

# The Theory of Sobolev Spaces and its Applications to Second-Order PDEs

Jun Hao (Simon) Hu \*

August 13, 2018

## Abstract

In this paper, I motivate the development of the Sobolev space  $W^{k,p}(\Omega)$  for the purpose of solving partial differential equations (PDEs). Sobolev spaces possess excellent mathematical structure that allows us to use tools developed from functional analysis to study PDEs. Applications of Sobolev spaces to second-order elliptic and parabolic equations will be presented. In particular, I examine topological properties of solutions, strong and weak, to second-order elliptic and parabolic equations.

## Acknowledgments.

I would like to acknowledge my friends and family for their continued support in all of my academic endeavors. Additionally, I would like to acknowledge the tremendous support my instructors have provided me. Their support, criticism, and knowledge have made it possible for me to understand the theory of Sobolev spaces and present their applications in this manner. Furthermore, I am eternally grateful for the opportunity to study Sobolev spaces and their applications to PDEs.

---

\*Department of Electrical and Electronics Engineering, University of California at San Diego

# Contents

<b>1</b>	<b>Introduction.</b>	<b>3</b>
<b>2</b>	<b>Weak Derivatives.</b>	<b>3</b>
<b>3</b>	<b>Function Analytic Properties of Sobolev Spaces.</b>	<b>7</b>
3.1	Dense Subspaces of $W^{k,p}(\Omega)$ .	8
3.2	The Space $W^{1,p}(\Omega)$ .	10
3.3	Extensions of Sobolev Spaces.	10
3.4	Traces.	13
3.5	Sobolev Embeddings.	15
3.6	Compactness.	20
<b>4</b>	<b>Second-Order Elliptic Equations.</b>	<b>20</b>
4.1	Motivation for Weak Solutions.	21
4.2	Existence of Weak Solutions.	22
<b>5</b>	<b>Second-Order Parabolic Equations.</b>	<b>23</b>
5.1	Weak Solutions.	23
<b>6</b>	<b>Appendix.</b>	<b>23</b>
6.1	Results from Functional Analysis.	23
6.2	The Space $H^{-1}$ .	25
6.3	Mollification Theory.	26
<b>7</b>	<b>References.</b>	<b>28</b>

# 1 Introduction.

It is widely accepted that PDEs are not easy to solve, yet they are useful in almost every discipline of study. When analysts solve a PDE problem, they are not interested in solutions that lack a specified degree of smoothness. This is due to practicality, and the application it is being used for. Unfortunately, it is rarely the case that we are able to find such smooth solutions, at least for PDEs with interesting applications. It is in these cases, that analysts turn to numerical methods to solve the problem. Common numerical methods perform a discretization of the domain and then reduce the problem to solving a matrix equation. A deep understanding of the function space in which the solution(s) lie is critical for developing an intuition for the expected outcome of the numerical method. Without such intuition, it would be difficult to confirm our results. Numerical solutions to PDE are often enough for applications in engineering, biology, and physics, but they lack differentiability, continuity, and other topological properties. Analysts are often interested in solutions that have these properties.

It turns out, these solutions often exist but it requires us to *weaken* the notion of a derivative. The theory of Sobolev spaces introduces the concept of *weak* derivatives, which generalize the notion of a derivative by moving away from the classical, or *strong*, derivative that we are used to in undergraduate calculus courses. The theory of Sobolev spaces give rise to the concept of a *weak* solution to the PDE, which are aptly named after the weak derivative. To understand these weak solutions, we will borrow ideas from functional analysis. Sobolev spaces are useful in this sense, because we can use tools from functional analysis to shine some light onto the PDE problem.

In this paper, I assume that the reader has already had some exposure to functional analysis. Furthermore, I assume that the reader has background knowledge on the theories of Banach, Hilbert, and  $L^p$  spaces. A rudimentary introduction to mollifiers is helpful in understanding the material, though not required. If the reader requires a review, they should consult (insert reference here). Standard results, like inequalities and elementary theorems from functional analysis, may not be proved in the paper. However, a select few of these results will be proved in the appendix. For example, I will make use of the Minkowski, Young, and Hölder's inequalities, but they will not be proved in the paper. Rather, they appear in the appendix. If the reader would like proof, they should consult any standard analysis text. Finally, I assume the reader has previous exposure to PDE theory, of which an introductory course will suffice.

This paper is divided into five sections, not withstanding this one. In section 2, we explore the concept of a weak derivative and familiarize the reader with the notation used in this paper. Furthermore, in this section, we introduce the Sobolev space. In section 3, we explore the function analytic properties of the Sobolev space, which is what makes it so useful for PDEs, and discuss properties of weak derivatives. Furthermore, we will discuss approximations and extensions, theories which will be useful in our later discussions. The proofs in these sections are long and detailed; the reader should keep this in mind. In section 4, we discuss second-order elliptic PDEs. In section 5, we discuss parabolic PDEs. In both sections 4 and 5, we will use the theory of Sobolev spaces to derive fundamental results that are important in the discussion of PDE.

## 2 Weak Derivatives.

*Remark 2.1.* For the remainder of this paper, let  $\Omega \subset \mathbb{R}^n$  be a bounded and open set.

We begin by first introducing the notation for multiindices, which will be useful for the remainder of the paper.

**Definition 2.1.** Let  $\alpha$  be a multiindex such that  $|\alpha| \leq k$ , the degree of the index. Then we define the partial differential operator

$$D^\alpha(u) := \frac{d^\alpha}{dx_1^{\alpha_1} dx_2^{\alpha_2} \dots dx_n^{\alpha_n}} u.$$

**Definition 2.2.** Let  $\alpha, \beta$  be any two multiindices of degree  $k$ . Then we define the following:

1.  $\alpha \leq \beta$  if  $\alpha_i \leq \beta_i$  for all  $1 \leq i \leq n$ .
2. If  $\alpha \leq \beta$ , define the difference as  $\alpha - \beta := \gamma = (\alpha_1 - \beta_1, \dots, \alpha_n - \beta_n)$ .
3. Define the factorial of a multiindex as  $\alpha! = \alpha_1! \dots \alpha_n!$ .

We now define the *test functions*, which will be used to motivate weak derivatives.

**Definition 2.3.** Let  $C_c^\infty(\Omega)$  denote the space of all infinitely differentiable functions  $\phi : \Omega \rightarrow \mathbb{R}$  that have compact support in  $\Omega$ . We call this  $\phi \in C_c^\infty(\Omega)$  a test function.

