cases}$$

Then $\|u^*\|_{1,p,\mathbb{R}^n} \leq 2\|u\|_{1,p,\mathbb{R}^n_+}$ and $u^* \in W^{1,p}(\mathbb{R}^n)$.

Proof. Observe that $u^* \in L^p(\mathbb{R}^n)$ because

$$\begin{aligned}\|u^*\|^p_{p,\mathbb{R}^n} &= \int_{\mathbb{R}^n_+} |u(x',x_n)|^p\,dx + \int_{\mathbb{R}^n_-} |u(x',-x_n)|^p\,dx \\ &= \int_{\mathbb{R}^n_+} |u(x',x_n)|^p\,dx + \int_{\mathbb{R}^{n-1}} \int_0^\infty |u(x',y_n)|^p\,dx'dy_n \\ &= 2\int_{\mathbb{R}^n_+} |u(x)|^p\,dx = 2\|u\|^p_{p,\mathbb{R}^n_+}.\end{aligned}$$

Thus, $\|u^*\|_{p,\mathbb{R}^n} = 2^{1/p}\|u\|_{p,\mathbb{R}^n_+}$. We now show that for $\alpha = e_i$, $1 \le i \le n-1$, $D^\alpha u^* = (D^\alpha u)^*$. Consider, for $\phi \in \mathcal{D}(\mathbb{R}^n)$,

$$\begin{aligned}D^\alpha u^*(\phi) &= -\int_{\mathbb{R}^n} u^* D^\alpha \phi\,dx \\ &= -\int_{\mathbb{R}^n_+} u(x',x_n) D^\alpha\phi\,dx - \int_{\mathbb{R}^n_-} u(x',-x_n) D^\alpha \phi\,dx \\ &= -\int_{\mathbb{R}^n_+} u(x',x_n) D^\alpha\phi(x',x_n)\,dx - \int_{\mathbb{R}^n_+} u(x',x_n) D^\alpha\phi(x',-x_n)\,dx.\end{aligned}$$

Hence,

$$D^\alpha u^*(\phi) = -\int_{\mathbb{R}^n_+} u D^\alpha \psi(x)\,dx, \tag{2.5.1}$$

where $\psi(x',x_n) = \phi(x',x_n) + \phi(x',-x_n)$ when $x_n > 0$. In general, $\psi \notin \mathcal{D}(\mathbb{R}^n_+)$. Thus, we shall multiply ψ by a suitable cut-off function so that the product is in $\mathcal{D}(\mathbb{R}^n_+)$. Choose a $\{\zeta_m\} \in C^\infty(\mathbb{R})$ such that

$$\zeta_m(t) = \begin{cases} 0 & \text{if } t < 1/2m \\ 1 & \text{if } t > 1/m, \end{cases}$$

then $\zeta_m(x_n)\psi(x) \in \mathcal{D}(\mathbb{R}^n_+)$. Since ζ_m is independent of x_i, for $1 \le i \le n-1$, we have

$$D^\alpha u(\zeta_m\psi) = -\int_{\mathbb{R}^n_+} u(x) D^\alpha(\zeta_m(x_n)\psi(x))\,dx = -\int_{\mathbb{R}^n_+} u(x)\zeta_m(x_n) D^\alpha \psi(x)\,dx.$$

Passing to limit, as $m \to \infty$ both sides, we get

$$\int_{\mathbb{R}^n_+} D^\alpha u(x)\psi(x)\,dx = -\int_{\mathbb{R}^n_+} u(x) D^\alpha\psi(x)\,dx.$$

The RHS in above equation is same as the RHS obtained in (2.5.1). Hence, we have $D^\alpha u^*(\phi) = \int_{\mathbb{R}^n_+} D^\alpha u(x)\psi(x)\,dx$. By setting

$$(D^\alpha u)^*(x', x_n) := \begin{cases} D^\alpha u(x', x_n) & x_n > 0 \\ D^\alpha u(x', -x_n) & x_n < 0, \end{cases}$$

we get $D^\alpha u^* = (D^\alpha u)^*$. We now show a similar result for $\alpha = e_n$. Consider, for $\phi \in \mathcal{D}(\mathbb{R}^n)$,

$$\begin{aligned} D^{e_n} u^*(\phi) &= -\int_{\mathbb{R}^n} u^* D^{e_n}\phi\,dx \\ &= -\int_{\mathbb{R}^n_-} u(x', -x_n) D^{e_n}\phi\,dx - \int_{\mathbb{R}^n_+} u(x', x_n) D^{e_n}\phi\,dx \\ &= \int_{\mathbb{R}^n_+} u(x', x_n) D^{e_n}\phi(x', -x_n)\,dx - \int_{\mathbb{R}^n_+} u(x', x_n) D^{e_n}\phi(x', x_n)\,dx. \end{aligned}$$

Hence,

$$D^{e_n} u^*(\phi) = -\int_{\mathbb{R}^n_+} u D^{e_n}\psi(x)\,dx, \tag{2.5.2}$$

where $\psi(x', x_n) = \phi(x', x_n) - \phi(x', -x_n)$ when $x_n > 0$. Note that $\psi(x', 0) = 0$. Thus, by Mean value theorem, $|\psi(x', x_n)| \leq C|x_n|$. As before, in general, $\psi \notin \mathcal{D}(\mathbb{R}^n_+)$, so $\zeta_m(x_n)\psi(x) \in \mathcal{D}(\mathbb{R}^n_+)$. One such choice of ζ_m is by choosing $\zeta_m(t) = \zeta(mt)$ where

$$\zeta(t) = \begin{cases} 0 & \text{if } t < 1/2 \\ 1 & \text{if } t > 1. \end{cases}$$

Consider,

$$\begin{aligned} D^{e_n} u(\zeta_m \psi) &= -\int_{\mathbb{R}^n_+} u(x) D^{e_n}(\zeta_m(x_n)\psi(x))\,dx \\ &= -\int_{\mathbb{R}^n_+} u(x)\zeta_m'(x_n)\psi(x)\,dx - \int_{\mathbb{R}^n_+} u(x)\zeta_m(x_n) D^{e_n}\psi(x)\,dx. \end{aligned}$$

Passing to limit, as $m \to \infty$ both sides, we get

$$\int_{\mathbb{R}^n_+} D^{e_n} u(x)\psi(x)\,dx = -\lim_{m\to\infty} \int_{\mathbb{R}^n_+} u(x)\zeta_m'(x_n)\psi(x)\,dx - \int_{\mathbb{R}^n_+} u(x) D^{e_n}\psi(x)\,dx.$$

Let us handle the first term in RHS. Note that

$$\begin{aligned}
\left|\int_{\mathbb{R}^n_+} u(x)\zeta_m'(x_n)\psi(x)\,dx\right| &= m\left|\int_{\mathbb{R}^n_+} u(x)\zeta'(mx_n)\psi(x)\,dx\right| \\
&\leq mC\int_{\mathbb{R}^n_+} |u(x)||\zeta'||x_n|\,dx \\
&\leq C\|\zeta'\|_{\infty,[0,1]}\int_{\mathbb{R}^{n-1}}\int_{0<x_n<1/m} |u(x)|\,dx.
\end{aligned}$$

Therefore, $\lim_{m\to\infty}\int_{\mathbb{R}^n_+} u(x)\zeta_m'(x_n)\psi(x)\,dx = 0$ and hence

$$\int_{\mathbb{R}^n_+} D^{e_n}u(x)\psi(x)\,dx = -\int_{\mathbb{R}^n_+} u(x)D^{e_n}\psi(x)\,dx.$$

The RHS in above equation is same as the RHS obtained in (2.5.2). Hence, we have $D^{e_n}u^*(\phi) = \int_{\mathbb{R}^n_+} D^{e_n}u(x)\psi(x)\,dx$. By setting

$$(D^{e_n}u)^\sharp(x',x_n) := \begin{cases} D^{e_n}u(x',x_n) & x_n > 0 \\ -D^{e_n}u(x',-x_n) & x_n < 0, \end{cases}$$

we get $D^{e_n}u^* = (D^{e_n}u)^\sharp$. Therefore, we have the estimate,

$$\begin{aligned}
\|u^*\|_{1,p,\mathbb{R}^n} &= \|u^*\|_{p,\mathbb{R}^n} + \sum_{|\alpha|=1}\|D^\alpha u^*\|_{p,\mathbb{R}^n} \\
&= 2^{1/p}\|u\|_{p,\mathbb{R}^n_+} + \sum_{i=1}^{n-1}\|(D^{e_i}u)^*\|_{p,\mathbb{R}^n} + \|(D^{e_n}u)^\sharp\|_{p,\mathbb{R}^n} \\
&= 2^{1/p}\|u\|_{1,p,\mathbb{R}^n_+}.
\end{aligned}$$

Hence, $u^* \in W^{1,p}(\mathbb{R}^n)$. □

Theorem 2.5.5. *Let Ω be a bounded open subset of $\mathbb{R}^n$ with C^1 boundary. Then there is an extension operator $P : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$.*

Corollary 2.5.6. *For $1 \leq p < \infty$ and Ω be bounded open set with C^1 boundary, then $C^\infty(\overline{\Omega})$ is dense in $W^{k,p}(\Omega)$, for all $k \geq 1$.*

2.6 Topological Dual of Sobolev Spaces

We shall now introduce the topological dual of Sobolev spaces and motivate its notation. For any $1 \leq p < \infty$, $k = 0, 1, 2, \ldots$ and Ω be an open subset of $\mathbb{R}^n$. Let $X_{k,p}(\Omega)$ be the topological dual of $W_0^{k,p}(\Omega)$. Thus, if $F \in X_{k,p}(\Omega)$ then F is a continuous linear functional on $W_0^{k,p}(\Omega)$ and the norm of F is given as,

$$\|F\|_{X_{k,p}(\Omega)} := \sup_{\substack{u \in W_0^{k,p}(\Omega) \\ u \neq 0}} \frac{|F(u)|}{\|u\|_{k,p}}.$$

Note that the dual space is considered for $W_0^{k,p}(\Omega)$ and not for $W^{k,p}(\Omega)$. The reason is that $\mathcal{D}(\Omega)$ is dense in $W_0^{k,p}(\Omega)$ and hence $W_0^{k,p}(\Omega)$ will have a unique continuous extension (by Hahn-Banach) for any continuous linear functional defined on $\mathcal{D}(\Omega)$.

Example 2.11. In general the dual of $W^{k,p}(\Omega)$ may not even be a distribution. Note that, in general, $\mathcal{D}(\Omega)$ is not dense in $W^{k,p}(\Omega)$. Thus, its dual $[W^{k,p}(\Omega)]^*$ is not in the space of distributions $\mathcal{D}'(\Omega)$. Of course, the restriction to $\mathcal{D}(\Omega)$ of every $T \in [W^{k,p}(\Omega)]^*$ is a distribution but this restriction may not identify with T. For instance, consider $\mathbf{f} \in [L^2(\Omega)]^n$ with $|\mathbf{f}| \geq c > 0$ a.e. and $\operatorname{div}(\mathbf{f}) = 0$. Define

$$T(\phi) := \int_\Omega \mathbf{f} \cdot \nabla\phi \, dx.$$

Since $|T(\phi)| \leq \|\mathbf{f}\|_2 \|\nabla\phi\|_2$, we infer that $T \in [H^1(\Omega)]^*$. However, the restriction of T to $\mathcal{D}(\Omega)$ is the zero operator of $\mathcal{D}'(\Omega)$ because

$$\langle T, \phi \rangle = -\langle \operatorname{div}(\mathbf{f}), \phi \rangle = 0 \quad \forall \phi \in \mathcal{D}(\Omega).$$

Theorem 2.6.1 (Characterisation of $X_{1,p}(\Omega)$)**.** *Let $1 \leq p < \infty$ and let $F \in X_{1,p}(\Omega)$. Then there exist functions $f_0, f_1, \ldots, f_n \in L^q(\Omega)$ such that*

$$F(u) = \int_\Omega f_0 u \, dx + \sum_{i=1}^n \int_\Omega f_i \frac{\partial u}{\partial x_i} \quad \forall u \in W_0^{1,p}(\Omega)$$

and $\|F\|_{X_{1,p}(\Omega)} = \max_{0 \leq i \leq n} \|f_i\|_q$. Further, if Ω is bounded, one may assume $f_0 = 0$.

Proof. Recall that (cf. proof of Theorem 2.2.2) $\mathscr{T} : W_0^{1,p}(\Omega) \to (L^p(\Omega))^{n+1}$ defined as

$$\mathscr{T}(u) = \left(u, \frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_n}\right)$$

is an isometry. The norm in $(L^p(\Omega))^{n+1}$ is defined as $\|u\| := \left(\sum_{i=1}^{n+1} \|u_i\|_p^p\right)^{1/p}$, where $u = (u_1, u_2, \ldots, u_{n+1})$. Let $E = \mathscr{T}(W_0^{1,p}(\Omega)) \subset (L^p(\Omega))^{n+1}$. Observe that $F \circ \mathscr{T}^{-1}$ is a continuous linear functional on E. By Hahn Banach theorem, there is a continuous extension S of $F \circ \mathscr{T}^{-1}$ to all of $(L^p(\Omega))^{n+1}$. Now, by Riesz representation theorem, there exist $f_0, f_1, \ldots, f_n \in L^q(\Omega)$ such that

$$S(v) = \int_\Omega f_0 v_0 \, dx + \sum_{i=1}^n \int_\Omega f_i v_i \, dx \quad \forall v = (v_0, v_1, \ldots, v_n) \in (L^p(\Omega))^{n+1}$$

and $\|S\| = \|F \circ \mathscr{T}^{-1}\|$. Now, for any $u \in W_0^{1,p}(\Omega)$,

$$\begin{aligned} F(u) &= F \circ \mathscr{T}^{-1}\left(u, \frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_n}\right) \\ &= S\left(u, \frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_n}\right) \\ &= \int_\Omega f_0 u \, dx + \sum_{i=1}^n \int_\Omega f_i \frac{\partial u}{\partial x_i} \, dx. \end{aligned}$$

Also, $\|F\| = \|F \circ \mathscr{T}^{-1}\|$, by isometry of $\mathscr{T}$ and hence $\|F\| = \|S\| = \max_{0 \le i \le n} \|f_i\|_q$.

Further, if Ω is bounded we have by Poincaré inequality that the seminorm $\|\nabla u\|_p$ becomes a norm in $W_0^{1,p}(\Omega)$, thus by using the gradient map as an isometry from $W_0^{1,p}(\Omega)$ to $(L^p(\Omega))^n$ and arguing as above, we see that f_0 can be chosen to be zero. □

Remark 2.6.2. Let $F \in X_{1,p}(\Omega)$ and $\phi \in \mathcal{D}(\Omega)$, then

$$\begin{aligned} F(\phi) &= \int_\Omega f_0 \phi \, dx + \sum_{i=1}^n \int_\Omega f_i \frac{\partial \phi}{\partial x_i} \, dx \\ &= \int_\Omega f_0 \phi \, dx - \sum_{i=1}^n \int_\Omega \phi \frac{\partial f_i}{\partial x_i} \, dx. \end{aligned}$$

Now, since $\mathcal{D}(\Omega)$ is dense in $W_0^{1,p}(\Omega)$, the extension of F to $W_0^{1,p}(\Omega)$ should be unique. Henceforth, we shall identify any element $F \in X_{1,p}(\Omega)$ with the distribution

$$f_0 - \sum_{i=1}^n \frac{\partial f_i}{\partial x_i}.$$

The above remark motivates the right notation for the space $X_{k,p}(\Omega)$. Observe that if $u \in W^{k,p}(\Omega)$, then the first derivatives of $\frac{\partial u}{\partial x_i}$, for all i, are in $W^{k-1,p}$. To carry forward this feature in our notation, the above remark motivates to rewrite $X_{k,p}(\Omega)$ as $W^{-k,q}(\Omega)$, where q is the conjugate exponent corresponding to p.

Let us observe that the representation of F in terms f_i is not unique. Let $g \in W^{1,q}(\Omega)$ such that $\Delta g = 0$, i.e., $\sum_{i=1}^n \frac{\partial}{\partial x_i}\left(\frac{\partial g}{\partial x_i}\right) = 0$. Then

$$F = f_0 - \sum_{i=1}^n \frac{\partial\left(f_i + \frac{\partial g}{\partial x_i}\right)}{\partial x_i}$$

is also a representation of F.

Exercise 64. Show that

$$\|F\|_{W^{-1,q}(\Omega)} = \inf_{F=f_0-\sum_{i=1}^n \frac{\partial f_i}{\partial x_i}} \left\{ \left(\sum_{i=0}^n \|f_i\|_q^q \right)^{1/q} \right\}.$$

2.7 Fractional Order Sobolev Space

Let $1 \leq p < \infty$ and $0 < \sigma < 1$. The Sobolev spaces $W^{\sigma,p}(\Omega)$, for non-integral σ, is defined as:

$$W^{\sigma,p}(\Omega) := \left\{ u \in L^p(\Omega) \mid \frac{|u(x)-u(y)|}{|x-y|^{\sigma+(\frac{n}{p})}} \in L^p(\Omega \times \Omega) \right\} \tag{2.7.1}$$

with the obvious norm. For any positive real number s, set $k := \lfloor s \rfloor$, integral part and $\sigma := s$ is the fractional part. Note that $0 < \sigma < 1$.

$$W^{s,p}(\Omega) = \{u \in W^{k,p}(\Omega) | D^\alpha u \in W^{\sigma,p}(\Omega) \quad \text{for all } |\alpha| = k\}. \tag{2.7.2}$$

We denote by $W_0^{s,p}(\Omega)$, the closure of $\mathcal{D}(\Omega)$ in $W^{s,p}(\Omega)$ and $W^{-s,p'}(\Omega)$ is dual of $W_0^{s,p}(\Omega)$.

We begin by giving a characterisation of the space $H^1(\mathbb{R}^n)$ in terms of Fourier transform. Recall that for any $u \in L^1(\mathbb{R}^n)$ its Fourier transform $\hat{u}$ is defined as

$$\hat{u}(\xi) = \int_{\mathbb{R}^n} e^{-\imath 2\pi\xi\cdot x} u(x)\,dx. \tag{2.7.3}$$

Recall that $\hat{u} \in C_0(\mathbb{R}^n)$ and if $\hat{u} \in L^1(\mathbb{R}^n)$, one can invert the Fourier transform to obtain u from $\hat{u}$ by the following formula:

$$u(x) = \int_{\mathbb{R}^n} e^{\imath \xi \cdot x} \hat{u}(\xi)\, d\xi. \tag{2.7.4}$$

In particular, by Fourier transform, if $u \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ then $\hat{u} \in L^2(\mathbb{R}^n)$ and

$$\|\hat{u}\|_{L^2(\mathbb{R}^n)} = \|u\|_{L^2(\mathbb{R}^n)}. \tag{2.7.5}$$

Theorem 2.7.1. *The Sobolev space*

$$H^1(\mathbb{R}^n) = \{u \in L^2(\mathbb{R}^n) \mid (1 + |\xi|^2)^{\frac{1}{2}} \hat{u} \in L^2(\mathbb{R}^n)\}.$$

Further, the H^1 is given as

$$\|u\|_{H^1(\mathbb{R}^n)} = \|(1 + |\xi|^2)^{\frac{1}{2}} \hat{u}\|_{2,\mathbb{R}^n}.$$

Proof. Consider $u \in \mathcal{D}(\mathbb{R}^n)$. By definition,

$$\widehat{\frac{\partial u}{\partial x_k}}(\xi) = \int_{\mathbb{R}^n} e^{-\imath 2\pi \xi \cdot x} \frac{\partial u}{\partial x_k}(x)\, dx.$$

Using integration by parts and the fact that u has compact support, we obtain

$$\widehat{\frac{\partial u}{\partial x_k}}(\xi) = 2\pi \imath \xi_k \hat{u}(\xi). \tag{2.7.6}$$

By the density of $\mathcal{D}(\mathbb{R}^n)$ in $H^1(\mathbb{R}^n)$ (cf. Theorem 2.3.1), for any $u \in H^1(\mathbb{R}^n)$, there exists a sequence $\{u_m\}_{m\in\mathbb{N}} \subset \mathcal{D}(\mathbb{R}^n)$ such that $u_m \to u$ in the norm topology of $H^1(\mathbb{R}^n)$. So, for each $m \in \mathbb{N}$,

$$\widehat{\frac{\partial u_m}{\partial x_k}}(\xi) = 2\pi \imath \xi_k \hat{u}_m(\xi). \tag{2.7.7}$$

Since $u_m \to u$ in the $H^1(\mathbb{R}^n)$ norm, $u_m \to u$ in $L^2(\mathbb{R}^n)$ and $\frac{\partial u_m}{\partial x_k} \to \frac{\partial u}{\partial x_k}$ in $L^2(\mathbb{R}^n)$, we can extract a subsequence u_{m_ℓ} such that

$$\hat{u}_{m_\ell}(\xi) \to \hat{u}(\xi) \quad \text{for a.e. } \xi \in \mathbb{R}^n$$

and

$$\widehat{\frac{\partial u_{m_\ell}}{\partial x_k}}(\xi) \to \widehat{\frac{\partial u}{\partial x_k}}(\xi) \quad \text{for a.e. } \xi \in \mathbb{R}^n.$$

By the continuity of Fourier transform, we obtain

$$\widehat{\frac{\partial u}{\partial x_k}}(\xi) = 2\pi\imath\xi_k\hat{u}(\xi) \quad \text{for a.e. } \xi \in \mathbb{R}^n. \tag{2.7.8}$$

Now, if $u \in H^1(\mathbb{R}^n)$ then $\frac{\partial u}{\partial x_k} \in L^2(\mathbb{R}^n)$, for all $k = 1, 2, \ldots, n$. Thus, the Fourier transform of $\frac{\partial u}{\partial x_k}$ is well-defined and

$$\widehat{\frac{\partial u}{\partial x_k}}(\xi) = 2\pi\imath\xi_k\hat{u}(\xi).$$

Hence, $\xi_k\hat{u}(\xi) \in L^2(\mathbb{R}^n)$, for all $k = 1, 2, \ldots, n$. Conversely, if $u \in L^2(\mathbb{R}^n)$ such that $\xi_k\hat{u}(\xi) \in L^2(\mathbb{R}^n)$, for all $k = 1, 2, \ldots, n$, then $\frac{\partial u}{\partial x_k} \in L^2(\mathbb{R}^n)$ for all $k = 1, 2, \ldots, n$ and, hence, $u \in H^1(\mathbb{R}^n)$. Therefore, $u \in H^1(\mathbb{R}^n)$ if and only if $\hat{u} \in L^2(\mathbb{R}^n)$ and $\imath\xi_k\hat{u} \in L^2(\mathbb{R}^n)$, for all $k = 1, 2, \ldots, n$. This is equivalent to saying that $(1 + |\xi|^2)^{\frac{1}{2}}\hat{u} \in L^2(\mathbb{R}^n)$ (cf. Lemma 2.7.2). Further,

$$\begin{aligned}
\|u\|^2_{H^1(\mathbb{R}^n)} &= \int_{\mathbb{R}^n} \left(u^2(x) + \sum_{k=1}^{n} \left| \frac{\partial u}{\partial x_k}(x) \right|^2 \right) dx \\
&= \int_{\mathbb{R}^n} \left(|\hat{u}(\xi)|^2 + \sum_{k=1}^{n} \left| \widehat{\frac{\partial u}{\partial x_k}}(\xi) \right|^2 \right) d\xi \\
&= \int_{\mathbb{R}^n} (1 + |\xi|^2)|\hat{u}(\xi)|^2 \, d\xi = \|(1 + |\xi|^2)^{\frac{1}{2}}\hat{u}\|^2_{L^2(\mathbb{R}^n)}.
\end{aligned}$$

□

Lemma 2.7.2. *There exist positive constants C_1 and C_2 depending only on k and n such that*

$$C_1(1 + |\xi|^2)^k \leq \sum_{|\alpha| \leq k} |\xi^\alpha|^2 \leq C_2(1 + |\xi|^2)^k \quad \forall \xi \in \mathbb{R}^n.$$

Proof. Note that $|\xi|^2 = \xi_1^2 + \ldots + \xi_n^2$ and $|\xi^\alpha| = |\xi_1|^{\alpha_1} \ldots |\xi_n|^{\alpha_n}$. By induction argument on k, we can see that same powers of ξ occur in $(1 + |\xi|^2)^k$ and $\sum_{|\alpha| \leq k} |\xi^\alpha|^2$, albeit with different coefficients, which depend only on n and k. Since the number of terms is finite and depends again only on n and k, the result follows. □

Owing to the above lemma one can define the space $H^k(\mathbb{R}^n)$ as follows:

$$H^k(\mathbb{R}^n) = \{u \in L^2(\mathbb{R}^n) | (1+|\xi|^2)^{\frac{k}{2}} \hat{u}(\xi) \in L^2(\mathbb{R}^n)\}$$

and

$$\|u\|_{H^k(\mathbb{R}^n)} = \left(\int_{\mathbb{R}^n} (1+|\xi|^2)^k |\hat{u}(\xi)|^2 \, d\xi \right)^{\frac{1}{2}}.$$

A major interest of the above approach is that it suggests a natural definition of the space $H^s(\mathbb{R}^n)$, for $s \in \mathbb{R}$. The central point being that the integral

$$\int_{\mathbb{R}^n} (1+|\xi|^2)^s |\hat{u}(\xi)|^2 \, d\xi$$

is finite, for any $s \in \mathbb{R}$, since a^s makes sense for any $s \in \mathbb{R}$ when $a > 0$. In this case, $a = 1 + |\xi|^2$ is positive. Thus, we are motivated to give the following definition.

Definition 2.7.3. *Let $s \geq 0$ be non-negative real number. We define*

$$H^s(\mathbb{R}^n) = \{u \in L^2(\mathbb{R}^n) \mid (1+|\xi|^2)^{\frac{s}{2}} \hat{u} \in L^2(\mathbb{R}^n)\},$$

equipped with the scalar product, for any $u, v \in H^s(\mathbb{R}^n)$

$$\langle u, v \rangle_{H^s(\mathbb{R}^n)} := \int_{\mathbb{R}^n} (1+|\xi|^2)^s \hat{u}(\xi) \overline{\hat{v}}(\xi) \, d\xi$$

and the corresponding norm,

$$\|u\|_{H^s(\mathbb{R}^n)} := \left(\int_{\mathbb{R}^n} (1+|\xi|^2)^s |\hat{u}(\xi)|^2 \, d\xi \right)^{\frac{1}{2}}.$$

Theorem 2.7.4. *For any $s \in [0, \infty)$, $H^s(\mathbb{R}^n)$ is a Hilbert space. If $s = k \in \mathbb{N}$ then $H^s(\mathbb{R}^n) = H^k(\mathbb{R}^n) = W^{k,2}(\Omega)$ is the classical Sobolev space.*

Proof. Recall that the Fourier transfrom is an isomorphism from $H^s(\mathbb{R}^n)$ onto Lebesgue space $L^2_\mu(\mathbb{R}^n)$ where $\mu := (1+|\xi|^2)^s \, dx$ is the measure with density $(1+|\xi|^2)^s$ w.r.t the Lebesgue measure dx on $\mathbb{R}^n$. Moreover, the Fourier transform is an isometry and the Hilbert structure of the weighted Lebesgue space $L^2_\mu(\mathbb{R}^n)$ is passed on to $H^s(\mathbb{R}^n)$. Hence, $H^s(\mathbb{R}^n)$ is isomorphic to $L^2_\mu(\mathbb{R}^n)$. $\square$

For $s > 0$, we define $H^{-s}(\mathbb{R}^n)$ as the dual of $H^s(\mathbb{R}^n)$. The negative order Sobolev spaces has the following characterization:

Theorem 2.7.5. *Let $s \in (0, \infty)$. Then*

$$H^{-s}(\mathbb{R}^n) = \{u \in \mathcal{S}'(\mathbb{R}^n) \mid (1 + |\xi|^2)^{-\frac{s}{2}} \hat{u}(\xi) \in L^2(\mathbb{R}^n)\} \tag{2.7.9}$$

Proof. We shall give the proof for $s = 1$. If $u \in H^{-1}(\mathbb{R}^n)$ then

$$u = f_0 + \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i}, \quad f_0, f_1 \dots, f_n \in L^2(\mathbb{R}^n).$$

Hence, u is a tempered distribution and

$$\hat{u} = \hat{f}_0 + \sum_{i=1}^{n} (2\pi\imath)\xi_i \hat{f}_i.$$

Then $(1 + |\xi|^2)^{-\frac{1}{2}} \hat{u} \in L^2(\mathbb{R}^n)$, proving one inclusion in (2.7.9). To prove the reverse inclusion, consider $u \in \mathcal{S}'(\mathbb{R}^n)$ such that $(1 + |\xi|^2)^{-\frac{1}{2}} \hat{u}(\xi) \in L^2(\mathbb{R}^n)$. Let $\phi \in \mathcal{D}(\mathbb{R}^n)$. Then there exists $\psi \in \mathcal{S}(\mathbb{R}^n)$ such that $\phi = \hat{\psi}$. Set

$$k(\xi) := (1 + |\xi|^2)^{\frac{1}{2}} \text{ and } k_{-1}(\xi) := (1 + |\xi|^2)^{-\frac{1}{2}}.$$

Note that both k and k_{-1} are in $C^\infty(\mathbb{R}^n)$, we write

$$u(\phi) = u(\hat{\psi}) = \hat{u}(\psi) = (kk_{-1})\hat{u}(\psi) = k_{-1}\hat{u}(k\psi).$$

But $k_{-1}\hat{u} \in L^2(\mathbb{R}^n)$ and, hence,

$$u(\phi) = \int_{\mathbb{R}^n} (1 + |\xi|^2)^{-\frac{1}{2}} \hat{u}(\xi)(1 + |\xi|^2)^{\frac{1}{2}} \psi(\xi)\, d\xi.$$

Therefore,

$$|u(\phi)| \leq |(1 + |\xi|^2)^{-\frac{1}{2}} \hat{u}(\xi)|_{0,\mathbb{R}^n} |(1 + |\xi|^2)^{\frac{1}{2}} \psi(\xi)|_{0,\mathbb{R}^n}.$$

But

$$\begin{aligned}
\left|(1 + |\xi|^2)^{\frac{1}{2}} \psi(\xi)\right|^2_{0,\mathbb{R}^n} &= \int_{\mathbb{R}^n} (1 + |\xi|^2)\psi^2(\xi)\, d\xi \\
&= \int_{\mathbb{R}^n} (1 + |\xi|^2)\psi^2(-\xi)\, d\xi \\
&= \int_{\mathbb{R}^n} (1 + |\xi|^2)\hat{\hat{\psi}}^2(\xi)\, d\xi \\
&= \int_{\mathbb{R}^n} (1 + |\xi|^2)(\hat{\phi}(\xi))^2\, d\xi = \|\phi\|^2_{H^1(\mathbb{R}^n)}.
\end{aligned}$$

Thus u defines a continuous linear functional on $H^1(\mathbb{R}^n)$ and so $u \in H^{-1}(\mathbb{R}^n)$. Also,

$$\|u\|_{H^{-1}(\mathbb{R}^n)} = |(1 + |\xi|^2)^{-\frac{1}{2}} \hat{u}(\xi)|_{0,\mathbb{R}^n}.$$

$\square$

For an open subset Ω of $\mathbb{R}^n$, we may define the Sobolev spaces $H^s(\Omega)$, for real s, as the restrictions to Ω of elements of $H^s(\mathbb{R}^n)$.