Suppose that we are given a function  $u \in C^1(\Omega)$  and a test function  $\phi$ . Then, we notice that

$$\int_{\Omega} u D\phi \, dx = - \int_{\Omega} Du \phi \, dx.$$

The above result comes from integration by parts. Since  $\phi$  has compact support in  $\Omega$ , it vanishes near the boundary,  $\partial\Omega$ . The above result can be generalized to functions that are  $k$  times continuously differentiable. Let  $u \in C^k(\Omega)$ ,  $\alpha$  be a multiindex of degree  $k$ . Integrating by parts, we can see that

$$\int_{\Omega} u D^{\alpha} \phi \, dx = (-1)^k \int_{\Omega} D^{\alpha} u \phi \, dx. \quad (2.1)$$

We require that  $u \in L_{\text{loc}}(\Omega)$  so that the left hand side of the above equality holds for every test function  $\phi$ . The right hand side on the other hand, does not pose an issue as long as  $u \in C^k(\Omega)$ . The concept of a weak derivative arises when we seek functions  $u$ , that are not in  $C^k(\Omega)$ , such that the right hand side is valid. If we seek said functions, a fundamental issue must be tackled: how do we define the term  $D^{\alpha}u$  if  $u$  is not  $k$  times continuously differentiable? We take advantage of this ambiguity to define the weak derivative. In the sequel, we assume that there exists a function  $v \in L_{\text{loc}}(\Omega)$  such that  $D^{\alpha}u = v$ , and use this assumption to define the weak derivative.

**Definition 2.4.** Let  $u, v \in L_{\text{loc}}(\Omega)$ ,  $\alpha$  be a multiindex of degree  $k$ , and  $\phi \in C_c^{\infty}(\Omega)$  be a test function. We call  $v$  the  $\alpha$ -th partial derivative of  $u$  if

$$\int_{\Omega} u D^{\alpha} \phi \, dx = (-1)^k \int_{\Omega} v \phi \, dx$$

for all test functions  $\phi$ .

To make notation easier, we will discuss properties of weak derivatives *after* we define the Sobolev spaces. In the sequel, fix  $1 \leq p \leq \infty$  and let  $k \in \mathbb{N}$ . We now define the Sobolev spaces, which are function spaces whose functions have weak derivatives up to certain orders.

**Definition 2.5.** The Sobolev space  $W^{k,p}(\Omega)$  is the space consisting of all functions  $u : \Omega \rightarrow \mathbb{R}$ ,  $u \in L_{\text{loc}}^p(\Omega)$  such that for each multiindex  $\alpha$  with  $|\alpha| \leq k$ , the weak derivative  $D^{\alpha}u$  exists and is an element of  $L^p(\Omega)$ . In other words,

$$W^{k,p}(\Omega) = \left\{ u \in L_{\text{loc}}^p(\Omega) : \|D^{\alpha}u\|_{L^p(\Omega)} < \infty \, \forall \, |\alpha| \leq k \right\}.$$

In a similar manner, we can define the local Sobolev space.

**Definition 2.6.** Let  $\omega \subset \subset \Omega$ . Let  $u : \Omega \rightarrow \mathbb{R}$  and  $\alpha$  satisfy the same assumptions made in the previous definition. Then the local Sobolev space  $W_{\text{loc}}^{k,p}(\Omega)$  is the space consisting of all  $u$  such that the weak derivative  $D^{\alpha}u$  exists and is an element of  $L^p(\omega)$ , for any choice of  $\omega$ . In other words,

$$W_{\text{loc}}^{k,p}(\Omega) = \left\{ u \in L_{\text{loc}}^p(\Omega) : \|D^{\alpha}u\|_{L^p(\omega)} < \infty \, \forall \, |\alpha| \leq k, \text{ for any arbitrary } \omega \right\}.$$

With the Sobolev spaces defined, we can now explore properties of weak derivatives, which will be explored in no particular order.

**Theorem 2.1.** (*Essential properties of weak derivatives*). Let  $u, v \in W^{k,p}(\Omega)$  and let  $\alpha$  be a multiindex such that  $|\alpha| \leq k$ . Then

1. The weak derivative, if it exists, is unique.
2. For each  $\lambda, \mu \in \mathbb{R}$ ,  $\lambda u + \mu v \in W^{k,p}(\Omega)$ . Furthermore,  $D^{\alpha}(\lambda u + \mu v) = \lambda D^{\alpha}u + \mu D^{\alpha}v$ .
3.  $D^{\alpha}u \in W^{k-|\alpha|,p}(\Omega)$ .
4.  $D^{\alpha}(D^{\beta}u) = D^{\beta}(D^{\alpha}u)$  provided that  $|\alpha| + |\beta| \leq k$ .
5.  $D^{\alpha}(D^{\beta}u) = D^{\alpha+\beta}u$  provided that  $|\alpha| + |\beta| \leq k$ .
6. (*Leibniz' Formula*). Let  $\xi \in C_c^{\infty}(\Omega)$  be a smooth and compactly supported function. Then  $\xi u \in W^{k,p}(\Omega)$  and the explicit formula is given by

$$D^{\alpha}(\xi u) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} D^{\beta} \xi D^{\alpha-\beta} u \quad (2.2)$$

which is exactly Leibniz's formula for weak derivatives.

*Proof.* (1). Let  $w, \bar{w} \in L_{\text{loc}}(\Omega)$  be two weak derivatives of  $u$ . Then

$$\int_{\Omega} u D^{\alpha} \phi \, dx = (-1)^{|\alpha|} \int_{\Omega} w \phi \, dx = (-1)^{|\alpha|} \int_{\Omega} \bar{w} \phi \, dx$$

and thus

$$\int_{\Omega} (w - \bar{w}) \phi \, dx = 0$$

whence  $(w - \bar{w}) = 0$  for all measurable sets  $\Omega$ .

(2). First, we show the linearity property. We write,

$$\begin{aligned} \int_{\Omega} (\lambda u + \mu v) D^{\alpha} \phi \, dx &= \int_{\Omega} \lambda u D^{\alpha} \phi \, dx + \int_{\Omega} \mu v D^{\alpha} \phi \, dx \\ &= (-1)^{|\alpha|} \int_{\Omega} \lambda D^{\alpha} u \phi \, dx + (-1)^{|\alpha|} \int_{\Omega} \mu D^{\alpha} v \phi \, dx \\ &= (-1)^{|\alpha|} \int_{\Omega} (\lambda D^{\alpha} u + \mu D^{\alpha} v) \phi \, dx \end{aligned}$$

by the linearity of integration. By integration by parts, we also write

$$\int_{\Omega} (\lambda u + \mu v) D^{\alpha} \phi \, dx = (-1)^{|\alpha|} \int_{\Omega} D^{\alpha} (\lambda u + \mu v) \phi \, dx.$$

Hence, since the above two equalities hold for all test functions  $\phi$ , we have

$$D^{\alpha} (\lambda u + \mu v) = \lambda D^{\alpha} u + \mu D^{\alpha} v$$

in the weak sense, which is the desired result. Next, we show that the function  $\lambda u + \mu v$  is an element of  $W^{k,p}$ . It is trivial to conclude that if  $u$  and  $v$  are both locally integrable functions, then  $\lambda u + \mu v$  is also a locally integrable function. To show that the weak derivative is bounded with respect to the  $L^p$  norm, consider

$$\begin{aligned} \|D^{\alpha} (\lambda u + \mu v)\|_{L^p(\Omega)} &= \|\lambda D^{\alpha} u + \mu D^{\alpha} v\|_{L^p(\Omega)} \\ &\leq |\lambda| \|D^{\alpha} u\|_{L^p(\Omega)} + |\mu| \|D^{\alpha} v\|_{L^p(\Omega)} \\ &< \infty \end{aligned}$$

where the inequality is due to Minkowski's inequality. The strict inequality follows from the assumption that  $u, v \in W^{k,p}$ .

(3) and (4) are easy to prove, and will be excluded.

(5). We write

$$\begin{aligned} \int_{\Omega} D^{\beta} u D^{\alpha} \phi \, dx &= (-1)^{|\beta|} \int_{\Omega} u D^{\alpha+\beta} \phi \, dx \\ &= (-1)^{|\beta|} (-1)^{|\alpha+\beta|} \int_{\Omega} D^{\alpha+\beta} u \phi \, dx \\ &= (-1)^{|\alpha|} \int_{\Omega} D^{\alpha+\beta} u \phi \, dx \end{aligned}$$

where the last step is obtained by observing that  $|\alpha| + |\alpha + \beta| = 2|\alpha| + |\beta|$ , and since  $2|\alpha|$  is always an even number, we conclude that  $(-1)^{2|\alpha|} (-1)^{|\beta|} = (-1)^{|\beta|}$ . Thus we conclude that  $D^{\alpha} (D^{\beta} u) = D^{\alpha+\beta} u$  in the weak sense.

(6). Let  $\phi$  be a test function. We proceed by induction on  $|\alpha|$ . For the base case,  $|\alpha| = 1$ , we observe that

$$\int_{\Omega} \xi u D^{\alpha} \phi \, dx = \int_{\Omega} u D^{\alpha} (\xi \phi) \, dx - \int_{\Omega} u D^{\alpha} \xi \phi \, dx$$

which comes from applying Leibniz's formula for two smooth functions (since  $\xi$  and  $\phi$  are both smooth functions). Finally, since  $\phi$  has compact support, the second integral above can be simplified using integration by parts, to obtain

$$\begin{aligned} \int_{\Omega} \xi u D^{\alpha} \phi \, dx &= - \int_{\Omega} \xi D^{\alpha} u \phi \, dx - \int_{\Omega} u D^{\alpha} \xi \phi \, dx \\ &= - \int_{\Omega} (\xi D^{\alpha} u + u D^{\alpha} \xi) \phi \, dx. \end{aligned}$$

Hence,  $D^{\alpha} (\xi u) = \xi D^{\alpha} u + u D^{\alpha} \xi$  in the weak sense. Thus, the base case is proved.

We proceed to prove the induction step. Let  $\ell < k$  and assume that the Leibniz formula holds for all  $|\alpha| \leq \ell$ . Choose some multiindex  $\alpha = \ell + 1$  and let  $|\beta| = \ell$  and  $|\gamma| = 1$  be two multiindices, so that  $|\alpha| = |\beta| + |\gamma|$ . Then,

$$\begin{aligned} \int_{\Omega} \xi u D^{\alpha} \phi \, dx &= \int_{\Omega} \xi u D^{\beta} (D^{\gamma} \phi) \, dx \\ &= (-1)^{|\beta|} \int_{\Omega} \sum_{\sigma \leq \beta} \binom{\beta}{\sigma} D^{\sigma} \xi D^{\beta-\sigma} u D^{\alpha} \phi \, dx \quad (\text{induction hypothesis}) \end{aligned}$$

$$\begin{aligned}
&= (-1)^{|\beta|+|\gamma|} \int_{\Omega} \sum_{\sigma \leq \beta} \binom{\beta}{\sigma} D^{\gamma} (D^{\sigma} \xi D^{\beta-\sigma} u) \phi \, dx \quad (\text{definition}) \\
&= (-1)^{|\alpha|} \int_{\Omega} \sum_{\sigma \leq \beta} \binom{\beta}{\sigma} D^{\gamma} (D^{\sigma} \xi D^{\beta-\sigma} u) \phi \, dx \quad (|\alpha| = |\beta| + |\gamma|) \\
&= (-1)^{|\alpha|} \int_{\Omega} \sum_{\sigma \leq \beta} \binom{\beta}{\sigma} \left[ D^{\sigma+\gamma} \xi D^{\beta-\sigma} u + D^{\sigma} \xi D^{\beta-\sigma+\gamma} u \right] \phi \, dx
\end{aligned}$$

where we have again used the induction hypothesis. There is a common term  $\sigma + \gamma$ , so we will define  $\rho = \sigma + \gamma$  for convenience. Thus, the above can be rewritten as

$$\begin{aligned}
\int_{\Omega} \xi u D^{\alpha} \phi \, dx &= (-1)^{|\alpha|} \int_{\Omega} \sum_{\sigma \leq \beta} \binom{\beta}{\sigma} \left[ D^{\rho} \xi D^{\alpha-\rho} u + D^{\sigma} \xi D^{\alpha-\sigma} u \right] \phi \, dx \\
&= (-1)^{|\alpha|} \int_{\Omega} I \phi \, dx.
\end{aligned}$$