Example 2.12. If δ_0 is the Dirac distribution, we know that $\hat{\delta_0} \equiv 1$ and, hence, $\delta_0 \in H^{-s}(\mathbb{R}^n)$ if and only if $(1 + |\xi|^2)^{-\frac{s}{2}} \in L^2(\mathbb{R}^n)$. This is true for $s > \frac{n}{2}$ since the integral in polar coordinates is

$$\int_0^\infty \frac{r^{n-1}}{(1 + r^2)^s}\, dr$$

is finite only when $s > \frac{n}{2}$.

2.8 Imbedding Results

In this section, we shall prove various inequalities which, in turn, implies the imbedding of Sobolev spaces in to classical spaces. By the definition of Sobolev spaces and their norms, we observe that $W^{k,p}(\Omega) \hookrightarrow L^p(\Omega)$ is continuously imbedded (cf. Definition 2.1.6). Let us make a remark on the intuitive importance of Sobolev norms. Recall the uniform norm or essential supremum on the space of bounded continuous functions $C_b(\Omega)$. One can, intuitively, think of essential supremum norm as capturing the “height” of the function. In a similar sense, the L^p norms, for $p < \infty$, capture the “height” and “width” of a function. In mathematical terms, “width” is same as the measure of the support of the function. The Sobolev norms captures “height”, “width” and “oscillations”. The Fourier transform measures oscillation (or frequency or wavelength) by the decay of the Fourier transform, i.e., the “oscillation” of a function is translated to “decay” of the its Fourier transform. Sobolev norms measures “oscillation” via its derivatives (or regularity). Thus, Sobolev imbedding results are precisely statements about functions after incorporating its “oscillation” information.

We shall restrict ourselves to $W^{1,p}$, for all $1 \leq p \leq \infty$, to make the presentation clear and later state the results for the spaces $W^{k,p}$, $k \geq 2$.

Recall that the L^p spaces are actually equivalence classes of functions with equivalence relation being "equality almost everywhere". This motivates the following definition.

Definition 2.8.1. *For any $u \in L^p$ $(1 \le p \le \infty)$, we say $u^\star$ is a* representative[6] *of u if $u = u^\star$ a.e.*

Theorem 2.8.2 (One dimensional case)**.** *Let $(a, b) \subseteq \mathbb{R}$ be an open interval and $1 \le p \le \infty$. If $u \in W^{1,p}(a, b)$ then there is a* representative *of u, $u^\star$, which is absolutely continuous (is in $AC(a, b)$).*

Proof. Since $u \in W^{1,p}(a, b)$, u is weakly differentiable and $u' \in L^p(a, b)$. For each $x \in (a, b)$, we define $v : (a, b) \to \mathbb{R}$ as

$$v(x) := \int_a^x u'(t)\, dt.$$

By definition, $v \in BV(a, b)$ and is differentiable a.e. Hence $v' = u'$ and $(v - u)' = 0$ a.e. Thus, $u = v - c$ a.e., where c is some constant. Set $u^\star := v - c$. We claim that $u^\star$ is absolutely continuous because

$$u^\star(b) - u^\star(a) = v(b) - v(a) = \int_a^b u'(t)\, dt = \int_a^b v'(t)\, dt = \int_a^b (u^\star)'(t)\, dt.$$

Thus, $u^\star$ is the absolutely continuous *representative* of u. □

Example 2.13. In the above result, we have shown that $W^{1,p}(a, b) \subset AC(a, b)$, in the sense of representatives. Let us give an example of a continuous function on a bounded interval $I \subset \mathbb{R}$ which does not belong to $H^1(I)$. Take $I = (-1, 1)$. For non-zero γ in the range $-\frac{1}{2} < \gamma \le \frac{1}{2}$, the function $|x|^\gamma \notin H^1(I)$ but is in $L^2(I)$. If $0 < \gamma \le \frac{1}{2}$ then $|x|^\gamma$ is continuous function and not in $H^1(I)$.

2.8.1 Sobolev Inequality $(1 \le p < n)$

If there exists a positive constant $C > 0$ such that

$$\|u\|_{r,\mathbb{R}^n} \le C\|\nabla u\|_{p,\mathbb{R}^n} \quad \forall u \in W^{1,p}(\mathbb{R}^n) \tag{2.8.1}$$

[6]This is not a standard usage in literature and is introduced by the author of this manuscript for convenience sake

for some $r \in [1, \infty)$ and $1 \leq p < \infty$, then we have a continuous imbedding of $W^{1,p}(\Omega)$ in to $L^r(\mathbb{R}^n)$ because

$$\|u\|_{r,\mathbb{R}^n} \leq C\|\nabla u\|_{p,\mathbb{R}^n} \leq C\|\nabla u\|_{p,\mathbb{R}^n} + C\|u\|_{p,\mathbb{R}^n} = C\|u\|_{1,p,\mathbb{R}^n}.$$

However, it is obvious that (2.8.1) is not a sufficient condition for continuous imbedding. The equation (2.8.1) is a Sobolev inequality, a stronger necessary condition for continuous imbedding.

Before we prove an inequality like (2.8.1), let's check it validity. When can we expect such an inequality? If $u \in W^{1,p}(\mathbb{R}^n)$ satisfies (2.8.1) for some $C > 0$ and r, then $u_\lambda(x) := u(\lambda x)$, for any $\lambda > 0$, also satisfies (2.8.1) for the same C and r. Since, for $1 \leq p < \infty$, $\|u_\lambda\|_p = \frac{1}{\lambda^{n/p}}\|u\|_p$ and $\|\nabla u_\lambda\|_p = \frac{\lambda}{\lambda^{n/p}}\|\nabla u\|_p$, we have

$$\begin{aligned} \frac{1}{\lambda^{n/r}}\|u\|_r &\leq C\frac{\lambda}{\lambda^{n/p}}\|\nabla u\|_p \\ \|u\|_r &\leq C\lambda^{1+\frac{n}{r}-\frac{n}{p}}\|\nabla u\|_p. \end{aligned}$$

The above obtained inequality being true for $\lambda > 0$ would contradict (2.8.1) except when $1 + \frac{n}{r} - \frac{n}{p} = 0$. Consequently, to expect an inequality of the kind (2.8.1), we need to have $\frac{1}{r} = \frac{1}{p} - \frac{1}{n}$. Therefore, $\frac{1}{p} - \frac{1}{n} > 0$ and hence $1 \leq p < n$ and $r = \frac{np}{n-p}$. This discussion motivates the definition of *Sobolev conjugate*, $p^\star$, corresponding to the index p.

Definition 2.8.3. *If $1 \leq p < n$, the* Sobolev conjugate *of p is defined as*

$$p^\star := \frac{np}{n-p}$$

Equivalently, $\frac{1}{p^\star} = \frac{1}{p} - \frac{1}{n}$. Also, $p^\star > p$.

Lemma 2.8.4 (Loomis-Whitney Inequality)**.** *Let $n \geq 2$. Let $f_1, f_2, \ldots, f_n \in L^{n-1}(\mathbb{R}^{n-1})$. For $x \in \mathbb{R}^n$, set $\hat{x_i} = (x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in \mathbb{R}^{n-1}$ for $1 \leq i \leq n$. Define $f(x) = f_1(\hat{x_1}) \ldots f_n(\hat{x_n})$ for $x \in \mathbb{R}^n$. Then $f \in L^1(\mathbb{R}^n)$ and*

$$\|f\|_{1,\mathbb{R}^n} \leq \prod_{i=1}^{n} \|f_i\|_{n-1,\mathbb{R}^{n-1}}.$$

Proof. Let $n = 2$, then

$$\begin{aligned} \|f\|_{1,\mathbb{R}^2} = \int_{\mathbb{R}^2} |f(x_1, x_2)|\, dx_1\, dx_2 &= \int_{\mathbb{R}^2} |f_1(x_2)||f_2(x_1)|\, dx_1\, dx_2 \\ &= \int_{\mathbb{R}} |f_2(x_1)|\, dx_1 \int_{\mathbb{R}} |f_1(x_2)|\, dx_2 \\ &= \|f_2\|_{1,\mathbb{R}} \|f_1\|_{1,\mathbb{R}}. \end{aligned}$$

Let $n = 3$, then using Cauchy-Schwarz inequality twice we get,

$$\begin{aligned} \|f(x)\|_{1,\mathbb{R}^3} &= \int_{\mathbb{R}^3} |f_1(x_2, x_3)||f_2(x_1, x_3)||f_3(x_1, x_2)|\, dx_1\, dx_2\, dx_3 \\ &= \int_{\mathbb{R}^2} |f_3(x_1, x_2)| \left(\int_{\mathbb{R}} |f_1(x_2, x_3)||f_2(x_1, x_3)|\, dx_3 \right) dx_1\, dx_2 \\ &\le \int_{\mathbb{R}^2} |f_3(x_1, x_2)| \left[\prod_{i=1}^{2} \left(\int_{\mathbb{R}} |f_i(\hat{x}_i)|^2\, dx_3 \right)^{\frac{1}{2}} \right] dx_1\, dx_2 \\ &= \int_{\mathbb{R}^2} |f_3(x_1, x_2)| g(x_2)^{1/2} h(x_1)^{1/2}\, dx_1\, dx_2 \\ &\le \left(\int_{\mathbb{R}^2} |f_3(x_1, x_2)|^2\, dx_1\, dx_2 \right)^{1/2} \left(\int_{\mathbb{R}^2} g(x_2) h(x_1)\, dx_1\, dx_2 \right)^{1/2} \\ &= \|f_3\|_{2,\mathbb{R}^2} \|f_2\|_{2,\mathbb{R}^2} \|f_1\|_{2,\mathbb{R}^2}. \end{aligned}$$

The general case will be proved by induction. Assume the result for n. Let $x \in \mathbb{R}^{n+1}$. Fix x_{n+1} and $x' = (x_1, \ldots, x_n)$, then by Hölder's inequality,

$$\int_{\mathbb{R}^n} |f(x)|\, dx' \le \|f_{n+1}\|_{n,\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f_1 f_2 \ldots f_n|^{n'} dx_1\, dx_2\, \ldots\, dx_n \right)^{1/n'} \tag{2.8.2}$$

where $n' = \frac{n}{n-1}$ is the conjugate exponent of n. Recall that $f_1, \ldots, f_n \in L^n(\mathbb{R}^n)$. Thus, by treating x_{n+1} as a fixed parameter, $|f_1|^{n'}, \ldots, |f_n|^{n'} \in L^{n-1}(\mathbb{R}^{n-1})$. Therefore, by induction hypothesis,

$$\begin{aligned} \int_{\mathbb{R}^n} |f_1 f_2 \ldots f_n|^{n'} dx_1\, \ldots\, dx_n &\le \prod_{i=1}^{n} \left(\int_{\mathbb{R}^{n-1}} |f_i|^{n'(n-1)}\, dx_1 \ldots\, dx_n \right)^{1/(n-1)} \\ &= \prod_{i=1}^{n} \left(\int_{\mathbb{R}^{n-1}} |f_i|^n\, dx_1 \ldots\, dx_n \right)^{n'/n} \\ &= \prod_{i=1}^{n} \|f_i\|_{n,\mathbb{R}^{n-1}}^{n'} \end{aligned}$$

Now, substituting above inequality in (2.8.2), we get

$$\int_{\mathbb{R}^n} |f(x)|\,dx' \leq \|f_{n+1}\|_{n,\mathbb{R}^n} \prod_{i=1}^{n} \|f_i\|_{n,\mathbb{R}^{n-1}}.$$

Integrate both sides with respect to x_{n+1}. We get,

$$\int_{\mathbb{R}^{n+1}} |f(x)|\,dx \leq \|f_{n+1}\|_{n,\mathbb{R}^n} \prod_{i=1}^{n} \|f_i\|_{n,\mathbb{R}^n}.$$

□

Theorem 2.8.5 (Gagliardo-Nirenberg-Sobolev Inequality)**.** *Let $1 \leq p < n$. Then there exists a constant $C > 0$ (depending on p and n) such that*

$$\|u\|_{p^\star,\mathbb{R}^n} \leq C\|\nabla u\|_{p,\mathbb{R}^n} \quad \forall u \in W^{1,p}(\mathbb{R}^n).$$

In particular, we have the continuous imbedding

$$W^{1,p}(\mathbb{R}^n) \hookrightarrow L^{p^\star}(\mathbb{R}^n).$$

Proof. We begin by proving for the case $p = 1$. Note that $p^\star = 1^\star = \frac{n}{n-1}$. We first prove the result for the space of test functions which is dense in $W^{1,p}(\mathbb{R}^n)$. Let $\phi \in \mathcal{D}(\mathbb{R}^n)$, $x \in \mathbb{R}^n$ and $1 \leq i \leq n$, then

$$\begin{aligned}
\phi(x) &= \int_{-\infty}^{x_i} D^{e_i}\phi(x_1,\ldots,x_{i-1},t,x_{i+1},\ldots,x_n)\,dt \\
|\phi(x)| &\leq \int_{-\infty}^{\infty} |D^{e_i}\phi(x_1,\ldots,x_{i-1},t,x_{i+1},\ldots,x_n)|\,dt =: f_i(\hat{x_i}) \\
|\phi(x)|^n &\leq \prod_{i=1}^{n} f_i(\hat{x_i}) \\
|\phi(x)|^{n/n-1} &\leq \prod_{i=1}^{n} |f_i(\hat{x_i})|^{1/n-1}.
\end{aligned}$$

Now, integrating both sides with respect to x, we have

$$\int_{\mathbb{R}^n} |\phi|^{n/n-1}\,dx \leq \int_{\mathbb{R}^n} \prod_{i=1}^{n} |f_i|^{1/n-1}\,dx. \tag{2.8.3}$$

Observe that $g_i := |f_i|^{1/n-1} \in L^{n-1}(\mathbb{R}^{n-1})$ for each $1 \leq i \leq n$. Hence, by Lemma 2.8.4, we have

$$\begin{aligned}\int_{\mathbb{R}^n} \prod_{i=1}^n |f_i|^{1/n-1}\,dx = \int_{\mathbb{R}^n} \prod_{i=1}^n |g_i|\,dx &\leq \prod_{i=1}^n \left(\int_{\mathbb{R}^{n-1}} |g_i|^{n-1}\,dx\right)^{1/n-1} \\ &= \prod_{i=1}^n \left(\int_{\mathbb{R}^{n-1}} |f_i|\,dx\right)^{1/n-1} \\ &= \prod_{i=1}^n \|f_i\|_{1,\mathbb{R}^{n-1}}^{1/n-1} = \prod_{i=1}^n \|D^{e_i}\phi\|_{1,\mathbb{R}^n}^{1/n-1}.\end{aligned}$$

Thus, substituting above inequality in (2.8.3), we get

$$\|\phi\|_{1^\star,\mathbb{R}^n} \leq \prod_{i=1}^n \|D^{e_i}\phi\|_{1,\mathbb{R}^n}^{1/n}$$

and consequently, we get

$$\|\phi\|_{1^\star,\mathbb{R}^n} \leq \prod_{i=1}^n \|D^{e_i}\phi\|_{1,\mathbb{R}^n}^{1/n} \leq \prod_{i=1}^n \|\nabla\phi\|_{1,\mathbb{R}^n}^{1/n} = \|\nabla\phi\|_{1,\mathbb{R}^n}.$$

Hence the result proved for $p = 1$. Let $\psi := |\phi|^\gamma$, where $\gamma > 1$ will be chosen appropriately during the subsequent steps of the proof. Also, if $\phi \in \mathcal{D}(\mathbb{R}^n)$, then $\psi \in C_c^1(\mathbb{R}^n)$. We shall apply the $p = 1$ result to ψ. Therefore,

$$\begin{aligned}\|\psi\|_{1^\star,\mathbb{R}^n} &\leq \|\nabla\psi\|_{1,\mathbb{R}^n} \\ \left(\int_{\mathbb{R}^n} |\phi|^{n\gamma/n-1}\,dx\right)^{n-1/n} &\leq \int_{\mathbb{R}^n} |\nabla|\phi|^\gamma|\,dx \\ &= \gamma\int_{\mathbb{R}^n} |\phi|^{\gamma-1}|\nabla\phi|\,dx \\ &\leq \gamma\left(\int_{\mathbb{R}^n} |\phi|^{(\gamma-1)q}\,dx\right)^{1/q}\left(\int_{\mathbb{R}^n} |\nabla\phi|^p\,dx\right)^{1/p} \\ &\quad \text{(using Hölder's Inequality).}\end{aligned}$$

Since we want only the gradient term on the RHS, we would like to bring the q norm term to LHS. If we choose γ such that $\frac{n\gamma}{n-1} = (\gamma - 1)q$, then we can club their powers. Thus, we get $\gamma := \frac{p(n-1)}{n-p}$. The fact that $p > 1$ implies that $\gamma > 1$, as we had demanded. Thus, the inequality obtained above reduces to

$$\|\phi\|_{p^\star,\mathbb{R}^n} \leq \frac{p(n-1)}{n-p}\|\nabla\phi\|_{p,\mathbb{R}^n} \quad \forall \phi \in \mathcal{D}(\mathbb{R}^n).$$

Now, for any $u \in W^{1,p}(\mathbb{R}^n)$, there is a sequence $\{u_m\} \subset \mathcal{D}(\mathbb{R}^n)$ (cf. Theorem 2.3.1) such that $u_m \to u$ in $W^{1,p}(\mathbb{R}^n)$. Therefore, u_m is Cauchy in $W^{1,p}(\mathbb{R}^n)$ and hence is Cauchy in $L^{p^\star}(\mathbb{R}^n)$ by above inequality. Since $L^{p^\star}(\mathbb{R}^n)$ is complete, u_m converges in $L^{p^\star}(\mathbb{R}^n)$ and should converge to u, since $u_m = \phi_m(\rho_m * u)$ (cf. Theorem 2.3.1). Thus, $u \in L^{p^\star}(\mathbb{R}^n)$ and the inequality is satisfied. Hence the theorem is proved for any $1 \leq p < n$. In fact, in the proof we have obtained the constant C to be $C = \frac{p(n-1)}{n-p}$ (and this is not the best constant). □

Corollary 2.8.6. *For any $1 \leq p < n$, $W^{1,p}(\mathbb{R}^n) \hookrightarrow L^r(\mathbb{R}^n)$ is continuously imbedded, for all $r \in [p, p^\star]$.*

Proof. Let $u \in W^{1,p}(\mathbb{R}^n)$. By Theorem 2.8.5, we have $u \in L^{p^\star}(\mathbb{R}^n)$. We need to show that $u \in L^r(\mathbb{R}^n)$ for any $r \in (p, p^\star)$. Since $1/r \in [1/p^\star, 1/p]$, there is a $0 \leq \lambda \leq 1$ such that $1/r = \lambda/p + (1-\lambda)/p^\star$. Consider,

$$\begin{aligned}
\|u\|_r^r &= \int_{\mathbb{R}^n} |u|^r\,dx = \int_{\mathbb{R}^n} |u|^{\lambda r}|u|^{(1-\lambda)r}\,dx \\
&\leq \left(\int_{\mathbb{R}^n} |u|^p\,dx\right)^{\frac{\lambda r}{p}} \left(\int_{\mathbb{R}^n} |u|^{p^\star}\,dx\right)^{\frac{(1-\lambda)r}{p^\star}} \quad \text{(by Hölder inequality)} \\
\|u\|_r &\leq \|u\|_p^\lambda \|u\|_{p^\star}^{1-\lambda} \\
&\leq \lambda\|u\|_p + (1-\lambda)\|u\|_{p^\star} \quad \text{(By generalised AM-GM inequality)} \\
&\leq \|u\|_p + \|u\|_{p^\star} \\
&\leq \|u\|_p + C\|\nabla u\|_p \quad \text{(By Theorem 2.8.5)} \\
&\leq \max\{C, 1\}\|u\|_{1,p}.
\end{aligned}$$

Hence the continuous imbedding is shown for all $r \in [p, p^\star]$. □

We now extend the Sobolev inequality for a proper subset $\Omega \subset \mathbb{R}^n$ with smooth boundary, using the extension operator.

Theorem 2.8.7 (Sobolev Inequality for a Subset)**.** *Let Ω be an open subset of $\mathbb{R}^n$ with C^1 boundary. Also, let $1 \leq p < n$. Then $W^{1,p}(\Omega) \hookrightarrow L^{p^\star}(\Omega)$ where the constant C obtained depends on p, n and Ω. Further, $W^{1,p}(\Omega) \hookrightarrow L^r(\Omega)$ for all $r \in [p, p^\star]$.*

Proof. Let C denote a generic constant in this proof. Since Ω has a C^1 boundary, by Theorem 2.5.5, there is an extension operator P such that for some constant $C > 0$ (depending on Ω and p)

$$\|Pu\|_{1,p,\mathbb{R}^n} \leq C\|u\|_{1,p,\Omega} \quad \forall u \in W^{1,p}(\Omega).$$

Moreover, by Theorem 2.8.5, there exists a constant $C > 0$ (depending on p and n) such that

$$\|Pu\|_{p^\star,\mathbb{R}^n} \leq C\|\nabla(Pu)\|_{p,\mathbb{R}^n} \quad \forall u \in W^{1,p}(\Omega).$$

Let $u \in W^{1,p}(\Omega)$, then

$$\|u\|_{p^\star,\Omega} \leq \|Pu\|_{p^\star,\mathbb{R}^n} \leq C\|\nabla(Pu)\|_{p,\mathbb{R}^n} \leq C\|Pu\|_{1,p,\mathbb{R}^n} \leq C\|u\|_{1,p,\Omega}.$$

where the final constant C is dependent on p, n and Ω. Similar ideas work to prove the continuous imbedding in $L^r(\Omega)$ for $r \in [p, p^\star]$. $\square$

Note that the above result says

$$\|u\|_{p^\star} \leq C\|u\|_{1,p} \quad \forall u \in W^{1,p}(\Omega)$$

and not

$$\|u\|_{p^\star} \leq C\|\nabla u\|_p \quad \forall u \in W^{1,p}(\Omega)$$

because constant functions may belong to $W^{1,p}(\Omega)$ (as happens for bounded subsets Ω) whose derivatives are zero.

Corollary 2.8.8 (For $W_0^{1,p}(\Omega)$)**.** *Let Ω be an open subset of $\mathbb{R}^n$ (not necessarily bounded) and $1 \leq p < n$. Then there is a constant $C > 0$ (depending on p and n) such that*

$$\|u\|_{p^\star} \leq C\|\nabla u\|_p \quad \forall u \in W_0^{1,p}(\Omega)$$

and

$$\|u\|_r \leq C\|u\|_{1,p} \quad \forall u \in W_0^{1,p}(\Omega), r \in [p, p^\star].$$

Proof. Use the fact that $\mathcal{D}(\Omega)$ is dense in $W_0^{1,p}(\Omega)$ and follow the last step of the proof of Theorem 2.8.5 to get

$$\|u\|_{p^\star} \leq C\|\nabla u\|_p \quad \forall u \in W_0^{1,p}(\Omega).$$

The proof of the second inequality remains unchanged from the case of $\mathbb{R}^n$. One can also extend functions in $W_0^{1,p}(\Omega)$ by zero outside Ω and use the results proved for $\mathbb{R}^n$ and then restrict back to Ω. $\square$

Note that the second inequality in the statement of the above corollary involves the $W^{1,p}$-norm of u and not the L^p-norm of the gradient of u. However, for bounded sets one can hope to get the inequality involving gradient of u.

Corollary 2.8.9. *Let Ω be a* bounded *open subset of $\mathbb{R}^n$. For $1 \leq p < n$ there is a constant C (depending on p, n, r and Ω) such that*

$$\|u\|_r \leq C\|\nabla u\|_p \quad \forall u \in W_0^{1,p}(\Omega) \quad \forall r \in [1, p^\star].$$

Proof. Let $1 \leq p < n$ and $u \in W_0^{1,p}(\Omega)$. Then, by previous corollary, there is a constant $C > 0$ (depending on p and n) $\|u\|_{p^\star} \leq C\|\nabla u\|_p$. For any $1 \leq r \leq p^\star$, there is a constant $C > 0$ (depending on r and Ω) such that $\|u\|_r \leq C\|u\|_{p^\star}$ (since Ω is bounded). Therefore, we have a constant $C > 0$ (depending on p, n, r and Ω) $\|u\|_r \leq C\|\nabla u\|_p$ for all $r \in [1, p^\star]$. □

Exercise 65 (Poincaré Inequality). Using the corollary proved above show that, for $1 \leq p \leq \infty$, then there is a constant C (depending on p and Ω) such that

$$\|u\|_p \leq C\|\nabla u\|_p \quad \forall u \in W_0^{1,p}(\Omega).$$

2.8.2 Poincaré Inequality

The Poincaré inequality proved in previous section can be directly proved without using the results of previous section, as shown below.

Theorem 2.8.10 (Poincaré Inequality)**.** *Let Ω be a bounded open subset of $\mathbb{R}^n$, then there is a constant C (depending on p and Ω) such that*

$$\|u\|_{p,\Omega} \leq C\|\nabla u\|_{p,\Omega} \quad \forall u \in W_0^{1,p}(\Omega). \tag{2.8.4}$$

Proof. Let $a > 0$ and suppose $\Omega = (-a, a)^n$. Let $u \in \mathcal{D}(\Omega)$ and $x = (x', x_n) \in \mathbb{R}^n$. Then

$$u(x) = \int_{-a}^{x_n} \frac{\partial u}{\partial x_n}(x', t)\, dt.$$

Moreover, $u(x', -a) = 0$. Thus, by Hölder's inequality,

$$\begin{aligned}
|u(x)| &\leq \left(\int_{-a}^{x_n} \left|\frac{\partial u}{\partial x_n}(x', t)\right|^p dt\right)^{1/p} |x_n + a|^{1/q} \\
|u(x)|^p &\leq |x_n + a|^{p/q} \int_{-a}^{a} \left|\frac{\partial u}{\partial x_n}(x', t)\right|^p dt \\
|u(x)|^p &\leq (2a)^{p/q} \int_{-a}^{a} \left|\frac{\partial u}{\partial x_n}(x', t)\right|^p dt.
\end{aligned}$$

First integrating w.r.t x' and then integrating w.r.t x_n we get

$$\|u\|_{p,\Omega}^p \leq (2a)^{(p/q)+1} \left\|\frac{\partial u}{\partial x_n}\right\|_{p,\Omega}^p, \quad \forall u \in \mathcal{D}(\Omega)$$

and taking $(1/p)$-th power both sides, we get

$$\|u\|_{p,\Omega} \leq 2a \left\|\frac{\partial u}{\partial x_n}\right\|_{p,\Omega}, \quad \forall u \in \mathcal{D}(\Omega).$$

Thus,

$$\|u\|_{p,\Omega} \leq 2a \left\|\frac{\partial u}{\partial x_n}\right\|_{p,\Omega} \leq 2a\|\nabla u\|_{p,\Omega}, \quad \forall u \in \mathcal{D}(\Omega).$$

By the density of $\mathcal{D}(\Omega)$ in $W_0^{1,p}(\Omega)$, we get the result for all $u \in W_0^{1,p}(\Omega)$.