Focusing only on the integrand  $I$ , we observe

$$\begin{aligned}
I &= \sum_{\sigma \leq \beta} \binom{\beta}{\sigma} D^{\rho} \xi D^{\alpha-\rho} u + \sum_{\sigma \leq \beta} \binom{\beta}{\sigma} D^{\sigma} \xi D^{\alpha-\sigma} u \\
&= \sum_{\rho \leq \alpha} \binom{\beta}{\rho-\gamma} D^{\rho} \xi D^{\alpha-\rho} u + \sum_{\sigma \leq \beta} \binom{\beta}{\sigma} D^{\sigma} \xi D^{\alpha-\sigma} u \\
&= \sum_{\sigma \leq \alpha} \binom{\beta}{\sigma-\gamma} D^{\sigma} \xi D^{\alpha-\sigma} u + \sum_{\sigma \leq \beta} \binom{\beta}{\sigma} D^{\sigma} \xi D^{\alpha-\sigma} u
\end{aligned}$$

where we have changed the dummy variable from  $\rho$  to  $\sigma$ . Next, to collect like-terms, we pop out a term in the first summation. The second summation is left untouched. Thus,

$$\begin{aligned}
I &= \binom{\beta}{\beta} D^{\alpha} \xi D^0 u + \sum_{\sigma \leq \beta} \binom{\beta}{\sigma-\gamma} D^{\sigma} \xi D^{\alpha-\sigma} u + \sum_{\sigma \leq \beta} \binom{\beta}{\sigma} D^{\sigma} \xi D^{\alpha-\sigma} u \\
&= \binom{\beta}{\beta} D^{\alpha} \xi D^0 u + \sum_{\sigma \leq \beta} \left[ \binom{\beta}{\sigma-\gamma} + \binom{\beta}{\sigma} \right] D^{\sigma} \xi D^{\alpha-\sigma} u \\
&= \binom{\beta}{\beta} D^{\alpha} \xi D^0 u + \sum_{\sigma \leq \beta} \binom{\alpha}{\sigma} D^{\sigma} \xi D^{\alpha-\sigma} u
\end{aligned}$$

where we have used the elementary result

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

which is the recursive formula for the binomial coefficients. To complete the proof, we combine the terms with the summation. The loose term is exactly the summation evaluated when  $\sigma = \alpha$ . This justifies the combination. Thus,

$$I = \sum_{\sigma \leq \alpha} \binom{\alpha}{\sigma} D^{\sigma} \xi D^{\alpha-\sigma} u.$$

Finally, we can see that

$$\int_{\Omega} \xi u D^{\alpha} \phi \, dx = (-1)^{|\alpha|} \int_{\Omega} \sum_{\sigma \leq \alpha} \binom{\alpha}{\sigma} D^{\sigma} \xi D^{\alpha-\sigma} u \phi \, dx$$

which implies that

$$D^{\alpha}(\xi u) = \sum_{\sigma \leq \alpha} \binom{\alpha}{\sigma} D^{\sigma} \xi D^{\alpha-\sigma} u$$

which is the desired result. With an explicit formula for the weak derivative in hand, it remains to show that, with respect to the  $L^p(\Omega)$  norm, the weak derivative is bounded. For this portion of the proof, an application of Minkowski's inequality gives the desired result.  $\square$

### 3 Function Analytic Properties of Sobolev Spaces.

We move to examine the mathematical structure that the Sobolev space is equipped with, in particular it's function analytic properties. The Sobolev space is a linear vector space but it will be useful to view Sobolev spaces as *normed* linear vector spaces. To motivate this development, we will eventually show that the Sobolev space is indeed, for each  $1 \leq p \leq \infty$  and  $k = 1, \dots$ , a Banach space.

To begin our examination, we first define the norm that will be used in the sequel.

**Definition 3.1.** The norm of  $u \in W^{k,p}(\Omega)$  is defined in the following manner.

$$\|u\|_{W^{k,p}(\Omega)} := \begin{cases} \left( \sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^p(\Omega)}^p \right)^{1/p}, & 1 \leq p < \infty \\ \max_{|\alpha| \leq k} \left( \sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^\infty(\Omega)} \right), & p = \infty. \end{cases}$$

Let us confirm that this is indeed a norm.

*Proof.* Let  $u, v \in W^{k,p}(\Omega)$ . It is quite clear that the above satisfies the positive definite property, since it relies on the  $L^p$  norm. Furthermore, it is also clear that the above satisfies the absolute homogeneity property. The last property we need to show is the triangle inequality. For  $1 \leq p \leq \infty$ , the Minkowski's inequality implies

$$\begin{aligned} \|u + v\|_{W^{k,p}(\Omega)} &= \left( \sum_{|\alpha| \leq k} \|D^\alpha u + D^\alpha v\|_{L^p(\Omega)}^p \right)^{1/p} \\ &\leq \left( \sum_{|\alpha| \leq k} \left\{ \|D^\alpha u\|_{L^p(\Omega)}^p + \|D^\alpha v\|_{L^p(\Omega)}^p \right\} \right)^{1/p} \\ &\leq \left( \sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^p(\Omega)}^p \right)^{1/p} + \left( \sum_{|\alpha| \leq k} \|D^\alpha v\|_{L^p(\Omega)}^p \right)^{1/p} \\ &= \|u\|_{W^{k,p}(\Omega)} + \|v\|_{W^{k,p}(\Omega)}. \end{aligned}$$

Hence, our definition of the norm is valid. □

In order to show that the Sobolev space is a Banach space, we need to define convergence.

**Definition 3.2.** Let  $u \in W^{k,p}(\Omega)$  and let  $\{u_m\}_{m=1}^\infty$  be a sequence of functions. We say that  $u_m \rightarrow u$  in  $W^{k,p}(\Omega)$  if

$$\lim_{m \rightarrow \infty} \|u_m - u\|_{W^{k,p}(\Omega)} = 0.$$

A similar definition for local convergence in the Sobolev space is given by

**Definition 3.3.** Let  $u \in W_{\text{loc}}^{k,p}(\Omega)$  and let  $\{u_m\}_{m=1}^\infty$  be a sequence of functions. We say that  $u_m \rightarrow u$  in  $W_{\text{loc}}^{k,p}(\Omega)$  if

$$\lim_{m \rightarrow \infty} \|u_m - u\|_{W_{\text{loc}}^{k,p}(\Omega)} = 0.$$

Furthermore, we can define the closure of the space in the following manner.