Now, suppose Ω is not of the form $(-a, a)$, then $\Omega \subset (-a, a)$ for some $a > 0$, since Ω is bounded. Then any $u \in W_0^{1,p}(\Omega)$ can be extended to $W_0^{1,p}(-a, a)$ and use the result proved above. □

The Poincaré inequality makes the norm $\|\nabla u\|_p$ equivalent to $\|u\|_{1,p}$ in $W_0^{1,p}(\Omega)$.

Remark 2.8.11. Poincaré inequality is not true for $u \in W^{1,p}(\Omega)$. For instance, if $u \equiv c$, a constant, then $\nabla u = 0$ and hence $\|\nabla u\|_p = 0$ while $\|u\|_p > 0$. However, if $u \in W^{1,p}(\Omega)$ such that $u = 0$ on $\Gamma \subset \partial\Omega$, then Poincaré inequality remains valid for such u's.

Remark 2.8.12. Poincaré inequality is not true for unbounded domains. However, one can relax the bounded-ness hypothesis on Ω to bounded-ness along one particular direction, as seen from the proof.

2.8.3 Equality case, $p = n$

Theorem 2.8.13. *Let $n \geq 2$, $W^{1,n}(\mathbb{R}^n) \hookrightarrow L^r(\mathbb{R}^n)$ for all $r \in [n, \infty)$.*

Proof. Let $u \in \mathcal{D}(\mathbb{R}^n)$. Observe that the conjugate exponent of n is same as $1^\star$, the Sobolev conjugate of 1. Thus, we adopt the initial part of the proof of Theorem 2.8.5 to obtain

$$\|u\|_{1^\star,\mathbb{R}^n} \leq \|\nabla u\|_{1,\mathbb{R}^n}$$

where $1^\star = \frac{n}{n-1}$, the conjugate exponent of n. Applying this inequality to $|u|^\gamma$, for some $\gamma > 1$ as in Theorem 2.8.5, we get

$$\begin{aligned}
\left(\int_{\mathbb{R}^n} |u|^{\frac{n\gamma}{n-1}}\,dx\right)^{\frac{n-1}{n}} &\leq \gamma \left(\int_{\mathbb{R}^n} |u|^{\frac{(\gamma-1)n}{n-1}}\,dx\right)^{\frac{n-1}{n}} \left(\int_{\mathbb{R}^n} |\nabla u|^n\,dx\right)^{\frac{1}{n}} \\
\|u\|^\gamma_{\frac{n\gamma}{n-1}} &\leq \gamma \|u\|^{\gamma-1}_{\frac{n(\gamma-1)}{n-1}} \|\nabla u\|_n \\
&\leq \left(\|u\|_{\frac{n(\gamma-1)}{n-1}} + \|\nabla u\|_n\right)^\gamma \\
&\quad \text{Using } (a+b)^\gamma \geq \gamma a^{\gamma-1} b \text{ for } a, b \geq 0 \\
\|u\|_{\frac{n\gamma}{n-1}} &\leq \|u\|_{\frac{n(\gamma-1)}{n-1}} + \|\nabla u\|_n.
\end{aligned}$$

Now, by putting $\gamma = n$, we get

$$\|u\|_{\frac{n^2}{n-1}} \leq \|u\|_n + \|\nabla u\|_n = \|u\|_{1,n}.$$

and extending the argument, as done in Corollary 2.8.6, we get

$$\|u\|_r \leq C\|u\|_{1,n} \quad \forall r \in \left[n, \frac{n^2}{n-1}\right].$$

Now, repeating the argument for $\gamma = n + 1$, we get

$$\|u\|_r \leq C\|u\|_{1,n} \quad \forall r \in \left[\frac{n^2}{n-1}, \frac{n(n+1)}{n-1}\right].$$

Thus, continuing in similar manner for all $\gamma = n + 2, n + 3, \ldots$ we get

$$\|u\|_r \leq C\|u\|_{1,n} \quad \forall r \in [n, \infty).$$

The result extends to $W^{1,n}(\mathbb{R}^n)$ by similar density arguments of $\mathcal{D}(\mathbb{R}^n)$, as done for the $1 \leq p < n$ case. ∎

We now extend the results to proper subsets of $\mathbb{R}^n$. The proofs are similar to the equivalent statements from previous section.

Corollary 2.8.14 (For Subset)**.** *Let Ω be a bounded open subset of $\mathbb{R}^n$ with C^1 boundary and let $n \geq 2$. Then $W^{1,n}(\Omega) \hookrightarrow L^r(\Omega)$ for all $r \in [n, \infty)$.*

Corollary 2.8.15. *Let Ω be an open subset of $\mathbb{R}^n$ (not necessarily bounded) and $n \geq 2$. Then $W_0^{1,n}(\Omega) \hookrightarrow L^r(\Omega)$ for all $r \in [n, \infty)$. Further, if Ω is bounded, $W_0^{1,n}(\Omega) \hookrightarrow L^r(\Omega)$ for all $r \in [1, \infty)$.*

2.8.4 Morrey's Inequality ($n < p < \infty$)

Theorem 2.8.16 (Morrey's Inequality)**.** *Let $n < p < \infty$, then $W^{1,p}(\mathbb{R}^n) \hookrightarrow L^\infty(\mathbb{R}^n)$. Moreover, for any $u \in W^{1,p}(\mathbb{R}^n)$, there is a* representative *of u, $u^\star$, which is Hölder continuous with exponent $1 - n/p$ and there is a constant $C > 0$ (depending only on p and n) such that*

$$\|u^\star\|_{C^{0,\gamma}(\mathbb{R}^n)} \leq C\|u\|_{1,p}$$

where $\gamma := 1 - n/p$.

Proof. Let $u \in \mathcal{D}(\mathbb{R}^n)$ and let E be a cube of side a containing the origin and each of its sides being parallel to the coordinate axes of $\mathbb{R}^n$. Let $x \in E$. We have

$$\begin{aligned}
|u(x) - u(0)| &= \left|\int_0^1 \frac{d}{dt}(u(tx))\,dt\right| = \left|\int_0^1 \nabla u(tx)\cdot x\,dt\right| \\
&\leq \int_0^1 |\nabla u(tx)|\cdot|x|\,dt = \int_0^1 \sum_{i=1}^n |x_i| \left|\frac{\partial}{\partial x_i}u(tx)\right| dt \\
&\leq \int_0^1 \sum_{i=1}^n a \left|\frac{\partial}{\partial x_i}u(tx)\right| dt = a\int_0^1 \sum_{i=1}^n \left|\frac{\partial}{\partial x_i}u(tx)\right| dt
\end{aligned}$$

Let $\overline{u}$ denote the average of u over the cube E,

$$\overline{u} = \frac{1}{|E|}\int_E u(x)\,dx = \frac{1}{a^n}\int_E u(x)\,dx.$$

Consider,

$$\begin{aligned}
|\overline{u} - u(0)| &= \left| \frac{1}{a^n} \int_E u(x)\,dx - u(0) \right| \le \frac{1}{a^n} \int_E |u(x) - u(0)|\,dx \\
&\le \frac{a}{a^n} \int_E \int_0^1 \sum_{i=1}^n \left| \frac{\partial}{\partial x_i} u(tx) \right| dt\,dx \\
&= \frac{1}{a^{n-1}} \sum_{i=1}^n \int_0^1 \int_E \left| \frac{\partial}{\partial x_i} u(tx) \right| dx\,dt \quad \text{(Fubini's Theorem)} \\
&= \frac{1}{a^{n-1}} \sum_{i=1}^n \int_0^1 t^{-n} \int_{tE} \left| \frac{\partial}{\partial x_i} u(y) \right| dy\,dt \quad \text{(Change of variable)} \\
&\le \frac{1}{a^{n-1}} \sum_{i=1}^n \int_0^1 t^{-n} \left\| \frac{\partial u}{\partial x_i} \right\|_{p,E} (|tE|)^{1/q}\,dt \\
&\qquad \text{(By Hölder inequality and } tE \subset E \text{ for } 0 \le t \le 1) \\
&= \frac{a^{n/q}}{a^{n-1}} \|\nabla u\|_{p,E} \int_0^1 t^{-n} t^{n/q}\,dt \quad (q \text{ is conjugate exponent of } p) \\
&= a^{1-n/p} \|\nabla u\|_{p,E} \int_0^1 t^{-n/p}\,dt = \frac{a^{1-n/p}}{1 - n/p} \|\nabla u\|_{p,E}.
\end{aligned}$$

The above inequality is then true for any cube E of side length a with sides parallel to axes, by translating it in $\mathbb{R}^n$. Therefore, for any cube E of side a and $x \in E$, we have

$$|\overline{u} - u(x)| \le \frac{a^\gamma}{\gamma} \|\nabla u\|_{p,E} \tag{2.8.5}$$

where $\gamma := 1 - n/p$. Consider,

$$\begin{aligned}
|u(x)| &= |u(x) - \overline{u} + \overline{u}| \le \frac{1}{\gamma} \|\nabla u\|_{p,E} + \|u\|_{1,E} \\
&\le \frac{1}{\gamma} \|\nabla u\|_{p,E} + \|u\|_{p,E} \quad \text{(by Hölder's inequality)}
\end{aligned}$$

where we have used (2.8.5), in particular, for a unit cube E. Then

$$\|u\|_{\infty,\mathbb{R}^n} \le \frac{1}{\gamma} \|\nabla u\|_{p,\mathbb{R}^n} + \|u\|_{p,\mathbb{R}^n}.$$

Therefore,

$$\|u\|_{\infty,\mathbb{R}^n} \le C \|u\|_{1,p,\mathbb{R}^n}, \tag{2.8.6}$$

where $C = \max\{1, \frac{1}{\gamma}\}$, and hence $W^{1,p}(\mathbb{R}^n) \hookrightarrow L^\infty(\mathbb{R}^n)$.

Further, it follows from (2.8.5) that for any $x, y \in E$,

$$|u(x) - u(y)| \leq \frac{2a^\gamma}{\gamma}\|\nabla u\|_{p,E}.$$

Now, for any given $x, y \in \mathbb{R}^n$, one can always choose a cube E whose side $a = 2|x-y|$ and applying the above inequality, we get

$$|u(x) - u(y)| \leq \frac{2^{\gamma+1}}{\gamma}|x-y|^\gamma\|\nabla u\|_{p,E} \leq \frac{2^{\gamma+1}}{\gamma}|x-y|^\gamma\|\nabla u\|_{p,\mathbb{R}^n}.$$

Thus, u is Hölder continuous and its Hölder seminorm $p_\gamma(.)$ (cf.(2.1.2)) is bounded as below,

$$p_\gamma(u) \leq \frac{2^{\gamma+1}}{\gamma}\|\nabla u\|_{p,\mathbb{R}^n}.$$

and this together with (2.8.6) gives the bound for the γ-th Hölder norm,

$$\|u\|_{C^{0,\gamma}(\mathbb{R}^n)} \leq C\|u\|_{1,p,\mathbb{R}^n}.$$

By the density of $\mathcal{D}(\mathbb{R}^n)$ in $W^{1,p}(\mathbb{R}^n)$ we have a sequence $u_\varepsilon \to u$ in $W^{1,p}(\mathbb{R}^n)$. By the bound on Hölder norm, we find the sequence is also Cauchy in $C^{0,\gamma}(\mathbb{R}^n)$ and should converge to a *representative* of u, $u^\star$, in the γ-th Hölder norm. □

Remark 2.8.17. As usual, the results can be extended to $W^{1,p}(\Omega)$ for Ω bounded with C^1 smooth boundary and to $W_0^{1,p}(\Omega)$ for any open subset Ω.

Theorem 2.8.18 (Characterisation of $W^{1,\infty}$)**.** *For any $u \in W^{1,\infty}(\mathbb{R}^n)$ there is a representative $u^\star$ which is Lipschitz continuous from $\mathbb{R}^n$ to $\mathbb{R}$.*

Example 2.14. Let $p < n$. Consider the function $|x|^\delta$, for any choice of δ in $1 - \frac{n}{p} < \delta < 0$, is in $W^{1,p}(\Omega)$ which has no continuous representative.

Example 2.15. Let $p = n$ and $n \geq 2$. We shall give an example of a function in $W^{1,n}(\Omega)$ which has no continuous representative. We shall given an example for the case $n = 2$. Let $\Omega := \{x \in \mathbb{R}^n : |x| < R\}$ and $u(x) = (-\ln|x|)^\delta$ for $x \neq 0$. We have, using polar coordinates,

$$\int_{B_R(0)} u^n\,dx = R^{n-1}\omega_n\int_0^R(-\ln r)^{n\delta}r\,dr.$$

Using the change of variable $t = -\ln r$, we get

$$\int_{B_R(0)} u^n\,dx = R^{n-1}\omega_n \int_{-\ln R}^{+\infty} t^{n\delta} e^{-2t}\,dt < \infty \quad \text{for all } \delta.$$

Thus, $u \in L^p(\Omega)$ for every $\delta \in \mathbb{R}$. Further, for each $i = 1, 2, \dots, n$,

$$u_{x_i} = -\delta x_i |x|^{-2}(-\ln|x|)^{\delta-1}$$

and, therefore,

$$|\nabla u| = \left|\delta(-\ln|x|)^{\delta-1}\right| |x|^{-1}.$$

Thus, using polar coordinates, we get

$$\int_{B_R(0)} |\nabla u|^n\,dx = R^{n-1}\omega_n |\delta|^n \int_0^R |\ln r|^{n\delta-n} r^{1-n}\,dr.$$

Using the change of variable $t = -\ln r$, we get

$$\int_{B_R(0)} |\nabla u|^n\,dx = R^{n-1}\omega_n |\delta|^n \int_{-\ln R}^{\infty} |t|^{n(\delta-1)} e^{t(n-2)}\,dt.$$

If $n = 2$ then the integral is finite iff $n(1-\delta) > 1$ or $\delta < 1 - \frac{1}{n} = \frac{1}{2}$. In particular, ∇u represents the gradient of u in the distribution sense as well. We conclude that $u \in H^1(\Omega)$ iff $\delta < 1/2$. We point out that when $\delta > 0$, u is **unbounded** near 0. Thus, by taking $0 < \delta < \frac{1}{2}$, we obtain a function belonging to $H^1(\Omega)$ which blows up to $+\infty$ at the origin and which does not have a continuous representative.

2.8.5 Generalised Sobolev Imbedding

We shall now generalise the results of previous section to all derivative orders of $k \geq 2$.

Theorem 2.8.19. *Let $k \geq 1$ be an integer and $1 \leq p < \infty$. Then*

1. *If $p < n/k$, then $W^{k,p}(\mathbb{R}^n) \hookrightarrow L^r(\mathbb{R}^n)$ for all $r \in [p, np/(n-pk)]$.*

2. *If $p = n/k$, then $W^{k,n/k}(\mathbb{R}^n) \hookrightarrow L^r(\mathbb{R}^n)$ for all $r \in [n/k, \infty)$.*

3. *If $p > n/k$, then $W^{k,p}(\mathbb{R}^n) \hookrightarrow L^\infty(\mathbb{R}^n)$ and further there is a* representative *of u, say $u^\star$, whose k-th partial derivative is Hölder continuous with exponent γ and there is a constant $C > 0$ (depending only on p, k and n) such that*

$$\|u^\star\|_{C^{k-1-[n/p],\gamma}(\mathbb{R}^n)} \leq C\|u\|_{k,p}$$

where $\gamma := k - n/p - [k - n/p]$ and $[l]$ is the largest integer such that $[l] \leq l$.

Proof. 1. If $k = 1$ we know the above results are true. Let $k = 2$ and $p < n/2$. Then, for any $u \in W^{2,p}(\mathbb{R}^n)$, both $u, D^1u \in W^{1,p}(\mathbb{R}^n)$ and since $p < n/2 < n$, using Sobolev inequality, we have both $u, D^1u \in L^{p^\star}(\mathbb{R}^n)$ with $p^\star = np/(n-p)$. Thus, $u \in W^{1,p^\star}(\mathbb{R}^n)$. Now, since $p < n/2$, we have $p^\star < n$. Thus, using Sobolev inequality again again, we get $u \in L^{(p^\star)^\star}(\mathbb{R}^n)$. But $1/(p^\star)^\star = 1/p - 2/n$. Extending similar arguments for each case, we get the result. Note that when we say D^1u, we actually mean $D^\alpha u$ for each $|\alpha| = 1$. Such convention will be used throughout this proof.

2. We know the result for $k = 1$. Let $k = 2$ and $u \in W^{2,n/2}(\mathbb{R}^n)$. Then, $u, D^1u \in W^{1,n/2}(\mathbb{R}^n)$. Since $n/2 < n$ and $(n/2)^\star = n$, we have, using Sobolev inequality, $u \in W^{1,r}(\mathbb{R}^n)$ for all $r \in [n/2, n]$. Thus, $u \in W^{1,n}(\mathbb{R}^n)$, which is continuously imbedded in $L^r(\mathbb{R}^n)$ for all $r \in [n, \infty)$.

3. Let $k = 2$ and $p > n/2$. $W^{2,p}(\mathbb{R}^n)$ is continuously imbedded in $W^{1,p}(\mathbb{R}^n)$. If $p > n > n/2$, using Morrey's inequality, we have $W^{1,p}(\mathbb{R}^n)$ is continuously imbedded in $L^\infty(\mathbb{R}^n)$ and the Hölder norm estimate is true for both u and ∇u. If $n/2 < p < n$ then $p^\star > n$ and, by Sobolev inequality, any $u \in W^{2,p}(\mathbb{R}^n)$ is also in $W^{1,r}(\mathbb{R}^n)$ for all $r \in [p, p^*]$. Now, by Morrey's inequality, $W^{1,p^\star}(\mathbb{R}^n)$ is continuously imbedded in $L^\infty(\mathbb{R}^n)$ and the Hölder norm estimate is true for both u and ∇u. Let $p = n > n/2$ then for any $u \in W^{2,n}(\mathbb{R}^n)$, we have $u \in W^{1,r}(\mathbb{R}^n)$ for all $r \in [n, \infty)$. Thus for $r > n$, we have the required imbedding in $L^\infty(\mathbb{R}^n)$. It now only remains to show the Hölder estimate.

□

Remark 2.8.20. As usual, the results can be extended to $W^{1,p}(\Omega)$ for Ω bounded with C^k smooth boundary and to $W_0^{1,p}(\Omega)$ for any open subset Ω.

2.8.6 Compact Imbedding

Our aim in this section is to isolate those continuous imbedding which are also compact. We first note that we cannot expect compact imbedding for unbounded domains.

Example 2.16. We shall construct a bounded sequence in $W^{1,p}(\mathbb{R})$ and show that it can not converge in $L^r(\mathbb{R})$, for all those r for which the imbedding is continuous. Let $I = (0,1) \subset \mathbb{R}$ and $I_j := (j, j+1)$, for all $j = 1, 2, \ldots$. Choose any $f \in C^1(\mathbb{R})$ with support in I and set $f_j(x) := f(x-j)$. Thus, f_j is same as f except that its support is now contained in I_j. Hence $\|f\|_{1,p} = \|f_j\|_{1,p}$ for all j. Now, set $g := \frac{f}{\|f\|_{1,p}}$ and $g_j := \frac{f_j}{\|f_j\|_{1,p}}$. Note that $\{g_j\}$ is a bounded sequence (norm being one) in $W^{1,p}(\mathbb{R})$. We know that $W^{1,p}$ is continuously imbedded in $L^\infty(\mathbb{R})$ for $1 < p < \infty$ and with $p = 1$ is imbedded in $L^r(\mathbb{R})$ for all $1 \leq r < \infty$. Since g_j have compact support, $\{g_j\} \subset L^r(\mathbb{R})$ for all $1 \leq r \leq \infty$ (depending on p), by the continuous imbedding. We will show that $\{g_j\}$ do not converge strongly in $L^r(\mathbb{R})$. Note that $\|g_j\|_{r,\mathbb{R}} = \|g_j\|_{r,I_j} = \|g\|_{r,I} = c > 0$. Consider for any $i \neq j$,

$$\|g_i - g_j\|^r_{r,\mathbb{R}} = \|g_i\|^r_{r,I_i} + \|g_j\|^r_{r,I_j} + \|g_i - g_j\|^r_{r,\mathbb{R}\setminus(I_i \cup I_j)} = 2c^r.$$

Thus, the sequence is not Cauchy in $L^r(\mathbb{R})$ (as seen by choosing $\varepsilon < 2^{1/r}c$ for all i, j). The arguments can be generalised to $\mathbb{R}^n$.

Theorem 2.8.21. *Let $1 \leq p \leq \infty$. For all $u \in W^{1,p}(\mathbb{R}^n)$ the following inequality holds:*

$$\|\tau_h u - u\|_p \leq \|\nabla u\|_p |h|, \quad \forall h \in \mathbb{R}^n.$$

Proof. It is enough to prove the inequality for $u \in \mathcal{D}(\mathbb{R}^n)$ due to the density of $\mathcal{D}(\mathbb{R}^n)$ in $W^{1,p}(\mathbb{R}^n)$. Consider

$$\begin{aligned}(\tau_h u)(x) - u(x) &= u(x-h) - u(x) \\ &= -\int_0^1 [\nabla u(x - th)]\, h \, dt.\end{aligned}$$

Hence, by the Cauchy-Schwarz inequality

$$|(\tau_h u)(x) - u(x)| \leq \int_0^1 |\nabla u(x - th)||h| \, dt,$$

and then, by Hölder's inequality,

$$|(\tau_h u)(x) - u(x)|^p \leq |h|^p \int_0^1 |\nabla u(x - th)|^p \, dt.$$

Integrating over $\mathbb{R}^n$ we have

$$\int_{\mathbb{R}^n} |(\tau_h u)(x) - u(x)|^p \, dx \leq |h|^p \int_{\mathbb{R}^n} \left(\int_0^1 |\nabla u(x - th)|^p \, dt \right) dx.$$

By Fubini-Tonelli theorem and the invariance, under translation, of the Lebesgue measure in $\mathbb{R}^n$ to obtain

$$\int_{\mathbb{R}^n} |\tau_h u - u|^p \, dx \leq |h|^p \int_{\mathbb{R}^n} |\nabla u(x)|^p \, dx.$$

$\square$

Theorem 2.8.22 (Rellich-Kondrasov). *Let $1 \leq p < \infty$ and let Ω be a bounded domain with C^1 boundary, then*

(i) If $p < n$, then $W^{1,p}(\Omega) \subset\subset L^r(\Omega)$ for all $r \in [1, p^\star)$.

(ii) If $p = n$, then $W^{1,n}(\Omega) \subset\subset L^r(\Omega)$ for all $r \in [1, \infty)$.

(iii) If $p > n$, then $W^{1,p}(\Omega) \subset\subset C(\overline{\Omega})$.

Proof. (i) We first prove the case $p < n$. Let B be the unit ball in $W^{1,p}(\Omega)$. We shall verify conditions (i) and (ii) of Theorem 1.3.31. Let $1 \leq q < p^*$. Then choose α such that $0 < \alpha \leq 1$ and

$$\frac{1}{q} = \frac{\alpha}{1} + \frac{1-\alpha}{p^*}.$$

If $u \in B$, $\Omega' \subset\subset \Omega$ and $h \in \mathbb{R}^n$ such that $|h| < \operatorname{dist}(\Omega', \mathbb{R}^n \setminus \Omega)$,

$$\begin{aligned} \|\tau_{-h} u - u\|_{q,\Omega'} &\leq \|\tau_{-h} u - u\|_{1,\Omega'}^{\alpha} \|\tau_{-h} u - u\|_{p^\star,\Omega'}^{1-\alpha} \\ &\leq (|h|^\alpha \|\nabla u\|_{1,\Omega}^{\alpha})(2\|u\|_{p^\star,\Omega})^{1-\alpha} \\ &\leq C|h|^\alpha. \end{aligned}$$

We choose h small enough such that $C|h|^\alpha < \varepsilon$. This will verify condition (i) of Theorem 1.3.31. Now, if $u \in B$ and $\Omega' \subset\subset \Omega$, it follows by Hölder's inequality that

$$\begin{aligned}
\|u\|_{q,\Omega\setminus\overline{\Omega}'} &\leq \|u\|_{p^\star,\Omega\setminus\overline{\Omega}'}|\Omega\setminus\overline{\Omega}'|^{1-(q/p^\star)} \\
&\leq C|\Omega\setminus\overline{\Omega}'|^{1-(q/p^\star)}
\end{aligned}$$

which can be made less than any given $\varepsilon > 0$ by choosing $\Omega' \subset\subset \Omega$ to be 'as closely filling Ω' as needed. This verifies condition (ii). Thus B is relatively compact in $L^q(\Omega)$ for $1 \leq q < p^\star$.

(ii) Assume for the moment that the result is true for $p < n$. Notice that as $p \to n$, $p^\star \to \infty$. Since Ω is bounded, $W^{1,n}(\Omega) \subset W^{1,n-\varepsilon}(\Omega)$, for every $\varepsilon > 0$. Since $n - \epsilon < n$, using the $p < n$ case, we get $W^{1,n-\varepsilon}(\Omega)$ is compactly imbedded in $L^r(\Omega)$ for all $r \in [1,(n-\varepsilon)^\star)$. Note that as $\varepsilon \to 0$, $(n-\varepsilon)^\star \to \infty$. Therefore, for any $r < \infty$ we can find small enough $\varepsilon > 0$ such that $1 \leq r < (n-\epsilon)^*$. We deduce that $W^{1,n}(\Omega)$ is compactly imbedded in $L^r(\Omega)$ for any $1 \leq r < \infty$.

(iii) For $p > n$, the functions of $W^{1,p}(\Omega)$ are Hölder continuous. If B is the unit ball in $W^{1,p}(\Omega)$ then the functions in B are uniformly bounded and equicontinuous in $C(\bar{\Omega})$. Thus B is relatively compact in $C(\bar{\Omega})$ by the Ascoli-Arzela Theorem.

□

Remark 2.8.23. Note that the continuous inclusion for the $r = p^\star$ case is not compact. The above result can be extended to $W_0^{1,p}(\Omega)$ provided Ω is bounded and is a *connected* open subset (bounded domain) of $\mathbb{R}^n$.

Corollary 2.8.24 (Compact subsets of $W^{1,p}(\mathbb{R}^n)$)**.** *Let A be a subset of $W^{1,p}(\mathbb{R}^n)$, $1 \leq p < +\infty$ which satisfies the two following conditions:*

(i) A is bounded in $W^{1,p}(\mathbb{R}^n)$, i.e., $\sup_{f\in A}\|f\|_{W^{1,p}(\mathbb{R}^n)} < +\infty$.

(ii) A is L^p-equi-integrable at infinity, i.e., $\lim_{r\to+\infty}\int_{\{|x|>r\}}|f(x)|^p\,dx = 0$ uniformly with respect to $f \in A$. Then, A is relatively compact in $L^p(\mathbb{R}^n)$.

Exercise 66. If Ω is connected subset of $\mathbb{R}^n$ and $u \in W^{1,p}(\Omega)$ such that $\nabla u = 0$ a.e. in Ω, then show that u is constant a.e. in Ω

Theorem 2.8.25 (Poincaré-Wirtinger Inequality)**.** *Let Ω be a bounded, connected open subset of $\mathbb{R}^n$ with C^1 smooth boundary and let $1 \leq p \leq \infty$. Then there is a constant $C > 0$ (depending on p, n and Ω) such that*

$$\|u - \overline{u}\|_p \leq C\|\nabla u\|_p \quad \forall u \in W^{1,p}(\Omega),$$

where $\overline{u} := \frac{1}{|\Omega|}\int_\Omega u(x)\,dx$ is the average of u in Ω.