**Definition 3.4.** We let  $W_0^{k,p}(\Omega)$  denote the closure of  $C_c^\infty(\Omega)$  in  $W^{k,p}(\Omega)$ .  $W_0^{k,p}(\Omega)$  contains all functions  $u$  such that there exists a function  $u_m \in C_c^\infty(\Omega)$  such that  $u_m \rightarrow u$  in  $W^{k,p}(\Omega)$ .

Intuitively, the space  $W_0^{k,p}(\Omega)$  consists of all functions  $u \in W^{k,p}(\Omega)$  such that the  $\alpha^{th}$ -weak derivative vanishes at the boundary  $\partial\Omega$ , for all  $|\alpha| \leq k-1$ .

With the norm and convergence in  $W^{k,p}(\Omega)$  defined, we can now show that the Sobolev space is a Banach space.

**Theorem 3.1.** For each  $1 \leq p \leq \infty$  and  $k = 1, \dots$ , the Sobolev space  $W^{k,p}(\Omega)$  is a Banach space.

*Proof.* Let  $\{u_m\}_{m=1}^\infty$  be a Cauchy sequence. Since the norm on the Sobolev space is defined with respect to the  $L^p$  norm, it is clear that  $\{D^\alpha u_m\}_{m=1}^\infty$  is a Cauchy sequence in  $L^p(\Omega)$ . Now since  $L^p(\Omega)$  is a *complete* space, there exists a function  $u_\beta \in L^p(\Omega)$  such that

$$D^\alpha u_m \rightarrow u_\beta \text{ in } L^p(\Omega)$$

for each multiindex  $\alpha$  such that  $|\alpha| \leq k$ . In particular, there exists a multiindex  $\alpha = (0, \dots, 0)$  such that, for  $u \in L^p(\Omega)$ ,

$$u_m \rightarrow u_\alpha =: u \text{ in } L^p(\Omega).$$

Let us furthermore assume that  $u \in W^{k,p}(\Omega)$ . Fixing some test function  $\phi \in C_c^\infty(\Omega)$ , we find that

$$\begin{aligned} \int_{\Omega} u D^\alpha \phi \, dx &= \lim_{m \rightarrow \infty} \int_{\Omega} u_m D^\alpha \phi \, dx \\ &= \lim_{m \rightarrow \infty} (-1)^{|\alpha|} \int_{\Omega} D^\alpha u_m \phi \, dx \\ &= (-1)^{|\alpha|} \int_{\Omega} u_\beta \phi \, dx \end{aligned}$$

for each multiindex  $\alpha$  such that  $|\alpha| \leq k$ . Thus,

$$D^\alpha u = u_\beta \text{ in } W^{k,p}(\Omega).$$

From this, we conclude that  $D^\alpha u_m \rightarrow D^\alpha u$  and whenceforth  $u_m \rightarrow u$ . Thus, every Cauchy sequence in  $W^{k,p}(\Omega)$  converges to a vector that is also in  $W^{k,p}(\Omega)$ , and so the space is complete. Since, the Sobolev space is a complete normed space, it is therefore a Banach space as required.  $\square$

There is a deep connection between Sobolev spaces and Fourier Series. For the case where  $p = 2$ , we have that

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

for  $k = 1, \dots$ , where  $H^k(\Omega)$  is the space of all functions that are in  $L^2(\Omega)$  and whose weak derivatives up to an order  $k$  are also in  $L^2(\Omega)$ . The letter  $H$  is used because it is indeed a Hilbert space. The inner product with respect to  $H^k(\Omega)$  is given by

$$(f, g) = \int_{\Omega} f \bar{g} + Df \cdot D\bar{g} + D^2 f \cdot D^2 \bar{g} + \dots + D^k f \cdot D^k \bar{g} \, dx.$$

Note that if  $k = 0$ , then we have the classic  $L^2(\Omega)$  space. For many physics and engineering students, when studying the Fourier series and Fourier transforms, the space  $L^2(\Omega)$  space is commonly used as the underlying candidate space.

### 3.1 Dense Subspaces of $W^{k,p}(\Omega)$ .

We move to construct dense subspaces of  $W^{k,p}(\Omega)$ . By constructing dense subspaces, we will be able to approximate functions in the Sobolev space by elements of these dense subspaces. In particular, we will construct dense subspaces whose elements are smooth functions. The tool of mollifiers provides the mechanics we need to construct these dense subspaces. If you require a refresher course on mollifiers, refer to the appendix.

*Remark 3.1.* If  $\Omega \subset \mathbb{R}^n$  is an open set, then let  $\Omega_\epsilon = \{x \in \Omega : \text{dist}(x, \partial\Omega) > \epsilon\}$ .

We start with a theorem that outlines a systematic procedure for local approximations by smooth functions.

**Theorem 3.2.** Assume that  $u \in W^{k,p}(\Omega)$  for  $1 \leq p < \infty$ . Then,

1.  $u^\epsilon \in C^\infty(\Omega_\epsilon)$  for all  $\epsilon > 0$ .
2.  $u^\epsilon \rightarrow u$  in  $W_{\text{loc}}^{k,p}(\Omega)$ , as  $\epsilon \rightarrow 0$ .

*Proof.* (1). Proof of this statement is provided in the appendix.

(2). Proving this statement is equivalent to proving that

$$\|u^\epsilon - u\|_{W_{\text{loc}}^{k,p}(\Omega)}^p \rightarrow 0$$

as  $\epsilon \rightarrow 0$ . We first assert that the ordinary  $\alpha^{th}$  derivative of the smooth function  $u^\epsilon$  is the  $\epsilon$ -mollification of the  $\alpha^{th}$ -weak derivative of  $u$ . That is, we wish to show that

$$D^\alpha u^\epsilon = \eta_\epsilon * D^\alpha u \text{ in } \Omega_\epsilon.$$

To confirm this assertion, note that

$$D^\alpha u^\epsilon = D^\alpha \int_{\Omega} \eta_\epsilon(x-y) f(y) \, dy$$



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

for some multiindex  $\alpha$ . Note that by fixing  $x \in \Omega_{\epsilon}$ , we have that  $\eta_{\epsilon} \in C_c^{\infty}(\Omega)$ . Therefore, it is a test function. Continuing the above work, we have

$$\int_{\Omega} D_y^{\alpha} \eta_{\epsilon}(x-y) u(y) dy = (-1)^{|\alpha|} \int_{\Omega} \eta_{\epsilon}(x-y) D^{\alpha} u(y) dy.$$