Proof. Suppose the inequality was false, then for each positive integer m we have a $u_m \in W^{1,p}(\Omega)$ such that

$$\|u_m - \overline{u_m}\|_p > m\|\nabla u_m\|_p \quad \forall u \in W^{1,p}(\Omega). \tag{2.8.7}$$

Set for all m,

$$v_m := \frac{u_m - \overline{u_m}}{\|u_m - \overline{u_m}\|_p}.$$

Thus, $\|v_m\|_p = 1$ and $\overline{v_m} = 0$. Hence, by (2.8.7), we have

$$\|\nabla v_m\|_p < 1/m.$$

Therefore, $\{v_m\}$ are bounded in $W^{1,p}(\Omega)$ and, by Rellich-Kondrasov compact imbedding, there is a subsequence of $\{v_m\}$ (still denoted by m) and a function $v \in L^p(\Omega)$ such that $v_m \to v$ in $L^p(\Omega)$. Therefore, $\overline{v} = 0$ and $\|v\|_p = 1$. Also, for any $\phi \in \mathcal{D}(\Omega)$,

$$\begin{aligned}\int_\Omega v\frac{\partial\phi}{\partial x_i}\,dx &= \lim_{m\to\infty}\int_\Omega v_m\frac{\partial\phi}{\partial x_i}\,dx \\ &= -\lim_{m\to\infty}\int_\Omega \frac{\partial v_m}{\partial x_i}\phi\,dx \\ &\to 0.\end{aligned}$$

Thus, $\nabla v = 0$ a.e. and $v \in W^{1,p}(\Omega)$. Moreover v is constant, since Ω is connected. But $\overline{v} = 0$ implies that $v \equiv 0$ which contradicts the fact that $\|v\|_p = 1$. $\square$

Theorem 2.8.26 (Compactness for measures)**.** *$\mathcal{R}(\Omega)$ is compactly imbedded in $W^{-1,r}(\Omega)$ for all $r \in \left[1, \frac{n}{n-1}\right)$.*

Proof. Let $\{\mu_k\}_1^\infty$ be a bounded sequence in $\mathcal{R}(\Omega)$. By weak compactness, we extract a subsequence $\{\mu_{k_j}\}_{j=1}^\infty \subset \{\mu_k\}_{k=1}^\infty$ such that $\mu_{k_j} \rightharpoonup \mu$ weak-*

converges in $\mathcal{R}(\Omega)$, for some measure $\mu \in \mathcal{R}(\Omega)$. Set $s = \frac{r}{r-1}$ and denote by B the closed unit ball in $W_0^{1,s}(\Omega)$. Since $1 \le r < \frac{n}{n-1}$, we have $s > n$ and so B is compact in $C_0(\overline{\Omega})$. Thus, given $\varepsilon > 0$ there exist $\{\phi_i\}_{i=1}^{N(\varepsilon)} \subset C_0(\overline{\Omega})$ such that

$$\min_{1 \le i \le N(\varepsilon)} \|\phi - \phi_i\|_{C(\overline{\Omega})} < \varepsilon$$

for each $\phi \in B$. Therefore, if $\phi \in B$,

$$\left| \int_\Omega \phi \, d\mu_{k_j} - \int_\Omega \phi \, d\mu \right| \le 2\varepsilon \sup_j |\mu_{k_j}|(\Omega) + \left| \int_\Omega \phi_i \, d\mu_{k_j} - \int_\Omega \phi_i \, d\mu \right|$$

for some index $1 \le i \le N(\varepsilon)$. Consequently,

$$\lim_{j \to \infty} \sup_{\phi \in B} \left| \int_\Omega \phi \, d\mu_{k_j} - \int_\Omega \phi \, d\mu \right| = 0$$

and, hence, $\mu_{k_j} \to \mu$ strongly in $W^{-1,r}(\Omega)$. □

2.9 Trace Theory

Let $u : \overline{\Omega} \to \mathbb{R}$ be a continuous function, where Ω is an open subset of $\mathbb{R}^n$. The *trace* of u on the boundary $\partial\Omega$ is the continuous function $\gamma(u) : \partial\Omega \to \mathbb{R}$ defined by $\gamma(u)(x) = u(x)$ for all $x \in \partial\Omega$. Thus, we have an operator $\gamma : C(\overline{\Omega}) \to C(\partial\Omega)$ which is the restriction to the boundary. If $u \in L^2(\Omega)$ then there is no sufficient information to talk about u on $\partial\Omega$, because the Lebesgue measure of $\partial\Omega$ is zero. However, an additional information on u, viz., "$\frac{\partial u}{\partial x_i}$ belongs to $L^2(\Omega)$ for any $i = 1, \dots, n$", can give meaning to u restricted on $\partial\Omega$. Thus, in the Sobolev space $W^{1,p}(\Omega)$, where Ω is a *domain* in $\mathbb{R}^n$, the notion of *trace* or restriction to boundary can be defined on $\partial\Omega$, even for functions not continuous on $\overline{\Omega}$. The basis for this extension is the following observation: Let Ω be a domain in $\mathbb{R}^n$ and let $1 \le p < n$. Then one can show that the mapping

$$\gamma : C^\infty(\overline{\Omega}) \to L^{p^\sharp}(\partial\Omega)$$

is well defined and continuous if the space $C^\infty(\overline{\Omega})$ is endowed with the $W^{1,p}$-norm where $p^\sharp > 1$ is defined

$$\frac{1}{p^\sharp} := \frac{1}{p} - \frac{p-1}{p(n-1)} \quad \text{for } 1 \le p < n.$$

Since the space $C^\infty(\overline{\Omega})$ is dense in $W^{1,p}(\Omega)$ and $L^{p^\sharp}(\partial\Omega)$ is complete, there exists a unique continuous linear extension of γ from $C^\infty(\overline{\Omega})$ to $W^{1,p}(\Omega)$. This extension, still denoted as γ, is called the *trace operator* and each $\gamma(u) \in L^{p^\sharp}(\partial\Omega)$ is called the *trace* of the function $u \in W^{1,p}(\Omega)$.

Theorem 2.9.1. *Let Ω be a domain in $\mathbb{R}^n$ and $1 \leq p < \infty$. The trace operator γ satisfies the following:*

(a) $\gamma : W^{1,p}(\Omega) \to L^{p^\sharp}(\partial\Omega)$ is a continuous linear operator, for $1 \leq p < n$.

(b) For $p = n$, $\gamma : W^{1,n}(\Omega) \to L^q(\partial\Omega)$ is a continuous linear operator, for all $q \in [1,\infty)$.

(c) For $p > n$, $\gamma : W^{1,p}(\Omega) \to C(\partial\Omega)$ is a continuous linear operator.

Further, if $1 < p < n$ then the trace operator $\gamma : W^{1,p}(\Omega) \to L^q(\partial\Omega)$ is compact for all $q \in [1, p^\sharp)$.

A consequence of the above Theorem is that, independent of the dimension n, the trace map $\gamma : W^{1,p}(\Omega) \to L^p(\partial\Omega)$ is a continuous linear map for all $1 \leq p < \infty$.

Lemma 2.9.2. *Let $\Omega := \mathbb{R}^n_+ = \{x = (x', x_n) \in \mathbb{R}^{n-1} \times \mathbb{R} : x_n > 0\}$. Then, for any $1 \leq p < +\infty$, and $u \in \mathcal{D}(\bar{R}^n_+)$*

$$\|\gamma(u)\|_{L^p(\mathbb{R}^{n-1})} \leq p^{\frac{1}{p}} \|u\|_{W^{1,p}(\mathbb{R}^n_+)}.$$

Proof. Let $u \in \mathcal{D}(\bar{\mathbb{R}}^n_+)$. For any $x' \in \mathbb{R}^{n-1}$ we have

$$\begin{aligned} |u(x',0)|^p &= -\int_0^{+\infty} \frac{\partial}{\partial x_n} |u(x',x_n)|^p \, dx_n \\ &\leq p \int_0^{+\infty} |u(x',x_n)|^{p-1} \left| \frac{\partial u}{\partial x_n}(x',x_n) \right| dx_n. \end{aligned}$$

By Young's convexity inequality

$$ab \leq \frac{1}{p} a^p + \frac{1}{q} b^q$$

with $\frac{1}{p} + \frac{1}{q} = 1$ to the following situation:

$$a = \left| \frac{\partial u}{\partial x_n}(x',x_n) \right| \quad \text{and} \quad b = |u(x',x_n)|^{p-1}.$$

We obtain

$$|u(x',0)|^p \leq p\left[\int_0^{+\infty}\left(\frac{1}{p}\left|\frac{\partial u}{\partial x_n}(x',x_n)\right|^p + \frac{1}{q}|u(x',x_n)|^{(p-1)q}\right)dx_n\right].$$

By using the relation $(p-1)q = p$ we obtain

$$|u(x',0)|^p \leq (p-1)\int_0^{+\infty}|v(x',x_n)|^p\,dx_n + \int_0^{+\infty}\left|\frac{\partial u}{\partial x_n}(x',x_n)\right|^p\,dx_n.$$

Integrating over $\mathbb{R}^{n-1}$ yields

$$\begin{aligned}
\int_{\mathbb{R}^{n-1}}|u(x',0)|^p\,dx' &\leq (p-1)\int_{\mathbb{R}^n_+}|u(x)|^p\,dx + \int_{\mathbb{R}^n_+}\left|\frac{\partial u}{\partial x_n}(x)\right|^p\,dx\\
&\leq (p-1)\int_{\mathbb{R}^n_+}|u(x)|^p\,dx + \int_{\mathbb{R}^n_+}\sum_{i=1}^n\left|\frac{\partial u}{\partial x_i}\right|^p\,dx\\
&\leq p\int_{\mathbb{R}^n_+}\left(|u(x)|^p + \sum_{i=1}^n\left|\frac{\partial u}{\partial x_i}\right|^p\right)dx.
\end{aligned}$$

Hence

$$\|u(\cdot,0)\|_{L^p(\mathbb{R}^{n-1})} \leq p^{\frac{1}{p}}\|u\|_{W^{1,p}(\mathbb{R}^n_+)}.$$

□

Theorem 2.9.3. *Let Ω be bounded and $\partial\Omega$ is in the class C^k. Then there is a bounded linear operator $\gamma : W^{k,p}(\Omega) \to L^p(\partial\Omega)$ such that*

1. $\gamma(u) = u\,|_{\partial\Omega}$ *if $u \in W^{k,p}(\Omega) \cap C(\overline{\Omega})$,*
2. $\|\gamma(u)\|_{p,\partial\Omega} \leq C\|u\|_{1,p,\Omega}$ *for all $u \in W^{k,p}(\Omega)$ and C depending on p and Ω.*

We call $\gamma(u)$ to be the trace *of u on $\partial\Omega$.*

Proof. By the regularity property of the boundary $\partial\Omega$, we know that $\mathcal{D}(\bar{\Omega})$ is dense in $W^{1,p}(\Omega)$. For any $v \in \mathcal{D}(\bar{\Omega})$, we can define the restriction of u to $\partial\Omega$, setting $\gamma(u) := u|_{\partial\Omega}$. Suppose we prove that

$$\gamma : (\mathcal{D}(\bar{\Omega}),\, \|\cdot\|_{W^{1,p}(\Omega)}) \longrightarrow (L^p(\partial\Omega),\, \|\cdot\|_{L^p(\partial\Omega)})$$

is continuous, then, by the linearity of γ, it is uniformly continuous. The space $L^p(\partial\Omega)$ is a Banach space. Therefore, there exists a unique linear and continuous extension of γ

$$\gamma : W^{1,p}(\Omega) \longrightarrow L^p(\partial\Omega).$$

Thus, we just need to prove that γ is continuous. By Lemma 2.9.2, the result is true for half-space $\Omega = \mathbb{R}^n_+$. Let $\bar{\Omega} \subset \bigcup_{i=0}^k G_i$ with $\bar{G_0} \subset \Omega$, G_i open for all $i = 0, \ldots, k$, while $\phi_i : B(0,1) \to G_i$, $i = 1, 2, \ldots, k$, are the local coordinates $\{\alpha_0, \ldots, \alpha_i\}$, is an associated partition of unity, i.e., $\alpha_i \in \mathcal{D}(G_i)$, $\alpha_i \leq 0$, $\sum_{i=0}^k \alpha_i = 1$ on $\bar{\Omega}$.

$$w_i = \begin{cases} (\alpha_i v) \circ \phi_i & \text{on } B_+, \\ 0 & \text{on } \mathbb{R}^n_+ \backslash B_+. \end{cases}$$

Clearly w_i belongs to $\mathcal{D}(\mathbb{R}^n_+)$. By Lemma 2.9.2, we have

$$\|w_i(\cdot, 0)\|_{L^p(\mathbb{R}^{n-1})} \leq p^{\frac{1}{p}} \|w_i\|_{W^{1,p}(\mathbb{R}^n_+)}. \tag{2.9.1}$$

By using classical differential calculus rules (note that all the functions α_i, v, ϕ_i are continuously differentiable), one obtains the existence, for any $i = 1, \ldots, k$, of a constant C_i such that

$$\|w_i\|_{W^{1,p}(\mathbb{R}^n_+)} \leq C_i \|v\|_{W^{1,p}(\Omega)}. \tag{2.9.2}$$

Combining the two inequalities (2.9.1) and (2.9.2), we obtain

$$\|w_i(\cdot, 0)\|_{L^p(\mathbb{R}^{n-1})} \leq C_i p^{\frac{1}{p}} \|v\|_{W^{1,p}(\Omega)}. \tag{2.9.3}$$

We now use the definition of the $L^p(\partial\Omega)$ norm which is based on the use of local coordinates. One can show that an equivalent norm to the $L^p(\partial\Omega)$ norm can be obtained by using local coordinates: denoting by $\sim$ the extension by zero outside of $\mathbb{R}^{n-1} \backslash \{y \in \mathbb{R}^{n-1} \mid |y| < 1\}$, we have that

$$L^p(\partial\Omega) = \{u : \partial\Omega \to \mathbb{R} \mid \widetilde{(\alpha_i u)} \circ \phi_i(\cdot, 0) \in L^p(\mathbb{R}^{n-1}),\ 1 \leq i \leq k\}$$

and

$$u \mapsto \left(\sum_{i=1}^k \|\widetilde{(\alpha_i u)} \circ \phi_i\|^p_{L^p(\mathbb{R}^{n-1})} \right)^{\frac{1}{p}} \tag{2.9.4}$$

is an equivalent norm to the $L^p(\partial\Omega)$ norm.
This definition of the $L^p(\partial\Omega)$ norm and the inequality (2.9.3) (note that $w_i = (\widetilde{\alpha_i v}) \circ \phi_i$) yield

$$\|u\|_{L^p(\partial\Omega)} \leq C(p,n,\Omega)\|u\|_{W^{1,p}(\Omega)}$$

for some constant $C(p,n,\Omega)$. Thus, γ_0 is continuous. □

Theorem 2.9.4 (range of γ)**.** *Let Ω be an open bounded set in $\mathbb{R}^n$ whose boundary $\partial\Omega$ is of class $\mathbf{C}^1$. Then the trace operator γ is linear continuous and surjective from $H^1(\Omega) \to H^{\frac{1}{2}}(\partial\Omega)$.*

Proof. The definition of $H^{\frac{1}{2}}(\partial\Omega)$ is obtained by local coordinates. Thus, it is enough to prove when $\Omega = \mathbb{R}^n_+$ and $\partial\Omega = \mathbb{R}^{n-1}$. We do the proof in three steps.

First step: Let $w(x') = v(x',0)$. We first relate the Fourier transform (in $\mathbb{R}^{n-1}$) of w to that of v (in $\mathbb{R}^n$). We denote by $\widetilde{w}$ the Fourier transform of w in $\mathbb{R}^{n-1}$. By the Fourier inversion formula, if $v \in \mathcal{D}(\mathbb{R}^n)$,

$$\begin{aligned} v(x',0) &= \int_{\mathbb{R}^n} e^{2\pi\imath x'\cdot\xi'}\hat{v}(\xi)\,d\xi, \quad \xi = (\xi',\xi_n) \\ &= \int_{\mathbb{R}^n} e^{2\pi\imath x'\cdot\xi'}\left(\int_{-\infty}^{\infty}\hat{v}(\xi)\,d\xi_n\right)d\xi'. \end{aligned}$$

Now applying the same formula in $\mathbb{R}^{n-1}$, we get

$$v(x',0) = w(x') = \int_{\mathbb{R}^{n-1}} e^{2\pi\imath x'\cdot\xi'}\widetilde{w}(\xi')\,d\xi'.$$

By the uniqueness of this formula (since $w, \widetilde{w} \in \mathcal{S}(\mathbb{R}^{n-1})$ if $v \in \mathcal{D}(\mathbb{R}^n)$) we deduce that

$$\widetilde{w}(\xi') = \int_{-\infty}^{\infty}\hat{v}(\xi)\,d\xi_n.$$

Second step: To show that $v(x',0) = w(x') \in H^{\frac{1}{2}}(\mathbb{R}^{n-1})$ we need to show that $(1+|\xi'|^2)^{\frac{1}{2}}|\widetilde{w}(\xi')|^2$ is integrable. But

$$\begin{aligned}
&\int_{\mathbb{R}^{n-1}} (1+|\xi'|^2)^{\frac{1}{2}}|\widetilde{w}(\xi')|^2\,d\xi' \\
= &\int_{\mathbb{R}^{n-1}} (1+|\xi'|^2)^{\frac{1}{2}}\left|\int_{-\infty}^{\infty}\hat{v}(\xi)\,d\xi_n\right|^2 d\xi' \\
= &\int_{\mathbb{R}^{n-1}} (1+|\xi'|^2)^{\frac{1}{2}}\left|\int_{-\infty}^{\infty}\hat{v}(\xi)(1+|\xi|^2)^{\frac{-1}{2}}(1+|\xi|^2)^{\frac{1}{2}}\,d\xi_n\right|^2 d\xi' \\
\leq &\int_{\mathbb{R}^{n-1}} (1+|\xi'|^2)^{\frac{1}{2}}\left[\int_{-\infty}^{\infty}(1+|\xi|^2)|\hat{v}(\xi)|^2\,d\xi_n\int_{-\infty}^{\infty}(1+|\xi|^2)^{-1}\,d\xi_n\right] d\xi' \\
= &\;\pi\int_{\mathbb{R}^n}(1+|\xi|^2)|\hat{v}(\xi)|^2\,d\xi = \pi\|v\|^2_{H^1(\mathbb{R}^n)}
\end{aligned}$$

since

$$\int_{-\infty}^{\infty}(1+|\xi|^2)^{-1}\,d\xi_n = \int_{-\infty}^{\infty}\frac{d\xi_n}{1+|\xi'|^2+\xi_n^2} = \pi(1+|\xi'|^2)^{-\frac{1}{2}}.$$

The above formula is obtained by introducing the change of variable $\xi_n = (1+|\xi'|^2)^{\frac{1}{2}}\tan\theta$. Then

$$\frac{1}{(1+|\xi'|^2)^{\frac{1}{2}}}\int_{-\pi/2}^{\pi/2}\frac{\sec^2\theta}{1+\tan^2\theta}\,d\theta = (1+|\xi'|^2)^{\frac{-1}{2}}\int_{-\pi/2}^{\pi/2}d\theta = \pi(1+|\xi'|^2)^{\frac{-1}{2}}.$$

Thus, if $v \in \mathcal{D}(\mathbb{R}^n)$, $v(x',0) \in H^{\frac{1}{2}}(\mathbb{R}^{n-1})$ and by density the result follows for $v \in H^1(\mathbb{R}^n_+)$.

Third step: We now show that γ is onto $H^{\frac{1}{2}}(\mathbb{R}^{n-1})$. Let $h(x') \in H^{\frac{1}{2}}(\mathbb{R}^{n-1})$. Let $\widetilde{h}(\xi')$ be its Fourier transform. We define $u(x',x_n)$ by

$$\widetilde{u}(\xi',x_n) = e^{-(1+|\xi'|)x_n}\widetilde{h}(\xi'). \tag{2.9.5}$$

We must first show that u then belongs to $H^1(\mathbb{R}^n_+)$. To see this extend u by zero outside $\overline{\mathbb{R}^n_+}$. Now

$$\begin{aligned}
\hat{u}(\xi) &= \int_{\mathbb{R}^n} e^{-2\pi\imath x\cdot\xi} u(x)\,dx \\
&= \int_0^\infty \int_{\mathbb{R}^{n-1}} e^{-2\pi\imath x'\cdot\xi'} u(x',x_n) e^{-2\pi\imath x_n\xi_n}\,dx'\,dx_n \\
&= \int_0^\infty e^{-2\pi\imath x_n\xi_n} \widetilde{u}(\xi',x_n)\,dx_n \\
&= \widetilde{h}(\xi') \int_0^\infty e^{-(1+|\xi'|+2\pi\imath\xi_n} x_n)\,dx_n \\
&= \frac{\widetilde{h}(\xi')}{1+|\xi'|+2\pi\imath\xi_n}.
\end{aligned}$$

Now,

$$\begin{aligned}
\int_{\mathbb{R}^n} (1+|\xi'|^2)|\hat{u}(\xi)|^2\,d\xi &= \int_{\mathbb{R}^n} \frac{(1+|\xi'|)^2|\widetilde{h}(\xi')|^2}{(1+|\xi'|)^2+4\pi^2\xi_n^2}\,d\xi \\
&\leq \int_{\mathbb{R}^n} \frac{(1+|\xi'|)^2|\widetilde{h}(\xi')|^2}{1+|\xi'|^2+\xi_n^2}\,d\xi \\
&= \pi \int_{\mathbb{R}^{n-1}} (1+|\xi'|^2)^{\frac{1}{2}}|\widetilde{h}(\xi')|^2\,d\xi' < +\infty
\end{aligned}$$

since $h \in H^{\frac{1}{2}}(\mathbb{R}^{n-1})$. This proves that u (extended by zero) and $\frac{\partial u}{\partial x_i}$ are in $L^2(\mathbb{R}^n)$, for all $1 \leq i \leq n-1$. Hence, u, $\frac{\partial u}{\partial x_i}$ are in $L^2(\mathbb{R}^n_+)$, for all $1 \leq i \leq n-1$. For the case $i = n$, notice that by differentiating under the integral sign,

$$\widetilde{\left(\frac{\partial u}{\partial x_n}\right)}(\xi',x_n) = \frac{\partial \widetilde{u}}{\partial x_n}(\xi',x_n) = -(1+|\xi'|)\widetilde{u}\,(\xi',x_n).$$

Extend $\frac{\partial u}{\partial x_n}$ by zero outside $\mathbb{R}^n_+$. Then as before we get

$$\widehat{\left(\frac{\partial u}{\partial x_n}\right)}(\xi) = \frac{-(1+|\xi'|)\widetilde{h}\,(\xi')}{1+|\xi'|+2\pi\imath\xi_n}.$$

Then

$$\begin{aligned}
\int_{\mathbb{R}^n} \left|\widehat{\left(\frac{\partial u}{\partial x_n}\right)}(\xi)\right|^2 d\xi &\leq 2\int_{\mathbb{R}^n} \frac{(1+|\xi'|^2)|\widetilde{h}\,(\xi')|^2}{1+|\xi'|^2+\xi_n^2}\,d\xi \\
&= 2\pi \int_{\mathbb{R}^n} (1+|\xi'|^2)^{\frac{1}{2}}|\widetilde{h}\,(\xi')|^2\,d\xi' < \infty.
\end{aligned}$$

Thus $\frac{\partial u}{\partial x_n}$ (extended by zero) is in $L^2(\mathbb{R}^n)$ and so $\frac{\partial u}{\partial x_n} \in L^2(\mathbb{R}^n_+)$ and hence $u \in H^1(\mathbb{R}^n_+)$. Now (by the Fourier inversion formula)

$$\widetilde{u}(\xi', 0) = \widetilde{h}(\xi')$$

implies that $u(x', 0) = h(x')$ and so $\gamma(u) = h$. $\square$

Remark 2.9.5. If $v \in W^{2,p}(\Omega)$, by a similar argument one can give a meaning to $\frac{\partial v}{\partial \nu}$.Note that $\nabla v \in [W^{1,p}(\Omega)]^n$, and hence the trace of ∇v on $\partial\Omega$ belongs to $[L^p(\partial\Omega)]^n$. One defines

$$\frac{\partial v}{\partial \nu} := \gamma(\nabla v) \cdot \nu,$$

which belongs to $L^p(\partial\Omega)$. Indeed, one can show that

$$\frac{\partial v}{\partial \nu} \in W^{1-\frac{1}{p},p}(\partial\Omega).$$

For $p = 2$, for $v \in H^2(\Omega)$ we have $\frac{\partial v}{\partial \nu} \in H^{\frac{1}{2}}(\partial\Omega)$. One can also show that the operator $v \mapsto \{v|_{\partial\Omega}, \frac{\partial v}{\partial \nu}\}$ is linear continuous and onto from $W^{2,p}(\Omega)$ onto $W^{2-\frac{1}{p},p}(\partial\Omega) \times W^{1-\frac{1}{p},p}(\partial\Omega)$.

The relation $\gamma\left(W^{1,p}(\Omega)\right) \subsetneq L^{p^\sharp}(\partial\Omega)$ is the basis for defining the *trace spaces*:

$$W^{1-\frac{1}{p},p}(\partial\Omega) := \{\gamma(v) \in L^{p^\sharp}(\partial\Omega) \mid v \in W^{1,p}(\Omega)\} \quad \text{for } 1 \le p < n$$

which is the traces of all the functions in $W^{1,p}(\Omega)$, $1 \le p < n$.

Remark 2.9.6. By appropriate modifications one can easily prove that γ maps $H^k(\mathbb{R}^n_+)$ onto $H^{k-1/2}(\mathbb{R}^{n-1})$. In the same way, if $u \in H^2(\mathbb{R}^n_+)$ one can prove that $\frac{\partial u}{\partial x_n}(x', 0)$ is in $L^2(\mathbb{R}^{n-1})$ and is, in fact, in $H^{\frac{1}{2}}(\mathbb{R}^{n-1})$. Similarly, one can extend $-\frac{\partial u}{\partial x_n}(x', 0)$ to a bounded linear map $\gamma_1 : H^2(\mathbb{R}^n_+) \to L^2(\mathbb{R}^{n-1})$ whose range is $H^{\frac{1}{2}}(\mathbb{R}^{n-1})$. More generally, we have a collection of continuous linear maps $\{\gamma_j\}$ into $L^2(\mathbb{R}^{n-1})$ such that the map $\gamma = (\gamma_0, \gamma_1, \cdots, \gamma_{m-1})$ maps $H^k(\mathbb{R}^n_+)$ into $[L^2(\mathbb{R}^{n-1})]^k$ and the range in the space

$$\prod_{j=0}^{k-1} H^{k-j-1/2}(\mathbb{R}^{n-1}).$$

Definition 2.9.7. *We shall denote the range of the map T to be $W^{\frac{k}{2},p}(\partial\Omega)$. For $p = 2$, we denote $W^{\frac{k}{2},2}$ as $H^{\frac{k}{2}}(\partial\Omega)$.*

Thus, instead of defining the fractional power Sobolev spaces using Fourier transform, one can define them as range of trace operator γ.

Theorem 2.9.8. *$W^{\frac{k}{2},p}(\partial\Omega)$ is dense in $L^p(\partial\Omega)$.*

Theorem 2.9.9 (Trace zero)**.** *Let $\Omega \subset \mathbb{R}^n$ be bounded and $\partial\Omega$ is in C^k class. Then $u \in W_0^{k,p}(\Omega)$ iff $\gamma(u) = 0$ on $\partial\Omega$. In particular, ker$(\gamma) = W_0^{k,p}(\Omega)$.*

Proof. We first show the inclusion $W_0^{1,p}(\Omega) \subset \ker\gamma$. Take $v \in W_0^{1,p}(\Omega)$. By definition of $W_0^{1,p}(\Omega)$, there exists a sequence of functions $(v_n)_{n\in\mathbb{N}}$, $v_n \in \mathcal{D}(\Omega)$ such that $v_n \to v$ in $W^{1,p}(\Omega)$. Since $\gamma_0(v_n) = v_n|_{\partial\Omega} = 0$, by continuity of γ_0 we obtain that $\gamma_0(v) = 0$, i.e., $v \in \ker\gamma_0$.

The other inclusion is a bit more involved. We shall prove it in a sequence of lemma for $H^1(\mathbb{R}^n_+)$. The idea involved is following: Take $v \in W^{1,p}(\mathbb{R}^n_+)$ such that $\gamma_0(v) = 0$. Prove that $v \in W_0^{1,p}(\mathbb{R}^n_+)$. Let us first extend v by zero outside of $\mathbb{R}^n_+$. By using the information $\gamma_0(v) = 0$ one can verify that the so-obtained extension $\tilde{v}$ belongs to $W^{1,p}(\mathbb{R}^n)$. Then, let us translate $\tilde{v}$, and consider for any $h > 0$

$$\tau_h\tilde{v}(x', x_n) = \tilde{v}(x', x_n - h).$$

Finally, one regularizes by convolution the function $\tau_h\tilde{v}$. We have that for ε sufficiently small, $\rho_\varepsilon \star (\tau_h\tilde{v})$ belongs to $\mathcal{D}(\mathbb{R}^n_+)$ and $\rho_\varepsilon \star (\tau_h\tilde{v})$ tends to v in $W^{1,p}(\mathbb{R}^n_+)$ as $h \to 0$ and $\varepsilon \to 0$. Hence $v \in W_0^{1,p}(\mathbb{R}^n_+)$. □

Lemma 2.9.10 (Green's formula)**.** *. Let $u, v \in H^1(\mathbb{R}^n_+)$. Then*

$$\int_{\mathbb{R}^n_+} u\frac{\partial v}{\partial x_i} = -\int_{\mathbb{R}^n_+} \frac{\partial u}{\partial x_i}v \quad \text{if } 1 \le i \le (n-1) \tag{2.9.6}$$

$$\int_{\mathbb{R}^n_+} u\frac{\partial v}{\partial x_n} = -\int_{\mathbb{R}^n_+} \frac{\partial u}{\partial x_n}v - \int_{\mathbb{R}^{n-1}} \gamma_0(u)\gamma_0(v). \tag{2.9.7}$$

Proof. If $u, v \in \mathcal{D}(\mathbb{R}^n)$, then the relations (2.9.6) and (2.9.7) follow by integration by parts. The general case follows by the density of the restrictions of functions of $\mathcal{D}(\mathbb{R}^n)$ in $H^1(\mathbb{R}^n_+)$ and the continuity of the map $\gamma_0 : H^1(\mathbb{R}^n_+) \to L^2(\mathbb{R}^{n-1})$. □

Corollary 2.9.11. *If $u, v \in H^1(\mathbb{R}^n_+)$ and at least one of them is in ker (γ_0) then (2.9.6) holds for all $1 \leq i \leq n$.*

Lemma 2.9.12. *Let $v \in ker(\gamma_0)$. Then its extension by zero outside $\mathbb{R}^n_+$, denoted $\widetilde{v}$, is in $H^1(\mathbb{R}^n)$, and*

$$\frac{\partial \widetilde{v}}{\partial x_i} = \widetilde{\left(\frac{\partial v}{\partial x_i}\right)}, \quad 1 \leq i \leq n.$$

Proof. Let $\phi \in \mathcal{D}(\mathbb{R}^n)$. Then for $1 \leq i \leq n$,

$$\int_{\mathbb{R}^n} \widetilde{v}\frac{\partial \phi}{\partial x_i} = \int_{\mathbb{R}^n_+} v\frac{\partial \phi}{\partial x_n} = -\int_{\mathbb{R}^n_+} \frac{\partial v}{\partial x_i}\phi = \int_{\mathbb{R}^n} \widetilde{\left(\frac{\partial v}{\partial x_i}\right)}\phi$$

by the above corollary. Let $h > 0$ and consider $\bar{h} = he_n \in \mathbb{R}^n$ where e_n is the unit vector $(0, 0, ..., 0, 1)$. Consider the function $\tau_{\bar{h}}\widetilde{v}$, where $\widetilde{v}$ is the extension by zero outside $\mathbb{R}^n_+$ of $v \in ker(\gamma_0)$. Then $\tau_{\bar{h}}\widetilde{v}$ vanishes for all $x \in \mathbb{R}^n$ such that $x_n < h$. (Recall that $\tau_{\bar{h}}\widetilde{v}(x) = \widetilde{v}(x - \bar{h})$.) □

Lemma 2.9.13. *Let $1 \leq p < \infty$ and $\bar{h} \in \mathbb{R}^n$. Then if $f \in L^p(\mathbb{R}^n)$*

$$\lim_{\bar{h} \to 0} |\tau_{\bar{h}} f - f|_{0,p,\mathbb{R}^n} = 0$$

Proof. By the translation invariance of the Lebesgue measure, $\tau_{\bar{h}} f \in L^p(\mathbb{R}^n)$ as well. Let $\epsilon > 0$ be given and choose $\phi \in \mathcal{D}(\mathbb{R}^n)$ such that

$$|f - \phi|_{0,p,\mathbb{R}^n} < \frac{\epsilon}{3}. \tag{2.9.8}$$

Let $a > 0$ such that supp $(\phi) \subset [-a, a]^n$. Since ϕ is uniformly continuous, there exists $\delta < 0$ be given and choose $\phi \in \mathcal{D}(\mathbb{R}^n)$ such that

$$|\phi(x - \bar{h}) - \phi(x)| < \frac{\epsilon}{3}(2(a+1))^{-\frac{n}{p}}.$$

Then

$$\int_{\mathbb{R}^n} |\phi(x - \bar{h}) - \phi(x)|^p \, dx = \int_{[-(a+1),(a+1)]^n} |\phi(x - \bar{h}) - \phi(x)|^p dx < \left(\frac{\epsilon}{3}\right)^p.$$

Thus for $|\bar{h}| < \delta$,

$$|\tau_{\bar{h}}\phi - \phi|_{0,p,\mathbb{R}^n} < \frac{\epsilon}{3}. \tag{2.9.9}$$

Finally, again by the translation invariance of the Lebesgue measure, we have

$$|\tau_{\bar{h}} f - \tau_{\bar{h}} \phi|_{0,p,\mathbb{R}^n} = f - \phi|_{0,p,\mathbb{R}^n} < \frac{\epsilon}{3}. \tag{2.9.10}$$

The result now follows on combining (2.9.8), (2.9.9) and (2.9.10) by the triangle inequality. $\square$

Corollary 2.9.14. *If* $v \in H^1(\mathbb{R}^n)$ *then*

$$\lim_{h \to 0} \|\tau_{\bar{h}} v - v\|_{1,\mathbb{R}^n} = 0.$$

Proof. Clearly by the preceding lemma $\tau_{\bar{h}} v \to v$ in $L^2(\mathbb{R}^n)$. Also it is easy to check that for any $1 \leq i \leq n$,

$$\frac{\partial}{\partial x_i}(\tau_{\bar{h}} v) = \tau_{\bar{h}} \frac{\partial v}{\partial x_i}.$$

Thus again by the lemma, $\frac{\partial(\tau_{\bar{h}} v)}{\partial x_i} \to \frac{\partial v}{\partial x_i}$ in $L^2(\mathbb{R}^n)$. $\square$

Theorem 2.9.15. *ker* $(\gamma_0) = H_0^1(\mathbb{R}^n_+)$.

Proof. We already have seen that $H_0^1(\mathbb{R}^n_+) \subset \ker(\gamma_0)$. Let now $v \in \ker(\gamma_0)$. Then we have seen that its extension $\widetilde{v}$ by zero is in $H^1(\mathbb{R}^n)$. Using the cut-off functions $\{\zeta_k\}$ (cf. Theorem 2.1.3) we have that $\zeta_k \widetilde{v} \to \widetilde{v}$ as $k \to \infty$ in $H^1(\mathbb{R}^n)$. The functions $\zeta_k \widetilde{v}$ have compact support in $\mathbb{R}^n$ and vanish for $x_n < 0$. Now fix such a k so that

$$\|\widetilde{v} - \zeta_k \widetilde{v}\|_{1,\mathbb{R}^n} < \eta,$$

where $\eta > 0$ is a given positive quantity. Again we can choose h small enough so that if $\bar{h} = h e_n$, then

$$\|\tau_{\bar{h}}(\zeta_k \widetilde{v}) - \zeta_k \widetilde{v}\|_{1,\mathbb{R}^n} < \eta.$$

Now $\tau_{\bar{h}}(\zeta_k \widetilde{v})$ has compact support in $\mathbb{R}^n_+$ and vanishes for all $x \in \mathbb{R}^n$ with $x_n < h$. Let $\{\rho_\epsilon\}$ be the family of mollifers. If $\epsilon > 0$ is chosen small enough then $\rho_\epsilon * \tau_{\bar{h}}(\zeta_k \widetilde{v})$ will have support contained in the set

$$B(0;\epsilon) + K \cap \{x | x_n \geq h > 0\}$$

where $K = \text{supp}(\tau_{\bar{h}}(\zeta_k \widetilde{v}))$ is compact. Thus

$$\rho_\epsilon * \tau_{\bar{h}} \widetilde{(\zeta_k)} \in \mathcal{D}(\mathbb{R}^n_+)$$

and we know that as $\epsilon \downarrow 0$, $\rho_{\epsilon * \tau_{\bar{h}}}(\zeta_k \widetilde{v}) \to \tau_{\bar{h}}(\zeta_k \widetilde{v})$. Thus we can choose ϵ small enough such that

$$\|\rho_\epsilon * \tau_{\bar{h}}(\zeta_k \widetilde{v}) - \tau_{\bar{h}}(\zeta_k \widetilde{v})\|_{1,\mathbb{R}^n} < \eta.$$

Thus we have found a function $\phi_\eta \in \mathcal{D}(\mathbb{R}^n_+)$,

$$\phi_\eta = \rho_\epsilon * \tau_{\bar{h}}(\zeta_k \widetilde{v})$$

such that

$$\|\phi_\eta - v\|_{1,\mathbb{R}^n_+} \leq \|\phi_\eta - \widetilde{v}\|_{1,\mathbb{R}^n} < 3_\eta.$$

Thus, as η is arbitrary, it follows that

$$\ker(\gamma_0) \subset \overline{\mathcal{D}(\mathbb{R}^n_+)} = H^1_0(\mathbb{R}^n_+).$$

$\square$

Similarly, it can be proved that if $\gamma = (\gamma_0, \gamma_1, \cdots, \gamma_{m-1})$, then the kernel of γ in $H^m(\mathbb{R}^n_+)$ is precisely the set $H^1_0(\mathbb{R}^n_+)$.

Theorem 2.9.16. *Let $1 \leq p < \infty$. Then*

$$W^{2,p}_0(\Omega) = \left\{ u \in W^{2,p}(\Omega) \mid \gamma(u) = 0 \text{ and } \sum_{i=1}^{n} \nu_i \gamma(\frac{\partial u}{\partial x_i}) = 0 \right\},$$

where $(\nu_i)_{i=1}^n$ denotes the unit outer normal vector field along $\partial\Omega$.

Theorem 2.9.17 (Trace Theorem)**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set of class C^{m+1} with boundary $\partial\Omega$. Then there exists a trace map $\gamma = (\gamma_0, \gamma_1, \cdots, \gamma_{m-1})$ from $H^m(\Omega)$ into $(L^2(\Omega))^m$ such that*

(i) If $v \in C^\infty(\bar{\Omega})$, Then $\gamma_0(v) = v|_{\partial\Omega}$, $\gamma_1(v) = \frac{\partial v}{\partial \nu}|_{\partial\Omega}, \cdots$, and $\gamma_{m-1}(v) = \frac{\partial^{m-1}}{\partial \nu^{m-1}}(v)|_{\partial\Omega}$, where ν is the unit exterior normal to the boundary $\partial\Omega$

(ii) The range of γ is the space

$$\prod_{j=0}^{m-1} H^{m-j-\frac{1}{2}}(\partial\Omega).$$

(iii) The kernel of γ is $H^m_0(\Omega)$

Proof. Let us now turn to the case of a bounded open set Ω of class C^1. Let $\{U_j, T_j\}_{j=1}^k$ be an associated local chart for the boundary $\partial\Omega$ and let $\{\psi_j\}_{j=1}^k$ be a partition of unity subordinate to the cover $\{U_j\}$ of $\partial\Omega$. If $u \in H^1(\Omega)$, then $(\psi_j u|_{U_j\cap}\Omega)o\, T_j \in H^1(\mathbb{R}^n_+)$ and so we can define its trace as an element of $H^{\frac{1}{2}}(\mathbb{R}^{n-1})$. Coming back by T_j^{-1} we can define the trace on $U_j \bigcap \partial\Omega$. Piecing these together we get the trace $\gamma_0 u$ in $L^2(\partial\Omega)$ and the image (by definition of the spaces) will be precisely $H^{\frac{1}{2}}(\partial\Omega)$. Similarly if the boundary is smoother we can define the higher order traces γ_j. $\square$

Now that we have given a meaning to the functions restricted to the boundary of the domain, we intend to generalise the classical Green's identities (cf. Appendix **??**) to $H^1(\Omega)$. The trace theorem above helps us to obtain Green's theorem for functions in $H^1(\Omega)$, Ω of class C^1. If $\nu(x)$ denotes the unit outer-normal vector on the boundary $\partial\Omega$ (which is defined uniquely a.e. on $\partial\Omega$), we denote its components along the coordinate axes by $\nu_i(x)$. Thus we write generically,

$$\nu = (\nu_1, \cdots, \nu_n).$$

For example, if $\Omega = B(0; 1)$ then $\nu(x) = x$ for all $|x| = 1$. Thus $\nu_i(x) = x_i$ in this case. If $\Omega = B(0; R)$, then $\nu(x) = \frac{x}{R}$. If Ω has a part of its boundary, say, $x_n = 0$, then the unit outer normal is $\pm e_n$ depending on the side on which Ω lies.

Theorem 2.9.18 (Green's Theorem)**.** . *Let Ω be a bounded open set of $\mathbb{R}^n$ set of class C^1 lying on the same side of its boundary $\partial\Omega$. Let $u, v \in H^1(\Omega)$. Then for $1 \le i \le n$,*

$$\int_\Omega u\,\frac{\partial v}{\partial x_i} = -\int_\Omega \frac{\partial u}{\partial x_i} v + \int_{\partial\Omega} (\gamma_0 u)\,(\gamma_0 v)\,\nu_i. \tag{2.9.11}$$

Proof. Recall that $C^\infty(\bar{\Omega})$ is dense in $H^1(\Omega)$. If $u_m,\ v_m \in C^\infty(\bar{\Omega})$ then we have by the classical Green's Theorem

$$\int_\Omega u_m\,\frac{\partial v_m}{\partial x_i} = -\int_\Omega \frac{\partial u_m}{\partial x_i}\,v_m + \int_{\partial\Omega} u_m\,v_m\,\nu_i$$

and choosing $u_m \to u, v_m \to v$ in $H^1(\Omega)$ we deduce (2.9.11) by the continuity of the trace map γ_0. $\square$

Usually, we rewrite $\gamma_0(v)$ as just v on $\partial\Omega$ and understand it as the trace of v on $\partial\Omega$. Similarly if $\gamma_1 v$ is defined we will write it as $\frac{\partial v}{\partial \nu}$.

Theorem 2.9.19. *Let Ω be a bounded open subset of $\mathbb{R}^n$ with a C^1 boundary. Given a vector field $V = (v_1, \ldots, v_n)$ on Ω such that $v_i \in H^1(\Omega)$ for all $1 \le i \le n$, then*

$$\int_\Omega \nabla \cdot V \, dx = \int_{\partial\Omega} V \cdot \nu \, d\sigma.$$

Proof. Setting $u \equiv 1$ and $v = v_i \in H^1(\Omega)$, we get from (2.9.11)

$$\int_\Omega \frac{\partial v_i}{\partial x_i} = \int_\Omega v_i \, \nu_i.$$

and if $V = (v_i) \in (H^1(\Omega))^n$, we get on summing with respect to i,

$$\int_\Omega \text{div } (V) = \int_{\partial\Omega} V \cdot \nu \, d\sigma$$

which is the Gauss Divergence Theorem. □

Corollary 2.9.20. *Let Ω be a bounded open subset of $\mathbb{R}^n$ with a C^1 boundary. Let $u, v \in H^2(\Omega)$, then*

(i)

$$\int_\Omega (v\Delta u + \nabla v \cdot \nabla u) \, dx = \int_{\partial\Omega} v \frac{\partial u}{\partial \nu} \, d\sigma,$$

where $\frac{\partial u}{\partial \nu} := \nabla u \cdot \nu$ and $\Delta := \nabla \cdot \nabla$.

(ii)

$$\int_\Omega (v\Delta u - u\Delta v) \, dx = \int_{\partial\Omega} \left(v\frac{\partial u}{\partial \nu} - u\frac{\partial v}{\partial \nu} \right) \, d\sigma.$$

Proof. If we have $u \in H^2(\Omega)$ and use $\frac{\partial u}{\partial x_i}$ in place of u in (2.9.11), we get

$$\int_\Omega \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i} = -\int_\Omega \frac{\partial^2 u}{\partial x_i^2} v + \int_{\partial\Omega} \frac{\partial u}{\partial x_i} v\nu_i.$$

If u were smooth then $\sum_{i=1}^n \frac{\partial u}{\partial x_i}\nu_i = \frac{\partial u}{\partial \nu}$. Thus by continuity of the trace γ_1, we get, for $u \in H^2(\Omega)$, $v \in H^1(\Omega)$,

$$\int_\Omega \nabla u \cdot \nabla v = -\int_\Omega (\Delta u) \, v + \int_{\partial\Omega} v \, \frac{\partial u}{\partial \nu}.$$

□

Remark 2.9.21. The Green's formula holds for $u \in W^{1,p}(\Omega)$ and $a \in W^{1,q}(\Omega)$ with $\frac{1}{p} + \frac{1}{q} = 1$

2.10 Further reading

For a detailed study of Sobolev spaces, we refer to [Ada75, Maz85, Kes89].

Chapter 3

Bounded Variation Functions

Let $\Phi : \Omega \to \mathbb{R}^n$ and $\Phi = (\Phi_1, \ldots, \Phi_n)$ be the coordinates where $\Phi_i : \Omega \to \mathbb{R}$ for all i. Recall that the *divergence* of any $\Phi \in C^1(\Omega; \mathbb{R}^n)$ is defined as $\operatorname{div}(\Phi) = \nabla \cdot \Phi = \sum_{i=1}^m \frac{\partial \Phi_i}{\partial x_i}$. Also, the uniform norm of Φ is given as

$$\|\Phi\|_\infty = \sup_{x \in \Omega} \left(\sum_{i=1}^n \Phi_i^2(x) \right)^{1/2}.$$

Recall that any $u \in L^1_{\text{loc}}(\Omega)$ is a distribution and admits a distributional derivative.

Definition 3.0.1. *The total variation of $u \in L^1_{loc}(\Omega)$ is defined as*

$$V(u, \Omega) = \sup_{\substack{\Phi \in \mathcal{D}(\Omega; \mathbb{R}^n) \\ \|\Phi\|_\infty \le 1}} \left\{ \int_\Omega u(\nabla \cdot \Phi)\, dx \right\}$$

Definition 3.0.2. *For any open subset $\Omega \subset \mathbb{R}^n$, we define the space of bounded variations $BV(\Omega) := \{u \in L^1(\Omega) \mid V(u, \Omega) < \infty\}$ and the norm on it is given as*

$$\|u\|_{BV(\Omega)} = \|u\|_1 + V(u, \Omega).$$

Theorem 3.0.3. *The normed space $BV(\Omega)$ is complete.*

Definition 3.0.4. *We define the metric of* strict convergence *, d, on $BV(\Omega)$ as*

$$d(u, v) = \|u - v\|_1 + |V(u, \Omega) - V(v, \Omega)|.$$

Theorem 3.0.5. *Let Ω be an open subset of $\mathbb{R}^n$. Given any $u \in BV(\Omega)$, there exist functions $\{u_k\}_{k=1}^{\infty} \subset BV(\Omega) \cap C_c^{\infty}(\Omega)$ such that*

1. *$u_k \to u$ in $L^1(\Omega)$ and*
2. *$V(u_k, \Omega) \to V(u, \Omega)$ as $k \to \infty$*

Theorem 3.0.6. *Let Ω be an open subset of $\mathbb{R}^n$. For any function $u \in C_c^{\infty}(\Omega)$, the total variation of u satisfies the identity $V(u, \Omega) = \|\nabla u\|_{1,\Omega}$.*

Corollary 3.0.7. *Let $u \in W_{loc}^{1,1}(\Omega)$. Then $\nabla u \in L^1(\Omega; \mathbb{R}^n)$ iff $V(u, \Omega) < \infty$. Further, in that case, $\|\nabla u\|_{1,\Omega} = V(u, \Omega)$.*

Theorem 3.0.8 (Sobolev inequality for BV)**.** *There exists a constant $C > 0$ such that*

$$\|u\|_{\frac{n}{n-1},\mathbb{R}^n} \leq CV(u, \mathbb{R}^n) \quad \forall u \in BV(\mathbb{R}^n).$$

Proof. Choose $u_k \in C_c^{\infty}(\mathbb{R}^n)$ so that $u_k \to u$ in $L^1(\mathbb{R}^n)$ and $V(u_k, \mathbb{R}^n) \to V(u, \mathbb{R}^n)$ as $k \to \infty$. Applying Fatou's lemma on $\{|u_k|^{\frac{n}{n-1}}\}$, we get

$$\int_{\mathbb{R}^n} |u|^{\frac{n}{n-1}}\, dx \leq \liminf_{k\to\infty} \int_{\mathbb{R}^n} |u_k|^{\frac{n}{n-1}}\, dx.$$

Thus, $\|u\|_{\frac{n}{n-1},\mathbb{R}^n} \leq \liminf_{k\to\infty} \|u_k\|_{\frac{n}{n-1},\mathbb{R}^n}$. Since $u_k \in W^{1,1}(\mathbb{R}^n)$, applying GNS inequality for $p = 1$, we obtain

$$\|f\|_{\frac{n}{n-1},\mathbb{R}^n} \leq \lim_{k\to\infty} C\|\nabla u_k\|_{1,\mathbb{R}^n}.$$

But $\|\nabla u_k\|_{1,\mathbb{R}^n} = V(u_k, \mathbb{R}^n)$, for all k. Hence,

$$\|u\|_{\frac{n}{n-1},\mathbb{R}^n} \leq \lim_{k\to\infty} CV(u_k, \mathbb{R}^n) = CV(u, \mathbb{R}^n).$$

□

3.1 De Giorgi Perimeter

Theorem 3.1.1. *Let Ω be an open subset of $\mathbb{R}^n$ with C^1 boundary and the boundary is bounded. Then*

$$\int_{\partial\Omega} d\sigma = V(\chi_\Omega, \mathbb{R}^n).$$

Definition 3.1.2. *Let E be a measurable subset of Ω. The* De Giorgi perimeter *of E is defined as $P_\Omega(E) = V(\chi_E, \Omega)$. The set E is said to have* finite perimeter *in Ω if $\chi_E \in BV(\Omega)$.*

Corollary 3.1.3 (Isoperimetric Inequality)**.** *For any bounded set $E \subset \mathbb{R}^n$ of finite perimeter, we have*

$$|E|^{1-1/n} \leq CP_{\mathbb{R}^n}(E).$$

Proof. Let $u = \chi_E$. Then $u \in \mathrm{BV}(\mathbb{R}^n)$ and hence $\|u\|_{\frac{n}{n-1},\mathbb{R}^n} \leq CV(u, \mathbb{R}^n)$. But $\|u\|_{\frac{n}{n-1},\mathbb{R}^n} = |E|^{1-1/n}$ and $V(u, \mathbb{R}^n) = P_{\mathbb{R}^n}(E)$. □

Theorem 3.1.4 (Morse-Sard)**.** *Let Ω be an open subset of $\mathbb{R}^n$ and $u \in C^\infty(\Omega)$. Then the set*

$$\{t \in \mathbb{R} \mid u(x) = t \text{ and } \nabla u(x) = 0\}$$

has Lebesgue measure zero.

Theorem 3.1.5 (Co-area Formula)**.** *Let $u \in \mathcal{D}(\mathbb{R}^n)$ and $f \in C^1(\mathbb{R}^n)$. Then*

$$\int_{\mathbb{R}^n} f|\nabla u|\,dx = \int_0^\infty dt \int_{\{|u|=t\}} f d\sigma.$$

Moreover, for every open subset $\Omega \subset \mathbb{R}^n$,

$$\int_\Omega |\nabla u|\,dx = \int_0^\infty dt \int_{\{|u|=t\}} \chi_\Omega d\sigma.$$

Chapter 4

Second Order Elliptic PDE

4.1 Homogeneous Dirichlet Problem

Let Ω be an open subset of $\mathbb{R}^n$ and let $\partial\Omega$ be the topological boundary of Ω. Recall the homogeneous Dirichlet problem, i.e., given $f : \Omega \to \mathbb{R}$, find $u : \overline{\Omega} \to \mathbb{R}$ such that

$$\begin{cases} -\Delta u &= f \quad \text{in } \Omega \\ \quad u &= 0 \quad \text{on } \partial\Omega. \end{cases} \tag{4.1.1}$$

Note that the boundary condition imposed above is referred to as the Dirichlet boundary condition. The minus ($-$) sign before Laplacian has physical motivation.

Definition 4.1.1. *We say $u : \overline{\Omega} \to \mathbb{R}$ is a* classical *solution of the Dirichlet problem (4.1.1), if $u \in C^2(\overline{\Omega})$ and satisfies (4.1.1) pointwise for each $x \in \overline{\Omega}$.*

Let u be a classical solution of (4.1.1) and suppose Ω is bounded with C^1-smooth boundary (cf. Corollary **??**). Then, multiplying any $\phi \in C^1(\Omega)$ and integrating both sides in (4.1.1), we get

$$\begin{aligned} -\int_\Omega (\Delta u)\phi\,dx &= \int_\Omega f\phi\,dx \\ \int_\Omega \nabla\phi\cdot\nabla u\,dx - \int_{\partial\Omega}\phi\frac{\partial u}{\partial\nu}\,d\sigma &= \int_\Omega f\phi\,dx \\ &\quad \text{(By Green's identity)} \end{aligned}$$

Now, if we impose the Dirichlet boundary condition on u, we get

$$\int_\Omega \nabla\phi\cdot\nabla u\,dx = \int_\Omega f\phi\,dx,$$

since $u = 0$ implies that $\frac{\partial u}{\partial \nu} = 0$ on $\partial\Omega$. Thus, any classical solution u of (4.1.1) solves the following formulation of the problem: for any given f, find u such that

$$\int_\Omega \nabla u \cdot \nabla \phi \, dx = \int_\Omega f\phi \, dx \quad \forall \phi \in C^1(\Omega),$$

as long as the integrals make sense. Note that, in the above problem, it is enough for u to be in $C^1(\Omega)$. Also, in particular, any classical solution satisfies the identity

$$\int_\Omega |\nabla u|^2 \, dx = \int_\Omega fu \, dx.$$

Since $C^1(\Omega)$ which vanishes on boundary is dense in $H^1_0(\Omega)$, we have the classical solution u will also solve

$$\int_\Omega \nabla u \cdot \nabla v \, dx = \int_\Omega fv \, dx \quad \forall v \in H^1_0(\Omega),$$

as long as the integrals make sense. The integrals will make sense if $f \in L^2(\Omega)$. In fact, more generally, if $f \in H^{-1}(\Omega)$ then the integral on the right side can be replaced with the duality product. The arguments above motivates the notion of weak solution

Definition 4.1.2. *Given $f \in H^{-1}(\Omega)$, we say $u \in H^1_0(\Omega)$ is a* weak solution *of (4.1.1) if*

$$\int_\Omega \nabla u \cdot \nabla v \, dx = \langle f, v\rangle_{H^{-1}(\Omega), H^1_0(\Omega)}, \quad \forall v \in H^1_0(\Omega). \tag{4.1.2}$$

We have already observed that every classical solution is a weak solution for a bounded open set Ω with C^1-smooth boundary.

Theorem 4.1.3. *Let Ω be a bounded open subset of $\mathbb{R}^n$ and $f \in H^{-1}(\Omega)$. Then there is a unique weak solution $u \in H^1_0(\Omega)$ satisfying (4.1.2). Moreover, u minimizes the functional $J : H^1_0(\Omega) \to \mathbb{R}$ defined as,*

$$J(v) := \frac{1}{2}\int_\Omega |\nabla v|^2 \, dx - \langle f, v\rangle_{H^{-1}(\Omega), H^1_0(\Omega)}$$

in $H^1_0(\Omega)$.

Proof. Recall that $H^1(\Omega)$ is a Hilbert space endowed with the inner-product

$$(v, w) := \int_\Omega vw\,dx + \int_\Omega \nabla v \nabla w\,dx.$$

Since Ω is bounded, by Poincaré inequality (cf. (2.8.4)), there is a constant $C > 0$ such that

$$\|v\|_2 \leq C\|\nabla v\|_2 \quad \forall v \in H^1_0(\Omega).$$

Thus, $H^1_0(\Omega)$ is a Hilbert space endowed with the equivalent norm $\|v\|_{H^1_0(\Omega)} := \|\nabla v\|_2$ and the corresponding inner product is given as

$$((v, w)) := \int_\Omega \nabla v \cdot \nabla w\,dx.$$

We shall use the Lax-Milgram result (cf. Theorem **??**). We define the map $a : H^1_0(\Omega) \times H^1_0(\Omega) :\to \mathbb{R}$ as

$$a(v, w) := ((v, w)) = \int_\Omega \nabla v \cdot \nabla w\,dx.$$

Note that a is a bilinear. We need to show that a is continuous and coercive. Consider,

$$\begin{aligned}
|a(v, w)| &\leq \int_\Omega |\nabla v||\nabla w|\,dx \\
&\leq \left(\int_\Omega |\nabla v|^2\,dx\right)^{1/2} \left(\int_\Omega |\nabla w|^2\,dx\right)^{1/2} \\
&= \|v\|_{H^1_0(\Omega)}\|w\|_{H^1_0(\Omega)}.
\end{aligned}$$

Thus, a is a bilinear continuous form. Now, consider

$$a(v, v) = \int_\Omega |\nabla v|^2\,dx = \|\nabla v\|_2^2.$$

Hence a is coercive, bilinear continuous form. By Lax-milgram theorem, there is a $u \in H^1_0(\Omega)$ such that $a(u, v) = \langle f, v\rangle_{H^{-1}(\Omega), H^1_0(\Omega)}$. This is equivalent to (4.1.2) and also u minimizes the functional

$$= \frac{1}{2}\int_\Omega |\nabla v|^2\,dx - \langle f, v\rangle_{H^{-1}(\Omega), H^1_0(\Omega)}.$$

□

We have the shown the existence of weak solution to the homogeneous Dirichlet problem under the assumption that Ω is bounded. We now turn our attention to the inhomogeneous Dirichlet problem.