Thus,

$$D^{\alpha} u^{\epsilon} = \int_{\Omega} \eta_{\epsilon}(x-y) D^{\alpha} u(y) dy = \eta_{\epsilon} * D^{\alpha} u$$

as desired. Finally, pick any  $\omega \subset \subset \Omega$ . Since  $D^{\alpha} u^{\epsilon} = \eta_{\epsilon} * D^{\alpha} u$  and  $u \in W^{k,p}(\Omega)$ , we conclude that  $D^{\alpha} u^{\epsilon} \rightarrow D^{\alpha} u$  in  $L^p(\omega)$ . Consequently,

$$\|u^{\epsilon} - u\|_{W^{k,p}(\omega)}^p = \sum_{|\alpha| \leq k} \|D^{\alpha} u^{\epsilon} - D^{\alpha} u\|_{L^p(\omega)}^p \rightarrow 0$$

as  $\epsilon \rightarrow 0$ , for all  $|\alpha| \leq k$ . Since this is true for *any* compact subset  $\omega$  of  $\Omega$ , we conclude that

$$\|u^{\epsilon} - u\|_{W_{\text{loc}}^{k,p}(\Omega)}^p \rightarrow 0$$

as  $\epsilon \rightarrow 0$ , which is the desired result.  $\square$

It turns out, we can also find smooth functions that approximate functions in  $W^{k,p}(\Omega)$ , and not just in  $W_{\text{loc}}^{k,p}(\Omega)$ . This is a nice property of Sobolev spaces as, in general, finding a global approximation is better than finding a local approximation. We will borrow ideas from our discussion of local approximations, to create an outline for global approximations. The following theorem sums this concept up.

**Theorem 3.3.** *Assume that  $\Omega$  is also bounded, and  $u \in W^{k,p}(\Omega)$  for some  $1 \leq p < \infty$ . Then there exists functions  $v \in C^{\infty}(\Omega) \cap W^{k,p}(\Omega)$  such that  $v \rightarrow u$  in  $W^{k,p}(\Omega)$ .*

*Proof.* Our goal here is to essentially prove density. We will use mollification to prove density. We proceed by creating a smooth partition of unity. Let us partition  $\Omega$  into different parts. Define

$$\Omega_i := \{x \in \Omega : \text{dist}(x, \partial\Omega) \leq \epsilon_i\}.$$

Let us write  $\omega_i = \Omega_{i+3} - \bar{\Omega}_{i+1}$ . Next, choose some open set  $\omega_0 \subset \subset \Omega$  so that

$$\Omega = \bigcup_{i=1}^{\infty} \omega_i.$$

Let us then define  $\{\xi_i\}_{i=0}^{\infty}$  as the smooth partition of unity, subject to the conditions

$$\begin{cases} 0 \leq \xi_i \leq 1, & \text{for } i = 0, 1, \dots \\ \xi_i \in C_c^{\infty}(\omega_i), & \text{for } i = 0, 1, \dots \\ \sum_{i=0}^{\infty} \xi_i = 1, & \text{on } \Omega. \end{cases}$$

Pick up some  $u \in W^{k,p}(\Omega)$ . Recall, from our discussion of properties of weak derivatives, if  $\xi_i$  is a function with compact support, then  $\xi_i u \in W^{k,p}(\Omega)$ . Furthermore,  $\text{supp}(\xi_i u) \subset \omega_i$ . Define the function  $u^i = \eta_{\epsilon_i} * \xi_i u$  as the mollification of  $\xi_i u$  by  $\eta_{\epsilon_i}$ .

Define the set  $\Gamma_i = \Omega_{i+4} - \bar{\Omega}_i$ . Observe first, that  $\omega_i \subset \gamma_i$  for all  $i = 1, \dots$ . We choose the  $\gamma_i$  so that there is enough room to perform the mollification process. We know that the function  $\xi_i$  has support in  $\omega_i$ , but we cannot say that  $u^i$  has support in  $\omega_i$ . In fact, we require a slightly larger set, for which the set  $\gamma_i$  does the job.

Now that we have determined that the function  $u^i$  has support in  $\Gamma_i$ , we can use the previous theorem to determine that there exists some constant  $C, \delta > 0$  and  $\epsilon = \epsilon(C, \delta)$  such that

$$\|u^i - \xi_i u\|_{W^{k,p}(\gamma_i)} \leq C\delta.$$

Fix some  $\delta > 0$  and choose an  $\epsilon_i > 0$  small enough so that  $u^i$  has support in  $\gamma_i$  for  $i = 1, \dots$ . By the previous theorem, we have that

$$\|u^i - \xi_i u\|_{W^{k,p}(\gamma_i)} \leq \frac{\delta}{3^{i+1}} \text{ for } i = 0, 1, \dots$$

where the  $C = \frac{1}{3^{i+1}}$  was chosen for convenience. Let us now define the function  $v$  as

$$v := \sum_{i=0}^{\infty} u^i.$$

We note that this function  $v \in C^\infty(\Omega) \cap W^{k,p}(\Omega)$  is a very smooth function since, for each open set  $\omega \subset \subset \Omega$ , there are at most finitely many nonzero terms in the summation. Additionally, we note that the function  $v$  is also in the Sobolev space, since each  $u^i$  is in the Sobolev space. Thus, for each compact subset  $\omega \subset \subset \Omega$ , we have that

$$\begin{aligned} \|v - u\|_{W^{k,p}(\omega)} &\leq \|v - u\|_{W^{k,p}(\Omega)} \\ &= \left\| \sum_{i=0}^{\infty} u^i - \sum_{i=0}^{\infty} \xi_i u \right\|_{W^{k,p}(\Omega)} \\ &\leq \sum_{i=0}^{\infty} \|u^i - \xi_i u\|_{W^{k,p}(\Omega)} \\ &\leq \delta \sum_{i=0}^{\infty} \frac{1}{3^{i+1}} \\ &= \delta. \end{aligned}$$

Completing the proof, we take the supremum over all sets  $\omega \subset \subset \Omega$  to conclude that

$$\|v - u\|_{W^{k,p}(\Omega)} \leq \delta$$

which is the desired result.  $\square$

In the previous theorem, note that we did not make any assumptions about the boundary. However, in the next theorem, we will show that we can approximate a given function  $u \in W^{k,p}(\Omega)$  not only by a function that is smooth in  $\Omega$ , but is smooth in the closure of  $\Omega$ . Such a claim requires that we assume the boundary is smooth enough.