4.2 Inhomogeneous Dirichlet problem

Let Ω be an open subset of $\mathbb{R}^n$ and let $\partial\Omega$ be the topological boundary of Ω. The inhomogeneous Dirichlet problem is the following: Given $f : \Omega \to \mathbb{R}$ and $g : \partial\Omega \to \mathbb{R}$, find $u : \overline{\Omega} \to \mathbb{R}$ such that

$$\begin{cases} -\Delta u &= f \quad \text{in } \Omega \\ \quad u &= g \quad \text{on } \partial\Omega. \end{cases} \tag{4.2.1}$$

Definition 4.2.1. *We say $u : \overline{\Omega} \to \mathbb{R}$ is a* classical *solution of* (4.2.1)*, if $u \in C^2(\overline{\Omega})$ and satisfies* (4.2.1) *pointwise for each $x \in \overline{\Omega}$.*

We wish to give a weak notion of solution to the inhomogeneous Dirichlet problem. It follows from trace theorem (cf. Theorem 2.9.3) that for $u = g$ on $\partial\Omega$ to make sense for any $u \in H^1(\Omega)$, g should be in $H^{1/2}(\partial\Omega)$. Conversely, for any $g \in H^{1/2}(\partial\Omega)$, there is a $v_g \in H^1(\Omega)$ such that $Tv_g = g$. However, note that the choice of v_g is not unique.

Let $f \in L^2(\Omega)$ and $g \in H^{1/2}(\partial\Omega)$. For any given $g \in H^{1/2}(\partial\Omega)$ there is a $v_g \in H^1(\Omega)$ such that $Tv_g = g$. Let

$$K_g := \{v \in H^1(\Omega) \mid v - v_g \in H^1_0(\Omega)\}.$$

Exercise 67. Show that K_g is a closed convex non-empty subset of $H^1(\Omega)$ and $K_g = v_g + H^1_0(\Omega)$. (Hint: Use the linearity and continuity of the trace map T).

Note that if u is a classical solution of (4.2.1) and $g \in H^{1/2}(\partial\Omega)$, then $w := u - v_g$ is a weak solution of the homogeneous Dirichlet problem

$$\begin{cases} -\Delta w &= f + \Delta v_g \quad \text{in } \Omega \\ \quad w &= 0 \quad\quad\quad\; \text{on } \partial\Omega. \end{cases} \tag{4.2.2}$$

Definition 4.2.2. *Given $f \in H^{-1}(\Omega)$ and $g \in H^{1/2}(\partial\Omega)$, we say $u \in K_g$ is a* weak solution *of* (4.2.1) *if $w := u - v_g \in H^1_0(\Omega)$ is the weak solution of the homogeneous Dirichlet problem* (4.2.2)*, i.e., $u \in K_g$ is such that*

$$\int_\Omega \nabla u \cdot \nabla v \, dx = \langle f, v\rangle_{H^{-1}(\Omega), H^1_0(\Omega)} \, dx \quad \forall v \in H^1_0(\Omega). \tag{4.2.3}$$

Theorem 4.2.3. *Let Ω be a bounded open subset of $\mathbb{R}^n$, $g \in H^{1/2}(\partial\Omega)$ and $f \in H^{-1}(\Omega)$. Then there is a unique weak solution $u \in K_g$ of* (4.2.1)*.*

Moreover, if $f \in L^2(\Omega)$ then u minimizes the functional $J : K_g \to \mathbb{R}$ defined as,

$$J(v) := \frac{1}{2}\int_\Omega |\nabla v|^2\,dx - \int_\Omega fv\,dx$$

in K_g.

Proof. For any given $g \in H^{1/2}(\Omega)$, there is a $v_g \in H^1(\Omega)$. Note that the choice of v_g is not unique. We have already seen (cf. Theorem 4.1.3) the existence of weak solution for the homogeneous Dirichlet problem. Thus, there is a unique $w \in H^1_0(\Omega)$ satisfying (4.2.2). Then $u := w + v_g$ is a weak solution of (4.2.1). However, the uniqueness of u does not follow from the uniqueness of w, since the choice of v_g is not unique. If $u_1, u_2 \in K_g$ are two weak solutions of (4.2.1) then, by (4.2.3),

$$\int_\Omega \nabla(u_1 - u_2)\cdot\nabla v\,dx = 0 \quad \forall v \in H^1_0(\Omega).$$

In particular, for $v = u_1 - u_2$, we have $\|\nabla(u_1 - u_2)\|_2^2 = 0$. But, by Poincaré inequality, $\|u_1 - u_2\|_2^2 \le C\|\nabla(u_1 - u_2)\|_2^2 = 0$ implies that $u_1 = u_2$.

It now only remains to show that u minimizes J in K_g. Since K_g is a closed convex subset of $H^1(\Omega)$, by Theorem **??**, there is a unique $u^\star \in K_g$ such that

$$a(u^\star, v - u^\star) \ge \int_\Omega f(v - u^\star)\,dx, \quad \forall v \in K_g,$$

where $a(v, w) := \int_\Omega \nabla v \cdot \nabla w\,dx$. But $v \in K_g$ iff $v - u^\star \in H^1_0(\Omega)$. Thus, $a(u^\star, v) \ge \int_\Omega fv\,dx$ for all $v \in H^1_0(\Omega)$. Since $H^1_0(\Omega)$ is a subspace of $H^1(\Omega)$, we have $a(u^\star, v) = \int_\Omega fv\,dx$ for all $v \in H^1_0(\Omega)$. Thus, $u^\star$ satisfies (4.2.3) which by uniqueness of $u \in K_g$ we have $u^\star = u$. □

4.3 Linear Elliptic Operators

In this section, we introduce the generalization of the Laplace operator, called the *elliptic operators*. The elliptic operators come in two forms, divergence and non-divergence form, and we shall see that the notion of weak solution can be carried over to elliptic operator in divergence form.

Let Ω be an open subset of $\mathbb{R}^n$. Let $A = A(x) = (a_{ij}(x))$ be any given $n \times n$ matrix of functions, for $1 \le i, j \le n$. Let $\boldsymbol{b} = \boldsymbol{b}(x) = (b_i(x))$ be any given n-tuple of functions and let $c = c(x)$ be any given function.

Definition 4.3.1. *A second order operator L is said to be in* divergence form*, if L acting on some u has the form*

$$Lu := -\nabla \cdot (A(x)\nabla u) + \boldsymbol{b}(x) \cdot \nabla u + c(x)u.$$

Equivalently,

$$Lu := -\operatorname{div}(A(x)\nabla u) + \boldsymbol{b}(x) \cdot \nabla u + c(x)u.$$

(Hence the name divergence form). On the other hand, a second order operator L is said to be in non-divergence form*, if L acting some u has the form*

$$Lu := -\sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \boldsymbol{b}(x) \cdot \nabla u + c(x)u.$$

Observe that the operator L will make sense in the divergence form only if $a_{ij}(x) \in C^1(\Omega)$. Thus, if $a_{ij}(x) \in C^1$, then a divergence form equation can be rewritten in to a non-divergence form because

$$\nabla \cdot (A(x)\nabla u) = \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j} + \left(\sum_{j=1}^{n} \frac{\partial a_{ij}}{\partial x_j} \right)_i \cdot \nabla u.$$

Now, by setting $\widetilde{b}_i(x) = b_i(x) - \sum_{j=1}^{n} \frac{\partial a_{ij}}{\partial x_j}$, we have written a divergence L in non-divergence form. Also, note that for $A(x) = I$, $\boldsymbol{b}(x) = 0$ and $c(x) = 0$, $L = -\Delta$ is the usual Laplacian.

Definition 4.3.2. *We say a second order operator L is* elliptic *or* coercive *if there is a positive constant $\alpha > 0$ such that*

$$\alpha |\xi|^2 \le A(x)\xi.\xi \quad \text{a.e. in } x, \quad \forall \xi = (\xi_i) \in \mathbb{R}^n.$$

The second order operator L is said to be degenerate elliptic *if*

$$0 \le A(x)\xi.\xi \quad \text{a.e. in } x, \quad \forall \xi = (\xi_i) \in \mathbb{R}^n.$$

We shall now extend the notion of weak solution introduced for the Laplacian operator to a general second order elliptic equation in divergence form. We remark that for the integrals in the defintion of weak solution to make sense, the minimum hypotheses on $A(x), \boldsymbol{b}$ and c is that $a_{ij}, b_i, c \in L^\infty(\Omega)$.

Definition 4.3.3. *Let $a_{ij}, b_i, c \in L^\infty(\Omega)$ and let $f \in H^{-1}(\Omega)$, we say $u \in H^1_0(\Omega)$ is a* weak solution *of the homogeneous Dirichlet problem*

$$\begin{cases} -\mathrm{div}(A(x)\nabla u) + \boldsymbol{b}(x)\cdot\nabla u + c(x)u &= f \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega \end{cases} \tag{4.3.1}$$

whenever

$$\int_\Omega A\nabla u\cdot\nabla v\,dx + \int_\Omega (\boldsymbol{b}\cdot\nabla u)v\,dx + \int_\Omega cuv\,dx = \langle f, v\rangle_{H^{-1}(\Omega),H^1_0(\Omega)}, \quad \forall v \in H^1_0(\Omega). \tag{4.3.2}$$

We define the map $a(\cdot,\cdot) : H^1_0(\Omega)\times H^1_0(\Omega) \to \mathbb{R}$ as

$$a(v,w) := \int_\Omega A\nabla v\cdot\nabla w\,dx + \int_\Omega (\boldsymbol{b}\cdot\nabla v)w\,dx + \int_\Omega cvw\,dx.$$

It is easy to see that $a(\cdot,\cdot)$ is bilinear.

Lemma 4.3.4. *If $a_{ij}, b_i, c \in L^\infty(\Omega)$ then the bilinear map $a(\cdot,\cdot)$ is continuous on $H^1_0(\Omega)\times H^1_0(\Omega)$, i.e., there is a constant $c_1 > 0$ such that*

$$|a(v,w)| \le c_1\|v\|_{H^1_0(\Omega)}\|w\|_{H^1_0(\Omega)}.$$

Also, in addition, if Ω is bounded and L is elliptic, then there are constants $c_2 > 0$ and $c_3 \ge 0$ such that

$$c_2\|v\|^2_{H^1_0(\Omega)} \le a(v,v) + c_3\|v\|_2^2.$$

Proof. Consider,

$$\begin{aligned} |a(v,w)| &\le \int_\Omega |A(x)\nabla v(x)\cdot\nabla w(x)|\,dx + \int_\Omega |(\boldsymbol{b}(x)\cdot\nabla v(x))w(x)|\,dx \\ &\quad + \int_\Omega |c(x)v(x)w(x)|\,dx \\ &\le \max_{i,j}\|a_{ij}\|_\infty\|\nabla v\|_2\|\nabla w\|_2 + \max_i\|b_i\|_\infty\|\nabla v\|_2\|w\|_2 + \|c\|_\infty\|v\|_2\|w\|_2 \\ &\le m\|\nabla v\|_2(\|\nabla w\|_2 + \|w\|_2) + m\|v\|_2\|w\|_2, \\ &\qquad \text{where } m = \max\left(\max_{i,j}\|a_{ij}\|_\infty, \max_i\|b_i\|_\infty, \|c\|_\infty\right) \\ &\le c_1\|v\|_{H^1_0(\Omega)}\|w\|_{H^1_0(\Omega)}. \end{aligned}$$

Since L is elliptic, we have

$$\begin{aligned}
\alpha\|\nabla v\|_2^2 &\leq \int_\Omega A(x)\nabla v\cdot\nabla v\,dx \\
&= a(v,v) - \int_\Omega (\boldsymbol{b}\cdot\nabla v)v\,dx - \int_\Omega cv^2\,dx \\
&\leq a(v,v) + \max_i \|b_i\|_\infty \|\nabla v\|_2\|v\|_2 + \|c\|_\infty\|v\|_2^2.
\end{aligned}$$

If $\boldsymbol{b}=0$, then we have the result with $c_2=\alpha$ and $c_3=\|c\|_\infty$. If $\boldsymbol{b}\neq 0$, we choose a $\gamma>0$ such that

$$\gamma < \frac{2\alpha}{\max_i\|b_i\|_\infty}.$$

Then, we have

$$\begin{aligned}
\alpha\|\nabla v\|_2^2 &\leq a(v,v) + \max_i \|b_i\|_\infty \|\nabla v\|_2\|v\|_2 + \|c\|_\infty\|v\|_2^2 \\
&= a(v,v) + \max_i \|b_i\|_\infty \gamma^{1/2}\|\nabla v\|_2 \frac{\|v\|_2}{\gamma^{1/2}} + \|c\|_\infty\|v\|_2^2 \\
&\leq a(v,v) + \frac{\max_i\|b_i\|_\infty}{2}\left(\gamma\|\nabla v\|_2^2 + \frac{\|v\|_2^2}{\gamma}\right) + \|c\|_\infty\|v\|_2^2 \\
&\qquad \text{(using } ab\leq a^2/2+b^2/2\text{)} \\
\left(\alpha - \frac{\gamma}{2}\max_i\|b_i\|_\infty\right)\|\nabla v\|_2^2 &\leq a(v,v) + \left(\frac{1}{2\gamma}\max_i\|b_i\|_\infty + \|c\|_\infty\right)\|v\|_2^2.
\end{aligned}$$

By Poincaré inequality there is a constant $C>0$ such that $1/C\|v\|_{H_0^1(\Omega)}^2 \leq \|\nabla v\|_2^2$. Thus, we have

$$c_2\|v\|_{H_0^1(\Omega)}^2 \leq a(v,v) + c_3\|v\|_2^2.$$

$\square$

Theorem 4.3.5. *Let Ω be a bounded open subset of $\mathbb{R}^n$, $a_{ij}, c\in L^\infty(\Omega)$, $\boldsymbol{b}=0$, $c(x)\geq 0$ a.e. in Ω and $f\in H^{-1}(\Omega)$. Also, let A satisfy ellipticity condition. Then there is a unique weak solution $u\in H_0^1(\Omega)$ satisfying*

$$\int_\Omega A\nabla u\cdot\nabla v\,dx + \int_\Omega cuv\,dx = \langle f, v\rangle_{H^{-1}(\Omega),H_0^1(\Omega)}, \quad \forall v\in H_0^1(\Omega).$$

Further, if A is symmetric then u minimizes the functional $J : H_0^1(\Omega) \to \mathbb{R}$ defined as,

$$J(v) := \frac{1}{2}\int_\Omega A\nabla v \cdot \nabla v\, dx + \frac{1}{2}\int_\Omega cv^2\, dx - \langle f, v\rangle_{H^{-1}(\Omega), H_0^1(\Omega)}$$

in $H_0^1(\Omega)$.

Proof. We define the bilinear form as

$$a(v, w) := \int_\Omega A\nabla v \cdot \nabla w\, dx + \int_\Omega cvw\, dx.$$

It follows from Lemma 4.3.4 that a is a continuous. Now ,consider

$$\begin{aligned} \alpha\|\nabla v\|_2^2 &\leq \int_\Omega A(x)\nabla v \cdot \nabla v\, dx \\ &\leq \int_\Omega A(x)\nabla v \cdot \nabla v\, dx + \int_\Omega cv^2\, dx \quad (\text{since } c(x) \geq 0) \\ &= a(v, v). \end{aligned}$$

Thus, a is coercive in $H_0^1(\Omega)$, by Poincaré inequality. Hence, by Lax Milgram theorem (cf. Theorem **??**), $u \in H_0^1(\Omega)$ exists. Also, if A is symmetric, then u minimizes the functional J on $H_0^1(\Omega)$. □

Theorem 4.3.6. *Let Ω be a bounded open subset of $\mathbb{R}^n$, $a_{ij}, b_i, c \in L^\infty(\Omega)$ and $f \in L^2(\Omega)$. Also, let A satisfy ellipticity condition. Consider L as in (4.3.1). The space of solutions $\{u \in H_0^1(\Omega) \mid Lu = 0\}$ is finite dimensional. For non-zero $f \in L^2(\Omega)$, there exists a finite dimensional subspace $S \subset L^2(\Omega)$ such that (4.3.1) has solution iff $f \in S^\perp$, the orthogonal complement of S.*

Proof. It is already noted in Lemma 4.3.4 that one can find a $c_3 > 0$ such that $a(v, v) + c_3\|v\|_2^2$ is coercive in $H_0^1(\Omega)$. Thus, by Theorem 4.3.5, there is a unique $u \in H_0^1(\Omega)$ such that

$$a(u, v) + c_3\int_\Omega uv\, dx = \int_\Omega fv\, dx \quad \forall v \in H_0^1(\Omega).$$

Set the map $T : L^2(\Omega) \to H_0^1(\Omega)$ as $Tf = u$. The map T is a compact operator on $L^2(\Omega)$ because it maps u into $H_0^1(\Omega)$ which is compactly contained in $L^2(\Omega)$. Note that (4.3.1) is equivalent to $u = T(f + c_3u)$. Set

$v := f + c_3 u$. Then $v - c_3 Tv = f$. Recall that T is compact and $c_3 > 0$. Thus, $I - c_3 T$ is invertible except when c_3^{-1} is an eigenvalue of T. If c_3^{-1} is not an eigenvalue then there is a unique solution v for all $f \in L^2(\Omega)$. If c_3^{-1} is an eigenvalue then it has finite geometric multiplicity (T being compact). Therefore, by Fredhölm alternative (cf. Theorems **??** and **??**), solution exists iff $f \in N(I - c_3 T^*)^\perp$ and the dimension of $S := N(I - c_3 T^*)$ is same as $N(I - c_3 T)$. □

Theorem 4.3.7 (Regularity of Weak Solution)**.** *Let Ω be an open subset of class C^2. Let $a_{ij}, b_i, c \in L^\infty(\Omega)$ and $f \in L^2(\Omega)$. Let $u \in H^1_0(\Omega)$ be such that it satisfies (4.3.2). If $a_{ij} \in C^1(\overline{\Omega})$, $b_i \in C(\overline{\Omega})$ and $f \in L^2(\Omega)$ then $u \in H^2(\Omega)$. More generally, for $m \geq 1$, if $a_{ij} \in C^{m+1}(\overline{\Omega})$, $b_i \in C^m(\overline{\Omega})$ and $f \in H^m(\Omega)$ then $u \in H^{m+2}(\Omega)$.*

Exercise 68. Let $\{f_k\}_1^\infty$ be a bounded sequence in $W^{-1,p}(\Omega)$, for some $p > 2$. If $f_k = g_k + h_k$, where $\{g_k\}$ is relatively compact in $H^{-1}(\Omega)$ and $\{h_k\}$ is bounded in $\mathcal{R}(\Omega)$, then $\{f_k\}$ is relatively compact in $H^{-1}(\Omega)$.

Proof. For $k = 1, 2, ...$ let $u_k \in H^1_0(\Omega)$ be the weak solution of

$$\begin{cases} -\Delta u_k &= f_k \quad \text{in } \Omega, \\ \quad u_k &= 0 \quad \text{on } \partial\Omega. \end{cases}$$

Then $u_k = v_k + w_k$, where

$$\begin{cases} -\Delta v_k &= g_k \quad \text{in } \Omega, \\ \quad v_k &= 0 \quad \text{on } \partial\Omega \end{cases}$$

and

$$\begin{cases} -\Delta w_k &= h_k \quad \text{in } \Omega, \\ \quad w_k &= 0 \quad \text{on } \partial\Omega. \end{cases}$$

We observe that $\{v_k\}$ is relatively compact in $H^1_0(\Omega)$. By the compact imbedding of measures (above theorem), $\{w_k\}$ is relatively compact in $W^{1,r}_0(\Omega)$ for each $1 \leq r < \frac{n}{n-1}$. Thus, $\{u_k\}$ is relatively compact in $W^{1,r}_0(\Omega)$, and $\{f_k\}$ is relatively compact in $W^{-1,r}(\Omega)$. As $\{f_k\}$ is bounded in $W^{-1,p}(U)$ for some $p > 2$, the result follows. □

Theorem 4.3.8 (Weak Maximum Principle)**.** *Let Ω be a bounded open subset of $\mathbb{R}^n$ with sufficient smooth boundary $\partial\Omega$. Let $a_{ij}, c \in L^\infty(\Omega)$, $c(x) \geq 0$ and $f \in L^2(\Omega)$. Let $u \in H^1(\Omega) \cap C(\overline{\Omega})$ be such that it satisfies (4.3.2) with $\boldsymbol{b} \equiv 0$. Then the following are true:*

(i) *If* $f \geq 0$ *on* Ω *and* $u \geq 0$ *on* $\partial\Omega$ *then* $u \geq 0$ *in* Ω.

(ii) *If* $c \equiv 0$ *and* $f \geq 0$ *then* $u(x) \geq \inf_{y \in \partial\Omega} u(y)$ *for all* $x \in \Omega$.

(iii) *If* $c \equiv 0$ *and* $f \equiv 0$ *then* $\inf_{y \in \partial\Omega} u(y) \leq u(x) \leq \sup_{y \in \partial\Omega} u(y)$ *for all* $x \in \Omega$.

Proof. Recall that if $u \in H^1(\Omega)$ then $|u|$, u^+ and u^- are also in $H^1(\Omega)$.

(i) If $u \geq 0$ on $\partial\Omega$ then $u = |u|$ on $\partial\Omega$. Hence, $u^- \in H^1_0(\Omega)$. Thus, using $v = u^-$ in (4.3.2), we get

$$-\int_\Omega A\nabla u^- \cdot \nabla u^- \, dx - \int_\Omega c(x)(u^-)^2 \, dx = \int_\Omega f(x) u^-(x) \, dx$$

because u^+ and u^- intersect on $\{u = 0\}$ and, on this set, $u^+ = u^- = 0$ and $\nabla u^+ = \nabla u^- = 0$ a.e. Note that RHS is non-negative because both f and u^- are non-negative. Therefore,

$$0 \geq \int_\Omega A\nabla u^- \cdot \nabla u^- \, dx + \int_\Omega c(x)(u^-)^2 \, dx \geq \alpha \|\nabla u^-\|_2^2.$$

Thus, $\|\nabla u^-\|_2 = 0$ and, by Poincarè inequality, $\|u^-\|_2 = 0$. This implies $u^- = 0$ a.e and, hence, $u = u^+$ a.e. on Ω.

(ii) Let $m = \inf_{y \in \partial\Omega} u(y)$. Then $u - m \geq 0$ on $\partial\Omega$. Further, $c \equiv 0$ implies that $u - m$ satisfies (4.3.2) with $\boldsymbol{b} = 0$. By previous case, $u - m \geq 0$ on Ω.

(iii) If $f \equiv 0$ then $-u$ satisfies (4.3.2) with $\boldsymbol{b} = 0$. By previous case, we have the result.

□

Theorem 4.3.9 (Spectral Decomposition). *Let* A *be such that* $a_{ij}(x) = a_{ji}(x)$, *i.e., is a symmetric matrix and* $c(x) \geq 0$. *There exists a sequence of positive real eigenvalues* $\{\lambda_m\}$ *and corresponding orthonormal basis* $\{\phi_m\} \subset C^\infty(\Omega)$ *of* $L^2(\Omega)$, *with* $m \in \mathbb{N}$, *such that*

$$\begin{cases} -\operatorname{div}[A(x)\nabla\phi_m(x)] + c(x)\phi_m(x) &= \lambda_m \phi_m(x) \quad \text{in } \Omega \\ \phi_m &= 0 \quad \text{on } \partial\Omega \end{cases} \tag{4.3.3}$$

and $0 < \lambda_1 \leq \lambda_2 \leq \ldots$ *diverges.*

Proof. Let $T : L^2(\Omega) \to H_0^1(\Omega)$ defined as $Tf = u$ where u is the solution of

$$\begin{cases} -\text{div}[A(x)\nabla u(x)] + c(x)u(x) & = f(x) \quad \text{in } \Omega \\ \phi_m & = 0 \quad \text{on } \partial\Omega. \end{cases}$$

Thus,

$$\int_\Omega A(x)\nabla(Tf)\cdot\nabla v(x)\,dx + \int_\Omega c(x)(Tf)(x)v(x)\,dx = \int_\Omega f(x)v(x)\,dx.$$

Note that T is a compact operator on $L^2(\Omega)$ and T is self-adjoint because, for every $g \in L^2(\Omega)$,

$$\begin{aligned} \int_\Omega (Tf)(x)g(x)\,dx &= \int_\Omega A(x)\nabla(Tg)\cdot\nabla(Tf)\,dx + \int_\Omega c(x)(Tg)(x)(Tf)(x)\,dx \\ &= \int_\Omega (Tg)(x)f(x)\,dx. \end{aligned}$$

Further, T is positive definite because, for $f \not\equiv 0$,

$$\begin{aligned} \int_\Omega (Tf)(x)f(x)\,dx &= \int_\Omega A(x)\nabla(Tf)\cdot\nabla(Tf)\,dx + \int_\Omega c(x)(Tf)^2(x)\,dx \\ &\geq \alpha\|\nabla Tf\|_2^2 > 0. \end{aligned}$$

Thus, there exists an orthonormal basis of eigenfunctions $\{\phi_m\}$ in $L^2(\Omega)$ and a sequence of positive eigenvalues μ_m decreasing to zero such that $T\phi_m = \mu_m\phi_m$. Set $\lambda_m = \mu_m^{-1}$. Then $\phi_m = \lambda_m T\phi_m = T(\lambda_m\phi_m)$. Thus, $\phi_m \in H_0^1(\Omega)$ because range of T is $H_0^1(\Omega)$. Hence, ϕ_m satisfies (4.3.3). It now only remains to show that $\phi_m \in C^\infty(\Omega)$. For any $x \in \Omega$, choose $B_r(x) \subset \Omega$. Since $\phi_m \in L^2(B_r(x))$ and solves the eigen value problem, by interior regularity (cf. Theorem 4.3.7), $\phi_m \in H^2(B_r(x))$. Arguing similarly, we obtain $\phi_m \in H^k(B_r(x))$ for all k. Thus, by Sobolev imbedding results, $\phi_m \in C^\infty(B_r(x))$. Since $x \in \Omega$ is arbitrary, $\phi_m \in C^\infty(\Omega)$. □

Remark 4.3.10. Observe that if $H_0^1(\Omega)$ is equipped with the inner product $\int_\Omega \nabla u\cdot\nabla v\,dx$, then $\lambda_m^{-1/2}\phi_m$ is an orthonormal basis for $H_0^1(\Omega)$ where (λ_m, ϕ_m) is the eigen pair corresponding to $A(x) = I$ and $c \equiv 0$. With the usual inner product

$$\int_\Omega uv\,dx + \int_\Omega \nabla u\cdot\nabla v\,dx$$

in $H_0^1(\Omega)$, $(\lambda_m+1)^{-1/2}\phi_m$ forms an orthonormal basis of $H_0^1(\Omega)$. The set $\{\phi_m\}$ is dense in $H_0^1(\Omega)$ w.r.t both the norms mentioned above. Suppose $f \in H_0^1(\Omega)$ is such that $\langle f, \phi_m \rangle = 0$ in $H_0^1(\Omega)$, for all m. Then, from the eigenvalue problem, we get $\lambda_m \int_\Omega \phi_m f\, dx = 0$. Since ϕ_m is a basis for $L^2(\Omega)$, $f = 0$.

Definition 4.3.11. *The* Rayleigh quotient *map* $R : H_0^1(\Omega) \setminus \{0\} \to [0, \infty)$ *is defined as*

$$R(v) = \frac{\int_\Omega A(x)\nabla v \cdot \nabla v\, dx + \int_\Omega c(x) v^2(x)\, dx}{\|v\|_{2,\Omega}^2}.$$

4.4 Periodic Boundary Conditions

Let $Y = [0,1)^n$ be the unit cell of $\mathbb{R}^n$ and let, for each $i, j = 1, 2, \dots, n$, $a_{ij} : Y \to \mathbb{R}$ and $A(y) = (a_{ij})$. For any given $f : Y \to \mathbb{R}$, extended Y-periodically to $\mathbb{R}^n$, we want to solve the problem

$$\begin{cases} -\text{div}(A(y)\nabla u(y)) & = f(y) \quad \text{in } Y \\ \qquad\qquad u \quad \text{is} & \quad Y - \text{periodic.} \end{cases} \tag{4.4.1}$$

The condition u is Y-periodic is equivalent to saying that u takes equal values on opposite faces of Y. One may rewrite the equation on the n-dimensional unit torus $\mathbb{T}^n$ without the periodic boundary condition.

Let us now identify the solution space for (4.4.1). Let $C^\infty_{\text{per}}(Y)$ be the set of all Y-periodic functions in $C^\infty(\mathbb{R}^n)$. Let $H^1_{\text{per}}(Y)$ denote the closure of $C^\infty_{\text{per}}(Y)$ in the H^1-norm. Being a second order equation, in the weak formulation, we expect the weak solution u to be in $H^1_{\text{per}}(Y)$. Note that if u solves (4.4.1) then $u + c$, for any constant c, also solves (4.4.1). Thus, the solution will be unique up to a constant in the space $H^1_{\text{per}}(Y)$. Therefore, we define the quotient space $W_{\text{per}}(Y) = H^1_{\text{per}}(Y)/\mathbb{R}$ as solution space where the solution is unique.

Solving (4.4.1) is to find $u \in W_{\text{per}}(Y)$, for any given $f \in (W_{\text{per}}(Y))^\star$ in the dual of $W_{\text{per}}(Y)$, such that

$$\int_Y A\nabla u \cdot \nabla v\, dx = \langle f, v \rangle_{(W_{\text{per}}(Y))^\star, W_{\text{per}}(Y)} \quad \forall v \in W_{\text{per}}(Y).$$

The requirement that $f \in (W_{\text{per}}(Y))^\star$ is equivalent to saying that

$$\int_Y f(y)\, dy = 0$$

because f defines a linear functional on $W_{\mathrm{per}}(Y)$ and $f(0) = 0$, where $0 \in H^1_{\mathrm{per}}(Y)/\mathbb{R}$. In particular, the equivalence class of 0 is same as the equivalence class 1 and hence

$$\int_Y f(y)\,dy = \langle f, 1\rangle = \langle f, 0\rangle = 0.$$

Theorem 4.4.1. *Let Y be unit open cell and let $a_{ij} \in L^\infty(\Omega)$ such that the matrix $A(y) = (a_{ij}(y))$ is elliptic with ellipticity constant $\alpha > 0$. For any $f \in (W_{per}(Y))^\star$, there is a unique weak solution $u \in W_{per}(Y)$ satisfying*

$$\int_Y A\nabla u \cdot \nabla v\,dx = \langle f, v\rangle_{(W_{per}(Y))^\star, W_{per}(Y)} \quad \forall v \in W_{per}(Y).$$

Note that the solution u we find from above theorem is an equivalence class of functions which are all possible solutions. Any representative element from the equivalence class is a solution. All the elements in the equivalence differ by a constant. Let u be an element from the equivalence class and let c be the constant

$$c = \frac{1}{|Y|}\int_Y u(y)\,dy.$$

Thus, we have $u-c$ is a solution with zero mean value in Y, i.e., $\int_Y u(y)\,dy = 0$. Therefore, rephrasing (4.4.1) as

$$\begin{cases} -\mathrm{div}(A(y)\nabla u(y)) & = f(y) \quad \text{in } Y \\ u & \text{is} \quad Y-\text{periodic} \\ \frac{1}{|Y|}\int_Y u(y)\,dy & = 0 \end{cases}$$

we have unique solution u in the solution space

$$V_{\mathrm{per}}(Y) = \left\{ u \in H^1_{\mathrm{per}}(Y) \mid \frac{1}{|Y|}\int_Y u(y)\,dy = 0 \right\}.$$

4.5 Coercive Neumann Problem

Let Ω be an open subset of $\mathbb{R}^n$ and let $\partial\Omega$ be the topological boundary of Ω. The inhomogeneous Neumann problem, i.e., given $f : \Omega \to \mathbb{R}$, $g : \partial\Omega \to \mathbb{R}$, $a_{ij} \in L^\infty(\Omega)$ and $c \in L^\infty(\Omega)$ such that $c \not\equiv 0$ and $c(x) \geq 0$ for a.e. $x \in \Omega$, find $u : \overline{\Omega} \to \mathbb{R}$ such that

$$\begin{cases} -\mathrm{div}(A(x)\nabla u(x)) + c(x)u(x) & = f(x) \quad \text{in } \Omega \\ A(x)\nabla u \cdot \nu & = g \quad \text{on } \partial\Omega \end{cases} \tag{4.5.1}$$

where $\nu = (\nu_1, \ldots, \nu_n)$ is the unit outward normal. It is customary to denote $A\nabla u \cdot \nu$ as $\frac{\partial u}{\partial \nu_A}$. Note that the boundary condition imposed above is referred to as the Neumann boundary condition.

Definition 4.5.1. *We say $u : \overline{\Omega} \to \mathbb{R}$ is a* classical *solution of the Neumann problem* (4.5.1)*, if $u \in C^2(\overline{\Omega})$ and satisfies* (4.5.1) *pointwise for each x.*

Let u be a classical solution of (4.5.1) and suppose Ω is bounded with C^1-smooth boundary (cf. Corollary **??**). Then, multiplying any $\phi \in C^1(\Omega)$ and integrating both sides in (4.5.1), we get

$$\begin{aligned}
-\int_\Omega (\nabla \cdot (A\nabla u))\phi\, dx + \int_\Omega cu\phi\, dx &= \int_\Omega f\phi\, dx \\
\int_\Omega A\nabla u \cdot \nabla\phi\, dx - \int_{\partial\Omega} \phi \frac{\partial u}{\partial \nu_A}\, d\sigma + \int_\Omega c(x)u\phi\, dx &= \int_\Omega f\phi\, dx \\
&\quad \text{(By Green's identity)} \\
\int_\Omega A\nabla u \cdot \nabla\phi\, dx + \int_\Omega c(x)u\phi\, dx &= \int_\Omega f\phi\, dx + \int_{\partial\Omega} g\phi\, d\sigma \\
&\quad \left(\text{since } \frac{\partial u}{\partial \nu_A} = g \text{ on } \partial\Omega\right).
\end{aligned}$$

Thus, any classical solution u of (4.5.1) solves the problem: For any given f, find u such that

$$\int_\Omega A\nabla u \cdot \nabla\phi\, dx + \int_\Omega c(x)u\phi\, dx = \int_\Omega f\phi\, dx + \int_{\partial\Omega} g\phi\, d\sigma \quad \forall \phi \in C^1(\Omega),$$

as long as the integrals make sense. Note that, in the above problem, it is enough for u to be in $C^1(\Omega)$. Also, in particular, any classical solution satisfies the identity

$$\int_\Omega A\nabla u \cdot \nabla u\, dx + \int_\Omega c(x)u^2\, dx = \int_\Omega fu\, dx + \int_{\partial\Omega} gu\, d\sigma.$$

By the density of $C^1(\Omega)$ in $H^1(\Omega)$, the classical solution u will also solve

$$\int_\Omega A\nabla u \cdot \nabla v\, dx + \int_\Omega c(x)uv\, dx = \int_\Omega fv\, dx + \int_{\partial\Omega} gv\, d\sigma \quad \forall v \in H^1(\Omega),$$

where v inside the boundary integral is in the trace sense, as long as the integrals make sense. The integrals will make sense if $f \in L^2(\Omega)$ and $g \in L^2(\partial\Omega)$. In fact, one can consider $g \in H^{-1/2}(\partial\Omega)$, if we choose to replace the boundary integral with the duality product $\langle \cdot, \cdot \rangle_{H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega)}$. The arguments above motivates the notion of weak solution for (4.5.1).

Definition 4.5.2. *Given $f \in L^2(\Omega)$ and $g \in H^{-1/2}(\partial\Omega)$, we say $u \in H^1(\Omega)$ is a* weak solution *of* (4.5.1) *if*

$$\int_\Omega A\nabla u\cdot\nabla v\,dx+\int_\Omega c(x)uv\,dx = \int_\Omega fv\,dx+\langle g,v\rangle_{H^{-1/2}(\partial\Omega),H^{1/2}(\partial\Omega)} \quad \forall v \in H^1(\Omega). \tag{4.5.2}$$

We have already observed that every classical solution is a weak solution for a bounded open set Ω with C^1-smooth boundary. Let $\mathcal{M}(\alpha)$ denote the space of all $n \times n$ matrix $A = (a_{ij})$ such that $a_{ij} \in L^\infty(\Omega)$ for all i, j and

$$\alpha|\xi|^2 \leq A(x)\xi.\xi \quad \text{a.e. in } x, \quad \forall \xi = (\xi_i) \in \mathbb{R}^n.$$

Theorem 4.5.3. *Let Ω be a bounded open subset of $\mathbb{R}^n$, $f \in L^2(\Omega)$ and $g \in H^{-1/2}(\partial\Omega)$. Also, let $A \in \mathcal{M}(\alpha)$ and $\gamma > 0$ be such that $c(x) \geq \gamma$ for a.e. $x \in \Omega$. Then there is a unique weak solution $u \in H^1(\Omega)$ satisfying* (4.5.2). *Moreover, if A is symmetric, then u minimizes the functional $J : H^1(\Omega) \to \mathbb{R}$ defined as,*

$$J(v) := \frac{1}{2}\int_\Omega A\nabla v\cdot\nabla v\,dx + \frac{1}{2}\int_\Omega cv^2\,dx - \int_\Omega fv\,dx - \langle g,v\rangle_{H^{-1/2}(\partial\Omega),H^{1/2}(\partial\Omega)}$$

in $H^1(\Omega)$.

Proof. Recall that $H^1(\Omega)$ is a Hilbert space endowed with the inner-product

$$(v,w) := \int_\Omega vw\,dx + \int_\Omega \nabla v\nabla w\,dx.$$

We shall use the Lax-Milgram result (cf. Theorem **??**). We define the map $a : H^1(\Omega) \times H^1(\Omega) :\to \mathbb{R}$ as

$$a(v,w) := \int_\Omega A\nabla v\cdot\nabla w\,dx + \int_\Omega cvw\,dx.$$

Note that a is a bilinear. We need to show that a is continuous and coercive. Consider,

$$\begin{aligned}
|a(v,w)| &\leq \max_{i,j}\|a_{ij}\|_\infty\int_\Omega|\nabla v||\nabla w|\,dx + \|c\|_\infty\int_\Omega|vw|\,dx\\
&\leq C\left(\|\nabla v\|_2\|\nabla w\|_2 + \|v\|_2\|w\|_2\right)\\
&\quad\ \text{(where } C := \max\left\{\max_{i,j}\|a_{ij}\|_\infty, \|c\|_\infty\right\}\text{)}\\
&\leq 2C\|v\|_{H^1(\Omega)}\|w\|_{H^1(\Omega)}.
\end{aligned}$$

Thus, a is a bilinear continuous form. Now, consider

$$\begin{aligned} a(v,v) &= \int_\Omega A\nabla v\cdot\nabla v\,dx + \int_\Omega cv^2\,dx \\ &\geq \alpha\|\nabla v\|_2^2 + \gamma\|v\|_2^2 \\ &\geq \min\{\alpha,\gamma\}\|v\|^2_{H^1(\Omega)}. \end{aligned}$$

Hence a is coercive, bilinear continuous form. Also, note that the map

$$v \mapsto \int_\Omega fv\,dx + \langle g, v\rangle_{H^{-1/2}(\partial\Omega),H^{1/2}(\partial\Omega)}$$

is continuous linear functional on $H^1(\Omega)$, since

$$\left|\int_\Omega fv\,dx\right| \leq \|f\|_2\|v\|_2 \leq \|f\|_2\|v\|_{H^1(\Omega)}$$

and the linearity and continuity of the trace map T. Thus, by Lax-milgram theorem, there is a $u \in H^1(\Omega)$ such that

$$a(u,v) = \int_\Omega fv\,dx + \langle g, v\rangle_{H^{-1/2}(\partial\Omega),H^{1/2}(\partial\Omega)}.$$

This is equivalent to (4.5.2) and if A is symmetric then u minimizes the functional

$$J(v) := \frac{1}{2}\int_\Omega A\nabla v\cdot\nabla v\,dx + \frac{1}{2}\int_\Omega cv^2\,dx - \int_\Omega fv\,dx - \langle g, v\rangle_{H^{-1/2}(\partial\Omega),H^{1/2}(\partial\Omega)}$$

in $H^1(\Omega)$. □

The term *coercive* refers to the fact that both A and c are positive definite. In contrast to the Dirichlet boundary condition, observe that the Neumann boundary condition is not imposed *a priori*, but gets itself imposed naturally. In other words, if u is a weak solution of (4.5.1) and is in $H^2(\Omega)$, then for any $v \in H^1(\Omega)$,

$$\begin{aligned} \int_\Omega A\nabla u\cdot\nabla v\,dx + \int_\Omega c(x)uv\,dx &= \int_\Omega fv\,dx + \int_{\partial\Omega} gv\,d\sigma \\ -\int_\Omega \nabla\cdot(A\nabla u)v\,dx + \int_{\partial\Omega} v\frac{\partial u}{\partial\nu_A}\,d\sigma + \int_\Omega c(x)uv\,dx &= \int_\Omega fv\,dx + \int_{\partial\Omega} gv\,d\sigma \end{aligned}$$

implies that

$$\int_{\partial\Omega} v \left(\frac{\partial u}{\partial \nu_A} - g \right) d\sigma = 0 \quad \forall v \in H^1(\Omega).$$

But $v\,|_{\partial\Omega} \in H^{1/2}(\partial\Omega)$, since $v \in H^1(\Omega)$. Now, by the density of $H^{1/2}(\partial\Omega)$ in $L^2(\partial\Omega)$ (cf. Theorem 2.9.8), we have $\frac{\partial u}{\partial \nu_A} = g$ in $L^2(\partial\Omega)$. Thus, note that we have obtained the Neumann boundary condition as a consequence of u being a weak solution to the Neumann problem on the domain Ω. Thus, Dirichlet conditions are called *essential boundary condition* and Neumann conditions are called *natural boundary conditions*.

An astute reader should be wondering that if $H^1_0(\Omega)$ was the precise solution space for the Dirichlet problem (4.1.1), then the corresponding solution space for the Neumann problem (4.5.1) should be the closure of

$$V_g := \{\phi \in \mathcal{D}(\Omega) \mid A\nabla\phi \cdot \nu = g \text{ on } \partial\Omega\}$$

in $H^1(\Omega)$. The guess is right, solve the following exercise!

Exercise 69. $H^1(\Omega)$ is the closure of V_g in $H^1(\Omega)$-norm.

4.6 Semi-coercive Neumann Problem

Let Ω be an open subset of $\mathbb{R}^n$ and let $\partial\Omega$ be the topological boundary of Ω. We consider the inhomogeneous Neumann problem with $c \equiv 0$ (thus, the operator is semi-coercive), i.e., given $f : \Omega \to \mathbb{R}$, $g : \partial\Omega \to \mathbb{R}$ and $a_{ij} \in L^\infty(\Omega)$, find $u : \overline{\Omega} \to \mathbb{R}$ such that

$$\begin{cases} -\operatorname{div}(A(x)\nabla u(x)) &= f(x) \quad \text{in } \Omega \\ A(x)\nabla u \cdot \nu &= g \quad\quad\; \text{on } \partial\Omega \end{cases} \tag{4.6.1}$$

where $\nu = (\nu_1, \ldots, \nu_n)$ is the unit outward normal.

In contrast to the coercive Neumann problem, (4.6.1) is not well posed. This can be seen as follows:

Let u be a classical solution of (4.6.1) and suppose Ω is bounded with C^1-smooth boundary (cf. Corollary **??**). Then, multiplying any $1 \in C^1(\Omega)$

and integrating both sides in (4.6.1), we get

$$\begin{aligned}
-\int_\Omega (\nabla\cdot(A\nabla u))\,dx &= \int_\Omega f\,dx \\
-\int_{\partial\Omega} \frac{\partial u}{\partial \nu_A}\,d\sigma &= \int_\Omega f\,dx \\
&\quad \text{(By Green's identity)} \\
\int_\Omega f\,dx + \int_{\partial\Omega} g\,d\sigma &= 0 \\
&\quad \left(\text{since } \frac{\partial u}{\partial \nu_A} = g \text{ on } \partial\Omega\right).
\end{aligned}$$

Thus, any classical solution u of (4.6.1) necessarily satisfies

$$\int_\Omega f\,dx + \int_{\partial\Omega} g\,d\sigma = 0.$$

This is called the *compatibility* condition for the semi-coercive Neumann problem. We shall now see that the compatibility condition is also a sufficient condition. Moreover, observe that if u is a solution of (4.6.1), then for any constant C, $u + C$ is also a solution of (4.6.1). Thus, one can expect uniqueness only up to a constant. In particular, if u is a solution of (4.6.1), then $w := u - \frac{1}{|\Omega|}\int_\Omega u\,dx$ is also a solution of (4.6.1) such that $\int_\Omega w\,dx = 0$. Let us introduce the space

$$V := \left\{ v \in H^1(\Omega) \mid \int_\Omega v\,dx = 0 \right\}.$$

Exercise 70. Show that V is a closed subspace of $H^1(\Omega)$.

We expect the solution of (4.6.1) to be unique in V. We endow V with the inner-product of $H^1(\Omega)$, i.e.,

$$(v,w) := \int_\Omega vw\,dx + \int_\Omega \nabla v \nabla w\,dx.$$

Definition 4.6.1. *Let $f \in L^2(\Omega)$ and $g \in H^{-1/2}(\partial\Omega)$ be such that*

$$\int_\Omega f\,dx + \langle g, 1\rangle_{H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega)} = 0, \tag{4.6.2}$$

then we say $u \in V$ is a weak solution *of (4.6.1) if*

$$\int_\Omega A\nabla u\cdot\nabla v\,dx = \int_\Omega fv\,dx + \langle g, v\rangle_{H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega)} \quad \forall v \in V. \tag{4.6.3}$$

As usual, it is easy to check that every classical solution is a weak solution for a bounded open set Ω with C^1-smooth boundary.

Theorem 4.6.2. *Let Ω be a bounded, connected open subset of $\mathbb{R}^n$ with C^1 smooth boundary. Let $f \in L^2(\Omega)$ and $g \in H^{-1/2}(\partial\Omega)$ be given such that (4.6.2) is satisfied. Also, let $A \in \mathcal{M}(\alpha)$, then there is a unique weak solution $u \in V$ satisfying (4.6.3). Moreover, if A is symmetric, then u minimizes the functional $J : V \to \mathbb{R}$ defined as,*

$$J(v) := \frac{1}{2}\int_\Omega A\nabla v \cdot \nabla v \, dx - \int_\Omega f v \, dx - \langle g, v\rangle_{H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega)}.$$

Proof. We shall use the Lax-Milgram result (cf. Theorem **??**). We define the map $a : V \times V :\to \mathbb{R}$ as

$$a(v, w) := \int_\Omega A\nabla v \cdot \nabla w \, dx.$$

Note that a is a bilinear. We need to show that a is continuous and coercive. Consider,

$$\begin{aligned} |a(v,w)| &\leq \max_{i,j} \|a_{ij}\|_\infty \int_\Omega |\nabla v||\nabla w| \, dx \\ &\leq \max_{i,j} \|a_{ij}\|_\infty \|\nabla v\|_2 \|\nabla w\|_2 \\ &\leq \max_{i,j} \|a_{ij}\|_\infty \|v\|_{H^1(\Omega)} \|w\|_{H^1(\Omega)}. \end{aligned}$$

Thus, a is a bilinear continuous form. Consider

$$\begin{aligned} a(v,v) &= \int_\Omega A\nabla v \cdot \nabla v \, dx \\ &\geq \alpha \|\nabla v\|_2^2. \end{aligned}$$

Now, by (by Poincaré-Wirtinger, cf. Theorem 2.8.25) in V, there is a constant $C > 0$ such that

$$\|v\|_2 \leq C\|\nabla v\|_2 \quad \forall v \in V.$$

Thus,

$$\begin{aligned} a(v,v) &\geq \alpha \|\nabla v\|_2^2 \\ &\geq \frac{\alpha}{C^2} \|v\|_2^2 \\ &\geq \frac{\alpha}{C^2} \|v\|_{H^1(\Omega)}^2. \end{aligned}$$

Therefore, a is a coercive, bilinear continuous form on V. Also, note that the map

$$v \mapsto \int_\Omega fv\,dx + \langle g, v\rangle_{H^{-1/2}(\partial\Omega),H^{1/2}(\partial\Omega)}$$

is continuous linear functional on $H^1(\Omega)$, since

$$\left|\int_\Omega fv\,dx\right| \leq \|f\|_2\|v\|_2 \leq \|f\|_2\|v\|_{H^1(\Omega)}$$

and the linearity and continuity of the trace map T. Thus, by Lax-milgram theorem, there is a $u \in V$ such that

$$a(u,v) = \int_\Omega fv\,dx + \langle g, v\rangle_{H^{-1/2}(\partial\Omega),H^{1/2}(\partial\Omega)}.$$

This is equivalent to (4.6.3) and if A is symmetric then u minimizes the functional

$$J(v) := \frac{1}{2}\int_\Omega A\nabla v\cdot\nabla v\,dx - \int_\Omega fv\,dx - \langle g, v\rangle_{H^{-1/2}(\partial\Omega),H^{1/2}(\partial\Omega)}$$

in V. □

The term *semi-coercive* refers to the fact that A is positive definite and c is identically zero.

We introduce a different approach to prove the existence of weak solution for the semi-coercive Neumann problem (4.6.1). This approach is called the *Tikhonov Regularization Method*. The basic idea is to obtain the semi-coercive case as a limiting case of the coercive Neumann problem.

Theorem 4.6.3. *Let Ω be a bounded, connected open subset of $\mathbb{R}^n$ with C^1 smooth boundary. Let $f \in L^2(\Omega)$ and $g \in H^{-1/2}(\partial\Omega)$ be given such that (4.6.2) is satisfied. Let $A \in \mathcal{M}(\alpha)$ and, for any $\varepsilon > 0$, if $u_\varepsilon \in H^1(\Omega)$ is the weak solution of (4.5.1) with $c \equiv \varepsilon$, i.e.,*

$$\int_\Omega A\nabla u_\varepsilon\cdot\nabla v\,dx + \int_\Omega \varepsilon u_\varepsilon v\,dx = \int_\Omega fv\,dx + \langle g, v\rangle_{H^{-1/2}(\partial\Omega),H^{1/2}(\partial\Omega)} \quad \forall v \in H^1(\Omega), \tag{4.6.4}$$

then $u_\varepsilon \to u$ weakly *in $H^1(\Omega)$ where $u \in V$ is the weak solution of (4.6.1), as obtained in Theorem 4.6.2.*

Proof. Firstly, note that by choosing $v \equiv 1$ in (4.6.4), we have

$$\int_\Omega u_\varepsilon \, dx = 0.$$

Thus, $u_\varepsilon \in V$ for all ε. Now consider,

$$\begin{aligned}
\alpha \|\nabla u_\varepsilon\|_2^2 &\leq \int_\Omega A\nabla u_\varepsilon \cdot \nabla u_\varepsilon \, dx \\
&\leq \int_\Omega A\nabla u_\varepsilon \cdot \nabla u_\varepsilon \, dx + \varepsilon \|u_\varepsilon\|_2^2 \\
&= \int_\Omega f u_\varepsilon \, dx + \langle g, u_\varepsilon \rangle_{H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega)} \\
&\leq \|f\|_2 \|u_\varepsilon\|_2 + \|g\|_{H^{-1/2}(\partial\Omega)} \|Tu_\varepsilon\|_{H^{1/2}(\partial\Omega)} \\
&\leq C\left(\|f\|_2 + \|g\|_{H^{-1/2}(\partial\Omega)}\right) \|\nabla u_\varepsilon\|_2 \\
&\quad \text{(Using Poincaré-Wirtinger and continuity of } T) \\
\|\nabla u_\varepsilon\|_2 &\leq C\left(\|f\|_2 + \|g\|_{H^{-1/2}(\partial\Omega)}\right)
\end{aligned}$$

Hence, again by Poincaré-Wirtinger inequality (cf. Theorem 2.8.25), $\{u_\varepsilon\}$ is a bounded sequence in $H^1(\Omega)$, in fact a bounded sequence in V. Therefore, one can extract a weakly convergence subsequence of $\{u_\varepsilon\}$ (still denoted by u_ε) in V such that $u_\varepsilon \rightharpoonup u$ for some $u \in V$ (since V is closed). Thus, passing to the limit in (4.6.4), we get

$$\int_\Omega A\nabla u \cdot \nabla v \, dx = \int_\Omega f v \, dx + \langle g, v \rangle_{H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega)} \quad \forall v \in V.$$

Thus, $u \in V$ is a weak solution of (4.6.1). Suppose u_1 and u_2 are two solutions of (4.6.1), then $w := u_1 - u_2$ is a solution to (4.6.1) with $f, g \equiv 0$. Thus, w is equal to some constant. Also, $w \in V$ because both u_1 and u_2 are in V. Hence, $w = 0$, since $|\Omega| \neq 0$. Since the subsequence all converge to a unique u, the entire sequence $\{u_\varepsilon\}$ converges weakly to u in $H^1(\Omega)$. Moreover, u is the weak solution of (4.6.1). □

Observe that by choosing $v = u_\varepsilon$ in (4.6.4) and $v = u$ in (4.6.3), and passing to limit, we get

$$\begin{aligned}
\lim_{\varepsilon \to 0} \int_\Omega A\nabla u_\varepsilon \cdot \nabla u_\varepsilon \, dx &= \int_\Omega f u \, dx + \langle g, u \rangle_{H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega)} \\
&= \int_\Omega A\nabla u \cdot \nabla u \, dx.
\end{aligned}$$

Therefore, if A is identity matrix, then $\|\nabla u_\varepsilon\|_2^2 \to \|\nabla u\|_2^2$ and hence u_ε converges *strongly* to u in $H^1(\Omega)$. In fact, in general, we have the identity

$$\begin{aligned}\limsup_{\varepsilon\to 0}\|\nabla u_\varepsilon\|_2^2 &\leq \frac{1}{\alpha}\int_\Omega A\nabla u\cdot\nabla u\,dx\\ &\leq \frac{\max_{i,j}\|a_{ij}\|_\infty}{\alpha}\|\nabla u\|_2^2\\ &\leq \frac{\max_{i,j}\|a_{ij}\|_\infty}{\alpha}\liminf_{\varepsilon\to 0}\|\nabla u_\varepsilon\|_2^2.\end{aligned}$$

Corollary 4.6.4. *If A is symmetric, then the u obtained in the above Theorem minimizes the functional $J : V \to \mathbb{R}$ defined as,*

$$J(v) := \frac{1}{2}\int_\Omega A\nabla v\cdot\nabla v\,dx - \int_\Omega fv\,dx - \langle g, v\rangle_{H^{-1/2}(\partial\Omega),H^{1/2}(\partial\Omega)}.$$

Proof. If A is symmetric, then, for each $\varepsilon > 0$, u_ε minimizes the functional $J_\varepsilon : V \to \mathbb{R}$ defined as

$$J_\varepsilon(v) := \frac{1}{2}\int_\Omega A\nabla v\cdot\nabla v\,dx + \frac{\varepsilon}{2}\int_\Omega v^2\,dx - \int_\Omega fv\,dx - \langle g, v\rangle_{H^{-1/2}(\partial\Omega),H^{1/2}(\partial\Omega)}.$$

Equivalently, $J_\varepsilon(u_\varepsilon) \leq J_\varepsilon(v)$ for all $v \in V$. It easy to note that J_ε converges to J pointwise. Now taking limits both sides, we get,

$$\begin{aligned}\lim_{\varepsilon\to 0} J_\varepsilon(u_\varepsilon) &\leq \lim_{\varepsilon\to 0} J_\varepsilon(v) \quad \forall v \in V\\ \lim_{\varepsilon\to 0}\frac{1}{2}\left(\int_\Omega A\nabla u_\varepsilon\cdot\nabla u_\varepsilon\,dx + \varepsilon\int_\Omega u_\varepsilon^2\,dx\right) - & \\ \lim_{\varepsilon\to 0}\left(\int_\Omega fu_\varepsilon\,dx + \langle g, u_\varepsilon\rangle_{H^{-1/2}(\partial\Omega),H^{1/2}(\partial\Omega)}\right) &\leq J(v) \quad \forall v \in V\\ J(u) &\leq J(v) \quad \forall v \in V.\end{aligned}$$

Hence u minimizes J in V. □

4.7 Mixed and Robin Problem

Let Ω be an open subset of $\mathbb{R}^n$ and let $\partial\Omega$ be the topological boundary of Ω. Let $\Gamma_1 \subset \partial\Omega$ such that $\Gamma \neq \emptyset$. Let $\Gamma_2 := \partial\Omega \setminus \Gamma_1$. We consider the mixed Dirichlet-Neumann problem, i.e., given $f : \Omega \to \mathbb{R}$, $g : \Gamma_2 \to \mathbb{R}$ and $a_{ij} \in L^\infty(\Omega)$, find $u : \overline{\Omega} \to \mathbb{R}$ such that

$$\left\{\begin{array}{rll} -\operatorname{div}(A(x)\nabla u(x)) & = f(x) & \text{in } \Omega \\ u & = 0 & \text{on } \Gamma_1 \\ A(x)\nabla u \cdot \nu & = g & \text{on } \Gamma_2 \end{array}\right. \tag{4.7.1}$$

where $\nu = (\nu_1, \ldots, \nu_n)$ is the unit outward normal. Let $V_m := \{v \in H^1(\Omega) \mid Tv = 0 \text{ on } \Gamma_1\}$.