**Theorem 3.4.** *Assume that  $\Omega$  is bounded and  $\partial\Omega$  is  $C^1$ . Suppose that  $u \in W^{k,p}(\Omega)$  for some  $1 \leq p < \infty$ . Then there exists functions  $v \in C^\infty(\bar{\Omega})$  such that*

$$u_m \rightarrow u \in W^{k,p}(\Omega).$$

The proof of this theorem is left to the appendix.

## 3.2 The Space $W^{1,p}(\Omega)$ .

For the remainder of the paper, we will derive results for the space  $W^{1,p}(\Omega)$ . These proofs can be extended to include the other values of  $k$ , but we will focus solely on the case  $k = 1$  because it is particularly useful for our study of elliptic second-order PDEs. Additionally, the space  $W^{1,p}(\Omega)$  possesses properties that do not hold if  $k \neq 1$ .

## 3.3 Extensions of Sobolev Spaces.

In all of our previous discussions, we assumed that  $\Omega \subset \mathbb{R}^n$ . However, in the following theorem, we will show that for  $k = 1$ , we can actually extend the domain to be  $\mathbb{R}^n$  instead. This is actually not an easy task to complete. An obvious extension that comes to mind is the extension where, if  $u \in W^{1,p}(\Omega)$ , we set  $u(x) := 0$  for all  $x \in \mathbb{R}^n \setminus \Omega$ . We create a discontinuity that removes the property of weak differentiability. Thus, the extended function is no longer an element of  $W^{1,p}(\Omega)$ .

Extension theorems are crucial tools for studying PDEs. These theorems allow us to extend a function to a larger space via a linear, bounded, operator. The method of proving a result for a dense subset, and then using an extension theorem to extend the result to a larger space, is common in the analysis of PDEs. In fact, we will use this idea when discussing ways to assign values to the boundary.

**Theorem 3.5.** Assume that  $\Omega$  is bounded and  $\partial\Omega$  is  $C^1$ . Let  $\omega$  be a bounded open set,  $\omega \subset\subset \Omega$ . There exists a bounded linear operator

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

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

1.  $Eu = u$  almost everywhere in  $\Omega$
2.  $\text{supp}(Eu) \subset \omega$
3.  $\|Eu\|_{W^{1,p}(\mathbb{R}^n)} \leq C\|u\|_{W^{1,p}(\Omega)}$  such that  $C = C(p, \omega, \Omega)$ .

We call  $Eu$  an extension of  $u$  to  $\mathbb{R}^n$  by  $E$ .

*Proof.* We will only consider the case when  $1 \leq p < \infty$ . For the case where  $p = \infty$ , we have the space of Lipschitz continuous functions, and the proof is monumentally easier. If the reader is interested in this case, then they should consult any text on Sobolev spaces.

Fix some  $\hat{x} \in \partial\Omega$  and, though we will later consider the other case, let us assume that  $\partial\Omega$  is flat near  $\hat{x}$  and the boundary lies in the plane  $\{x_n = 0\}$ . Then, we may assume that there exists some open ball  $B = B(\hat{x}, r)$  such that

$$\begin{cases} B^+ := B \cap \{x_n \geq 0\} \subset \Omega \\ B^- := B \cap \{x_n < 0\} \subset \mathbb{R}^n \setminus \Omega. \end{cases}$$

Our goal is to eventually show that  $u \in C^1(B)$ . Let us first suppose that  $u \in C^\infty(\bar{\Omega})$ . Using the method of higher-order reflections, we construct the function  $\bar{u}$  as

$$\bar{u}(x) := \begin{cases} u(x), & x \in B^+ \\ -\frac{5}{3}u(x_1, \dots, x_{n-1}, -x_n) + \frac{8}{3}u(x_1, \dots, x_{n-1}, -\frac{x_n}{4}), & x \in B^-. \end{cases}$$

This reflection was carefully constructed with the goal in mind. The reader should be aware that is is just one particular construction. Any other construction can work, but the construction should keep the goal in mind.

Proceeding with our goal, we introduce the two functions

$$u^+(x) := \bar{u}|_{B^+}, \quad u^-(x) := \bar{u}|_{B^-}.$$

Hence, if we can show that  $u_{x_n}^+ = u_{x_n}^-$  and  $u_{x_i}^+ = u_{x_i}^-$  for  $i = 1, \dots, n-1$  on  $\{x_n = 0\}$ , then the desired goal follows.

We confirm first that  $u_{x_n}^+ = u_{x_n}^-$  on  $\{x_n = 0\}$ . Indeed, by our definition of the reflection of  $u$ , we have

$$\frac{\partial u^-}{\partial x_n}(x) = \frac{5}{3} \frac{\partial u}{\partial x_n}(x_1, \dots, x_{n-1}, -x_n) - \frac{2}{3} \frac{\partial u}{\partial x_n}(x_1, \dots, x_{n-1}, -\frac{x_n}{4})$$

and by our definition of  $u^-$ , we observe

$$\frac{\partial u^-}{\partial x_n}(x) = \frac{\partial u}{\partial x_n}(x_1, \dots, x_{n-1}, x_n) = \frac{\partial u^+}{\partial x_n}.$$

This confirms our assumption.

Next, we note that for any  $i = 1, \dots, n-1$ , we observe that

$$\frac{\partial u^-}{\partial x_i}(x) = -\frac{5}{3} \frac{\partial u}{\partial x_i}(x_1, \dots, x_{n-1}, -x_n) + \frac{8}{3} \frac{\partial u}{\partial x_i}(x_1, \dots, x_{n-1}, -\frac{x_n}{4})$$

so that

$$u_{x_i}^- = u_{x_i}^+$$

on  $\{x_n = 0\}$ . This confirms our assumption.