Exercise 71. Show that V_m is a closed subspace of $H^1(\Omega)$ endowed with the $H^1(\Omega)$-norm. (Hint: use the continuity of the trace operator T).

Definition 4.7.1. *Given $f \in L^2(\Omega)$ and $g \in H^{-1/2}(\Gamma_2)$, we say $u \in V_m$ is a* weak solution *of* (4.7.1) *if*

$$\int_\Omega A\nabla u \cdot \nabla v\,dx = \int_\Omega fv\,dx + \langle g, v\rangle_{H^{-1/2}(\Gamma_2),H^{1/2}(\Gamma_2)} \quad \forall v \in V_m. \tag{4.7.2}$$

Theorem 4.7.2. *Let Ω be a bounded open subset of $\mathbb{R}^n$, $f \in L^2(\Omega)$ and $g \in H^{-1/2}(\Gamma_2)$. Also, let $A \in \mathcal{M}(\alpha)$, then there is a unique weak solution $u \in V_m$ satisfying* (4.7.2)*. Moreover, if A is symmetric, then u minimizes the functional $J : V_m \to \mathbb{R}$ defined as,*

$$J(v) := \frac{1}{2}\int_\Omega A\nabla v \cdot \nabla v\,dx - \int_\Omega fv\,dx - \langle g, v\rangle_{H^{-1/2}(\Gamma_2),H^{1/2}(\Gamma_2)}.$$

We skip the proof of above theorem because the arguments are similar to those already introduced while studying homogeneous Dirichlet problem. The Poincaré inequality is valid in V_m (since u vanishes on section of the boundary, i.e., Γ_1). Thus, all earlier arguments can be carried over smoothly. Similarly, one can also study the mixed Dirichlet-Neumann problem with inhomogeneous condition on Γ_1, say h. In this case the arguments are similar to the Inhomogeneous Dirichlet problem with V_m replaced with its translate $V_m + h$.

Let $c : \partial\Omega \to \mathbb{R}^+$ be given. We consider the Robin problem, i.e., given $f : \Omega \to \mathbb{R}$, $g : \partial\Omega \to \mathbb{R}$ and $a_{ij} \in L^\infty(\Omega)$, find $u : \overline{\Omega} \to \mathbb{R}$ such that

$$\left\{\begin{array}{rll} -\operatorname{div}(A(x)\nabla u(x)) & = f(x) & \text{in } \Omega \\ c(x)u + A(x)\nabla u \cdot \nu & = g & \text{on } \partial\Omega \end{array}\right. \tag{4.7.3}$$

where $\nu = (\nu_1, \ldots, \nu_n)$ is the unit outward normal. Observe that the Robin problem incorporates most of the problems introduced in this chapter. For instance, if $c \equiv 0$, then we have the semi-coercive Neumann problem. If $c \equiv +\infty$, then (formally) we have the homogeneous Dirichlet problem. If $c \equiv +\infty$ on $\Gamma_1 \subset \partial\Omega$ and $c \equiv 0$ in $\partial\Omega \setminus \Gamma_1$, then we have the mixed Dirichlet-Neumann problem.

Definition 4.7.3. *Given $f \in L^2(\Omega)$ and $g \in H^{-1/2}(\partial\Omega)$, we say $u \in H^1(\Omega)$ is a* weak solution *of* (4.7.3) *if*

$$\int_\Omega A\nabla u \cdot \nabla v \, dx + \int_{\partial\Omega} cuv \, d\sigma = \int_\Omega fv \, dx + \langle g, v \rangle_{H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega)} \quad \forall v \in H^1(\Omega). \tag{4.7.4}$$

Theorem 4.7.4. *Let Ω be a bounded, connected open subset of $\mathbb{R}^n$ with C^1 smooth boundary. Let $f \in L^2(\Omega)$, $g \in H^{-1/2}(\partial\Omega)$ and $A \in \mathcal{M}(\alpha)$. Also, $c(x) \geq \gamma$ for a.e. $x \in \partial\Omega$, for some $\gamma > 0$. Then there is a unique weak solution $u \in H^1(\Omega)$ satisfying (4.7.4). Moreover, if A is symmetric, then u minimizes the functional $J : H^1(\Omega) \to \mathbb{R}$ defined as,*

$$J(v) := \frac{1}{2} \int_\Omega A\nabla v \cdot \nabla v \, dx + \frac{1}{2} \int_{\partial\Omega} cv^2 \, d\sigma - \int_\Omega fv \, dx - \langle g, v \rangle_{H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega)}.$$

Proof. The only non-trivial part of the proof is to show the coercivity of the bilinear map

$$a(v, w) := \int_\Omega A\nabla v \cdot \nabla w \, dx + \int_{\partial\Omega} cvw \, d\sigma$$

in $H^1(\Omega)$. □

4.8 p-Laplacian Operator

Let $1 < p < +\infty$. We define the p-Laplacian operator Δ_p, for $1 < p < +\infty$, as

$$\Delta_p v := \operatorname{div}(|\nabla v|^{p-2} \nabla v).$$

Note that when $p = 2$, we get the usual Laplacian operator, Δ. Let Ω be an open subset of $\mathbb{R}^n$ and let $\partial\Omega$ be the topological boundary of Ω. Let us first consider the Dirichlet problem for the p-Laplacian operator. Given $f : \Omega \to \mathbb{R}$, find $u : \overline{\Omega} \to \mathbb{R}$ such that

$$\begin{cases} -\Delta_p u &= f \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega. \end{cases} \tag{4.8.1}$$

Definition 4.8.1. *Let q be the conjugate exponent of p. Given $f \in W^{-1,q}(\Omega)$, we say $u \in W_0^{1,p}(\Omega)$ is a* weak solution *of* (4.8.1) *if*

$$\int_\Omega |\nabla u|^{p-2}\nabla u \cdot \nabla v \, dx = \langle f, v\rangle_{W^{-1,q}(\Omega),W_0^{1,p}(\Omega)}, \quad \forall v \in W_0^{1,p}(\Omega). \tag{4.8.2}$$

Theorem 4.8.2. *Let Ω be a bounded open subset of $\mathbb{R}^n$ and $f \in W^{-1,q}(\Omega)$. Then there is a unique weak solution $u \in W_0^{1,p}(\Omega)$ satisfying* (4.8.2)*. Moreover, u minimizes the functional $J : W_0^{1,p}(\Omega) \to \mathbb{R}$ defined as,*

$$J(v) := \frac{1}{p}\int_\Omega |\nabla v|^p \, dx - \langle f, v\rangle_{W^{-1,q}(\Omega),W_0^{1,p}(\Omega)}.$$

Proof. First prove J is coercive and continuous on $W_0^{1,p}(\Omega)$, then use Theorem **??** to show the existence of a minimizer $u \in W_0^{1,p}(\Omega)$.

Secondly, show J is convex, in fact strictly convex on $W_0^{1,p}(\Omega)$ and consequently conclude the uniqueness of minimizer.

Lastly, show that J is Gâteaux differentiable and use (i) of Proposition **??** to show that the minimizer u is a weak solution of (4.8.1). □

4.9 Stokes Problem

Let Ω be an open subset of $\mathbb{R}^n$ and let $\partial\Omega$ be the topological boundary of Ω. The Stokes problem for a viscous fluid flow is described as follows: Given n functions $f_1, f_2, \ldots, f_n : \Omega \to \mathbb{R}$, find $u_1, u_2, \ldots, u_n : \overline{\Omega} \to \mathbb{R}$ and $p : \Omega \to \mathbb{R}$ such that for each $i = 1, 2, \ldots, n$,

$$\left\{ \begin{array}{rll} -\mu\Delta u_i + \frac{\partial p}{\partial x_i} & = f_i & \text{in } \Omega \\ u_i & = 0 & \text{on } \partial\Omega, \end{array} \right. \tag{4.9.1}$$

where $\mu > 0$ is the viscosity coefficient (it is a scaler inversely proportional to Reynolds number). Observe that we have to find $n+1$ unknowns (u_i's and p) with n equations. The $n+1$ equation is introduced as the incompressibility condition of the fluid, i.e.,

$$\nabla \cdot \boldsymbol{u} = 0, \tag{4.9.2}$$

where $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$. Let $\boldsymbol{f} = (f_1, f_2, \ldots, f_n)$. Thus, the Stokes system can be written in the compressed form,

$$\left\{ \begin{array}{rll} -\mu\Delta \boldsymbol{u} + \nabla p & = \boldsymbol{f} & \text{in } \Omega \\ \operatorname{div}(\boldsymbol{u}) = 0 & \text{in } \Omega & \\ \boldsymbol{u} & = 0 & \text{on } \partial\Omega. \end{array} \right. \tag{4.9.3}$$

4.10 Elasticity System

Let $\Omega \subset \mathbb{R}^3$ denote an elastic homogeneous with boundary $\partial\Omega = \Gamma_1 \cup \Gamma_2$ such that $\Gamma_1 \cap \Gamma_2 = \emptyset$. Assume that the body is fixed along Γ_2 and the surface measure of Γ_2 is strictly positive. Let $\boldsymbol{f} = (f_1, f_2, f_3)$ denote the force applied on the body Ω and $\boldsymbol{g} = (g_1, g_2, g_3)$ denote the force applied on the surface of the body Γ_1. Let $\boldsymbol{u} = (u_1, u_2, u_3)$ denote the displacement vector for each point of Ω. One can derive the *strain* (deformation) tensor ϵ_{ij} as, for $1 \leq i, j \leq 3$,

$$\epsilon_{ij}(\boldsymbol{u}) = \frac{1}{2}\left(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j}\right).$$

Note that the strain tensor is 3×3 symmetric matrix. Let $\sigma_{ij}(\boldsymbol{u})$ denote the *stress* (measure of internal forces arising due to force f acting on Ω) tensor. Then the constitutive law (Hooke's law) states that

$$\sigma_{ij}(\boldsymbol{u}) = \lambda\left(\sum_{k=1}^{3} \epsilon_{kk}(\boldsymbol{u})\right)\delta_{ij} + 2\mu\epsilon_{ij}(\boldsymbol{u}),$$

where $\lambda \geq 0$ and $\mu > 0$ are called *Lame's coefficients* and δ_{ij} is the Kronecker delta defined as

$$\delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j. \end{cases}$$

Note that the stress tensor (σ_{ij}) is also a 3×3 symmetric matrix. Now, for each $i = 1, 2, 3$, the displacement vector is given as a solution of the BVP

$$\begin{cases} -\sum_{j=1}^{3} \frac{\partial}{\partial x_j}(\sigma_{ij}(\boldsymbol{u})) &= f_i & \text{in } \Omega \\ \sum_{j=1}^{3} \sigma_{ij}(\boldsymbol{u})\nu_j &= g_i & \text{on } \Gamma_1 \\ u_i &= 0 & \text{on } \Gamma_2 \end{cases} \tag{4.10.1}$$

4.11 Leray-Lions or Nonlinear Monotone Operators

Definition 4.11.1. *A function $a : \Omega \times \mathbb{R}^n \to \mathbb{R}^n$ is said to be a* Carathéodory function *if $a(\cdot, \xi)$ is measurable on Ω for every $\xi \in \mathbb{R}^n$ and $a(x, \cdot)$ is continuous on $\mathbb{R}^n$ for almost every $x \in \Omega$.*

Let $a(x, \xi)$ be a Carathéodory function satisfying the following hypotheses for almost every $x \in \Omega$:

H 1 (Coercive) There is a constant $\alpha > 0$ such that $a(x, \xi).\xi \geq \alpha|\xi|^p$, for every $\xi \in \mathbb{R}^n$.

H 2 (Bounded) There is a constant $\beta > 0$ and a non-negative function $h \in L^p(\Omega)$ such that $|a(x, \xi)| \leq \beta\left(h(x) + |\xi|^{p-1}\right)$ for every $\xi \in \mathbb{R}^n$.

H 3 (Monotone) $\left(a(x, \xi) - a(x, \xi')\right).(\xi - \xi') > 0$ for every $\xi, \xi' \in \mathbb{R}^n$ and $\xi \neq \xi'$.

Let $H(\alpha, \beta, h)$ denote the set of all Carathéodory functions satisfying the above three hypotheses. For every coercive (hypotheis **H**1) Carathéodory function, $a(x, 0) = 0$ for almost every $x \in \Omega$.

Definition 4.11.2. *Given $f \in W^{-1,q}(\Omega)$, we say $u \in W_0^{1,p}(\Omega)$ is a* weak solution *of*

$$\begin{cases} -\operatorname{div}(a(x, \nabla u)) & = f \quad \text{in } \Omega \\ u & = 0 \quad \text{on } \partial\Omega \end{cases} \tag{4.11.1}$$

whenever

$$\int_\Omega a(x, \nabla u) \cdot \nabla\phi \, dx = \langle f, \phi \rangle_{W^{-1,q}(\Omega), W_0^{1,p}(\Omega)}, \quad \forall \phi \in W_0^{1,p}(\Omega).$$

The map $u \mapsto -\operatorname{div}(a(x, \nabla u))$ is a coercive, continuous, bounded and monotone operator from $W_0^{1,p}(\Omega)$ to $W^{-1,q}(\Omega)$. The existence and uniqueness of solution is a standard theory of monotone operators ([LL65, Lio69]).

4.12 Solutions for Measure data

If $p > n$ then, by Theorem 2.8.26, $\mathcal{B}(\Omega)$ is compactly imbedded in $W^{-1,q}(\Omega)$ for all $q \in [1, \frac{n}{n-1})$. Hence, for $p > n$ the notion of weak solution carries forward to measure data, as well. However, for $p < n$, one cannot expect the solution to belong to $W_0^{1,p}(\Omega)$. For example, consider the Laplace equation in a ball, with μ as the Dirac mass at the centre. existence and uniqueness of sul

4.12.1 Stampacchia Solutions ($p = 2$ and $n > 2$)

For $n > 2$, we introduce the notion of solution for measure data introduced by G. Stampacchia in [Sta65, Definition 9.1].

Definition 4.12.1. *Given $\mu \in \mathcal{B}(\Omega)$ with finite variation, we say $u \in L^1(\Omega)$ is a* stampacchia solution *of*

$$\begin{cases} -\operatorname{div}(A(x)\nabla u) &= \mu \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega \end{cases} \tag{4.12.1}$$

whenever

$$\int_\Omega uf\,dx = \int_\Omega v\,d\mu, \quad \forall f \in L^\infty(\Omega),$$

where v solves

$$\begin{cases} -\operatorname{div}(A^t(x)\nabla v) &= f \quad \text{in } \Omega \\ v &= 0 \quad \text{on } \partial\Omega \end{cases} \tag{4.12.2}$$

The existence of v follows from Definition 4.3.3 and, by classical regularity results, v is in the class of Hölder continuous functions. The existence and uniqueness of stampacchia solution was shown in [Sta65, Theorem 9.1]. Also, one has

$$\|u\|_{W_0^{1,r}(\Omega)} \le C_0|\mu|, \quad \forall r \in \left[1, \frac{n}{n-1}\right)$$

where the constant C_0 depends on n, α and Ω.

4.12.2 Renormalized Solutions ($1 \le p < n$)

For every $k > 0$, we define the truncation function (truncation at height k) $T_k : \mathbb{R} \to \mathbb{R}$ by

$$T_k(s) = \begin{cases} s & \text{if } |s| \le k \\ \frac{ks}{|s|} & \text{if } |s| \ge k. \end{cases}$$

An equivalent way of stating the truncation at height k is $T_k(s) = \max(-k, \min(k, s))$.

Let $\mathcal{T}_0^{1,p}(\Omega)$ be the set all measurable functions $u : \Omega \to \overline{\mathbb{R}}$ on Ω and finite almost everywhere in Ω, such that $T_k(u) \in W_0^{1,p}(\Omega)$ for every $k > 0$. Every function $u \in \mathcal{T}_0^{1,p}(\Omega)$ will be identified with its p-quasi continuous representative.

Lemma 4.12.2 (Lemma 2.1 of [BBG+95]). *For every $u \in \mathcal{T}_0^{1,p}(\Omega)$, there exists a measurable function $v : \Omega \to \mathbb{R}^n$ such that $\nabla T_k(u) = v\chi_{\{|u| \leq k\}}$ almost everywhere in Ω, for every $k > 0$. Also v is unique up to almost everywhere equivalence.*

Thus, by above lemma, one can define a generalised gradient ∇u of $u \in \mathcal{T}_0^{1,p}(\Omega)$ by setting $\nabla u = v$. If $u \in W^{1,1}(\Omega)$, then this gradient coincides with the usual one. However, for $u \in L^1_{\text{loc}}(\Omega)$, then it may differ from the distributional gradient of u (cf. Remark 2.9 of [DMMOP99]).

Let $\text{Lip}_0(\mathbb{R})$ be the set of all Lipschitz continuous functions $l : \mathbb{R} \to \mathbb{R}$ whose derivative l' has compact support. Thus, every function $l \in \text{Lip}_0(\mathbb{R})$ is constant outside the support of its derivative, so that we can define the constants $l(+\infty) = \lim_{s\to+\infty} l(s)$ and $l(-\infty) = \lim_{s\to-\infty} l(s)$.

Appendices

Appendix A

Cantor Set and Cantor Function

Let us construct the Cantor set which plays a special role in analysis.

Consider $C_0 = [0, 1]$ and trisect C_0 and remove the middle open interval to get C_1. Thus, $C_1 = [0, 1/3] \cup [2/3, 1]$. Repeat the procedure for each interval in C_1, we get

$$C_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1].$$

Repeating this procedure at each stage, we get a sequence of subsets $C_i \subseteq [0, 1]$, for $i = 0, 1, 2, \ldots$. Note that each C_k is a compact subset, since it is a finite union of compact sets. Moreover,

$$C_0 \supset C_1 \supset C_2 \supset \ldots \supset C_i \supset C_{i+1} \supset \ldots.$$

The Cantor set C is the intersection of all the nested C_i's, $C = \cap_{i=0}^{\infty} C_i$.

Lemma A.0.1. *C is compact.*

Proof. C is countable intersection of closed sets and hence is closed. $C \subset [0, 1]$ and hence is bounded. Thus, C is compact. □

The Cantor set C is non-empty, because the end-points of the closed intervals in C_i, for each $i = 0, 1, 2, \ldots$, belong to C. In fact, the Cantor set cannot contain any interval of positive length.

Lemma A.0.2. *For any $x, y \in C$, there is a $z \notin C$ such that $x < z < y$. (Disconnected)*

Proof. If $x, y \in C$ are such that $z \in C$ for all $z \in (x, y)$, then we have the open interval $(x, y) \subset C$. It is always possible to find i, j such that

$$\left(\frac{j}{3^i}, \frac{j+1}{3^i}\right) \subseteq (x, y)$$

but does not belong $C_i \supset C$. $\square$

We show in example **??**, that C has length zero. Since C is non-empty, how 'big' is C? The number of end-points sitting in C are countable. But C has points other than the end-points of the closed intervals C_i for all i. For instance, $1/4$ (not an end-point) will never belong to the the intervals being removed at every step i, hence is in C. There are more! $3/4$ and $1/13$ are all in C which are not end-points of removed intervals. It is easy to observe these by considering the ternary expansion characterisation of C. Consider the ternary expansion of every $x \in [0, 1]$,

$$x = \sum_{i=1}^{\infty} \frac{a_i}{3^i} = 0.a_1a_2a_3\ldots_3 \quad \text{where } a_i = 0, 1 \text{ or } 2.$$

The decomposition of x in ternary form is not unique[1]. For instance, $1/3 = 0.1_3 = 0.022222\ldots_3$, $2/3 = 0.2_3 = 0.1222\ldots_3$ and $1 = 0.222\ldots_3$. At the C_1 stage, while removing the open interval $(1/3, 2/3)$, we are removing all numbers whose first digit in ternary expansion (in all possible representations) is 1. Thus, C_1 has all those numbers in $[0, 1]$ whose first digit in ternary expansion is not 1. Carrying forward this argument, we see that for each i, C_i contains all those numbers in $[0, 1]$ with digits upto ith place, in ternary expansion, is not 1. Thus, for any $x \in C$,

$$x = \sum_{i=1}^{\infty} \frac{a_i}{3^i} = 0.a_1a_2a_3\ldots_3 \quad \text{where } a_i = 0, 2.$$

Lemma A.0.3. *C is uncountable.*

Proof. Use Cantor's diagonal argument to show that the set of all sequences containing 0 and 2 is uncountable. $\square$

[1] This is true for any positional system. For instance, $1 = 0.99999\ldots$ in decimal system

Cantor Function

We shall now define the *Cantor function* $f_C : C \to [0,1]$ as,

$$f_C(x) = f_C\left(\sum_{i=1}^{\infty} \frac{a_i}{3^i}\right) = \sum_{i=1}^{\infty} \frac{a_i}{2} 2^{-i}.$$

Since $a_i = 0$ or 2, the function replaces all 2 occurrences with 1 in the ternary expansion and we interpret the resulting number in binary system. Note, however, that the Cantor function f_C is not injective. For instance, one of the representation of $1/3$ is $0.0222\ldots_3$ and $2/3$ is 0.2. Under f_C they are mapped to $0.0111\ldots_2$ and 0.1_2, respectively, which are different representations of the same point. Since f_C is same on the end-points of the removed interval, we can extend f_C to $[0,1]$ by making it constant along the removed intervals.

Alternately, one can construct the Cantor function step-by-step as we remove middle open intervals to get C_i. Consider f_1 to be a function which takes the constant value $1/2$ in the removed interval $(1/3, 2/3)$ and is linear on the remaining two intervals such that f_1 is continuous. In the second stage, the function f_2 coincides with f_1 in $1/3, 2/3$, takes the constant value $1/4$ and $3/4$ on the two removed intervals and is linear in the remaining four intervals such that f_2 is continuous. Proceeding this way we have a sequence of monotonically increasing continuous functions $f_k : [0,1] \to [0,1]$. Moreover, $|f_{k+1}(x) - f_k(x)| < 2^{-k}$ for all $x \in [0,1]$ and f_k converges uniformly to $f_C : [0,1] \to [0,1]$.

Exercise 72. The Cantor function $f_C : [0,1] \to [0,1]$ is uniformly continuous, monotonically increasing and is differentiable a.e. and $f'_C = 0$ a.e.

Exercise 73. The function f_C is not absolutely continuous.

Generalised Cantor Set

We generalise the idea behind the construction of Cantor sets to build Cantor-like subsets of $[0,1]$. Choose a sequence $\{a_k\}$ such that $a_k \in (0, 1/2)$ for all k. In the first step we remove the open interval $(a_1, 1-a_1)$ from $[0,1]$ to get C_1. Hence $C_1 = [0, a_1] \cup [1-a_1, 1]$. Let

$$C_1^1 := [0, a_1] \quad \text{and } C_1^2 := [1-a_1, 1].$$

Hence, $C_1 = C_1^1 \cup C_1^2$. Note that C_1^i are sets of length a_1 carved out from the end-points of C_0. We repeat step one for each of the end-points of C_1^i of length $a_1 a_2$. Therefore, we get four sets

$$C_2^1 := [0, a_1 a_2] \quad C_2^2 := [a_1 - a_1 a_2, a_1],$$
$$C_2^3 := [1 - a_1, 1 - a_1 + a_1 a_2] \quad C_2^4 := [1 - a_1 a_2, 1].$$

Define $C_2 = \cup_{i=1}^4 C_2^i$. Each C_2^i is of length $a_1 a_2$. Note that $a_1 a_2 < a_1$. Repeating the procedure successively for each term in the sequence $\{a_k\}$, we get a sequence of sets $C_k \subset [0,1]$ whose length is $2^k a_1 a_2 \dots a_k$. The "generalised" Cantor set C is the intersection of all the nested C_k's, $C = \cap_{k=0}^\infty C_k$ and each $C_k = \cup_{i=1}^{2^k} C_k^i$. Note that by choosing the constant sequence $a_k = 1/3$ for all k gives the Cantor set defined in the beginning of this Appendix. Similar arguments show that the generalised Cantor set C is compact. Moreover, C is non-empty, because the end-points of the closed intervals in C_k, for each $k = 0, 1, 2, \dots$, belong to C.

Lemma A.0.4. *For any $x, y \in C$, there is a $z \notin C$ such that $x < z < y$.*

Lemma A.0.5. *C is uncountable.*

We show in example **??**, that C has length $2^k a_1 a_2 \dots a_k$.

The interesting fact about generalised Cantor set is that it can have non-zero "length".

Proposition A.0.6. *For each $\alpha \in [0,1)$ there is a sequence $\{a_k\} \subset (0, 1/2)$ such that*

$$\lim_k 2^k a_1 a_2 \dots a_k = \alpha.$$

Proof. Choose $a_1 \in (0, 1/2)$ such that $0 < 2a_1 - \alpha < 1$. Use similar arguments to choose $a_k \in (0, 1/2)$ such that $0 < 2^k a_1 a_2 \dots a_k - \alpha < 1/k$. □

Generalised Cantor Function

We shall define the generalised Cantor function f_C on the generalised Cantor set C. Define the function $f_0 : [0,1] \to [0,1]$ as $f_0(x) = x$. f_0 is continuous on $[0,1]$. We define $f_1 : [0,1] \to [0,1]$ such that f is linear and continuous on C_1^i, and $1/2$ on $[a_1, 1 - a_1]$, the closure of removed open interval at first stage. We define $f_k : [0,1] \to [0,1]$ continuous $f_k(0) = 0, f_k(1) = 1$ such that $f_k(x) = i/2^k$ on the removed interval immediate right to C_k^i.

Theorem A.0.7. *Each f_k is continuous, monotonically non-decreasing and uniformly converges to some $f_C : [0,1] \to [0,1]$.*

Thus, f_C being uniform limit of continuous function is continuous and is the called the generalised Cantor function.

Bibliography

[Ada75] Robert A. Adams, *Sobolev spaces*, Academic Press, Boston, 1975.

[BBG+95] P Bénilan, L Boccardo, T Gallouët, R Gariepy, M Pierre, and J. L. Vazquez, *An L^1-theory of existence and uniqueness of solutions of nonlinear elliptic equations*, Ann. Scuola Normale Sup. Pisa Cl. Sci. **22** (1995), 241–273.

[CD44] G. Choquet and J. Deny, *Sur quelques propriétés de moyenne caractéristiques des fonctions harmoniques et polyharmoniques*, Bull. Soc. Math. de France **72** (1944), 118–140.

[Dir49] P. A. M. Dirac, *The principles of quantum mechanics*, 3 ed., University Press, Oxford, 1949.

[DMMOP99] Gianni Dal Maso, F. Murat, L. Orsina, and A. Prignet, *Renormalized solutions for elliptic equations with general measure data*, Ann. Scuola Normale Sup. Pisa Cl. Sci. **28** (1999), 741–808.

[Hor70] John Horvath, *An introduction to distributions*, American Mathematical Monthly **77** (1970), 227–240.

[Kes89] S. Kesavan, *Topics in functional analysis and applications*, New Age International, 1989, New edition/reprint might be avalaible.

[Lio69] J. L. Lions, *Quelques méthodes de résolution des problémes aux limites non linéaires*, Dunod et Gauthier-Villars, Paris, 1969.

[LL65] J. Leray and J. L. Lions, *Quelques résultats de visik sur les problémes elliptiques non linéaires par les méthodes de Minty-Browder*, Bull. Soc. Math. France **93** (1965), 97–107.

[Maz85] Vladimir Maz'ya, *Sobolev spaces*, Springer-Verlag, 1985.

[MS64] N. Meyers and J. Serrin, $h = w$, Proc. Nat. Acad. Sci., USA **51** (1964), 1055–1056.

[Rud91] Walter Rudin, *Functional analysis*, second ed., McGraw-Hill Book Company, 1991.

[Sch45] L. Schwartz, *Généralisation de la notion de fonction, de transformation de fourier et applications mathématiques et physiques*, Ann. Univ. Grenoble **21** (1945), 57–74.

[Sta65] Guido Stampacchia, *Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus*, Annales de l'institut Fourier **15** (1965), no. 1, 189–257.

[Tre70] F. Treves, *Applications of distributions to pde theory*, American Mathematical Monthly **77** (1970), 241–248.

Index