Combining the above two results, it is easy to see that for any multiindex  $\alpha$  with  $|\alpha| \leq 1$ ,

$$D^\alpha u^-|_{\{x_n=0\}} = D^\alpha u^+|_{\{x_n=0\}}.$$

In the previous case, we assumed that  $\partial\Omega$  was flat. Let us consider the case where the boundary is not flat. Suppose that the boundary is still  $C^1$ , but it is not flat. Pick up any  $\hat{x} \in \partial\Omega$ . Recall from calculus that there exists a  $C^1$  mapping  $\phi$  with an inverse,  $\psi$  such that the mapping  $\phi$  straightens out the boundary near the point  $\hat{x}$ .

Let  $U$  denote the mapping of the ball. That is,  $U = \phi(B)$ . Write  $y = \phi(x)$  and  $x = \psi(y)$  so that  $\hat{u}(y) = u(\psi(y))$ . Then, construct a small ball  $\hat{B}$ , just like we did before. Finally, proceed in the same manner as we did before, when we assumed that the boundary was flat near  $\hat{x}$ .

Next, we claim that

$$\|\bar{u}\|_{W^{1,p}(B)} \leq C\|u\|_{W^{1,p}(B^+)}$$

where the constant  $C$  is independent of  $u$ . We verify this claim. Observe that

$$\begin{aligned} \int_B |D\bar{u}|^p dx &= \int_{B^+} |D\bar{u}|^p dx + \int_{B^-} |D\bar{u}|^p dx \\ &= \int_{B^+} |Du|^p dx + \int_{B^-} |D\bar{u}|^p dx. \end{aligned}$$

For  $x \in B^-$ , we can use the definition to explicitly calculate  $D\bar{u}$ . It is given by

$$|D\bar{u}|^p \leq \left(\frac{5}{3}\right)^p |Du(x_1, \dots, -x_n)|^p + \left(\frac{8}{3}\right)^p |Du(x_1, \dots, -x_n/4)|^p.$$

The inequality is due to the scaling factor in front of  $-x_n$ . Finally, making a change of variables and multiplying by the Jacobian, we see that

$$\int_{B^-} |D\bar{u}|^p dx \leq \int_{B^+} \left(\frac{5}{3}\right)^p |Du(x_1, \dots, x_n)|^p dx + \int_{B^+} 4 \left(\frac{8}{3}\right)^p |Du(x_1, \dots, x_n)|^p dx$$

so that

$$\int_B |D\bar{u}|^p dx \leq \left(1 + \left(\frac{5}{3}\right)^p + 4 \left(\frac{8}{3}\right)^p\right) \int_{B^+} |Du|^p dx.$$

It then follows that

$$\|\bar{u}\|_{W^{1,p}(B)} \leq C\|u\|_{W^{1,p}(B^+)}$$

which asserts the validity of the claim.

Without loss of generality, we will assume that the boundary is not straight. Using the notation and language in the previous paragraphs, we let  $W := \psi(B)$ . Furthermore, we observe that

$$\|u\|_{W^{1,p}(B^+)} \leq \|u\|_{W^{1,p}(\Omega)}.$$

To verify this, consider

$$\int_{\Omega} |Du|^p dx = \int_{B^+} |Du|^p dx + \int_{\Omega \setminus B^+} |Du|^p dx.$$

For the second integral, note that  $\Omega \setminus B^+ \subset \Omega$  and is bounded. Call this new set  $U$ . Then,

$$\int_U |Du|^p dx = \|u\|_{W^{1,p}(U)}^p$$

which by our definition of the norm, this is strictly positive and bounded. Thus,

$$\int_{B^+} |Du|^p dx \leq \int_{\Omega} |Du|^p dx$$

from which our claim follows. From these above calculations, and with the ultimate goal in mind, we have the inequality

$$\|\bar{u}\|_{W^{1,p}(W)} \leq C\|u\|_{W^{1,p}(\Omega)}$$

where the constant  $C$  does not depend on the function  $u$ .

Up until this point, we outlined a procedure for extending one point on the boundary. But we assumed that the boundary was compact, and therefore there are a finite number of points  $\hat{x}_i$ , open sets  $W_i$  and extensions  $\bar{u}_i$  of  $u$  to  $W_i$  as done in the procedure above. Assuming that there are  $N$  points for which this can be done, we require that the boundary can be reconstructed by taking the union of open sets. Next, let us choose some  $W_0$  so that

$$\Omega \subset \bigcup_{i=0}^N W_i$$

and let  $\{\xi_i\}_{i=0}^N$  be the associated partition of unity. Define the function

$$\bar{u} = \sum_{i=0}^{\infty} \xi_i \bar{u}_i.$$

Then, we observe that

$$\begin{aligned}
\|\bar{u}\|_{W^{1,p}(\mathbb{R}^n)} &= \|\bar{u}_0\xi_0 + \bar{u}_1\xi_1 + \cdots + \bar{u}_N\xi_N\|_{W^{1,p}(\mathbb{R}^n)} \\
&\leq \|\bar{u}_0\xi_0\|_{W^{1,p}(\mathbb{R}^n)} + \|\bar{u}_1\xi_1\|_{W^{1,p}(\mathbb{R}^n)} + \cdots + \|\bar{u}_N\xi_N\|_{W^{1,p}(\mathbb{R}^n)} \\
&= \|\bar{u}_0\xi_0\|_{W^{1,p}(W_0)} + \|\bar{u}_1\xi_1\|_{W^{1,p}(W_1)} + \cdots + \|\bar{u}_N\xi_N\|_{W^{1,p}(W_N)} \\
&= C_0(\xi_0)\|u\|_{W^{1,p}(\Omega)} + C_1(\xi_1)\|u\|_{W^{1,p}(\Omega)} + \cdots + C_N(\xi_N)\|u\|_{W^{1,p}(\Omega)} \\
&= C(\xi_0, \xi_1, \dots, \xi_N)\|u\|_{W^{1,p}(\Omega)}.
\end{aligned}$$

In the third step, we have used that the  $\xi_i$  have support in  $W_i$ . In the fourth step, we have used the definition of  $\bar{u}$  and pulled out the  $\xi_i$  and absorbed it into a constant. In the final step, we have combined all the constants into one large constant that is independent of  $u$ . This confirms the inequality.

Finally, we define  $Eu := \bar{u}$ . Note that in the beginning of the proof, we assumed that  $u \in C^1(\bar{\Omega})$ . This is not necessary. Choose some  $u_m \in C^\infty(\Omega)$  that converges to  $u$  in  $W^{1,p}(\Omega)$ . This is made possible by the approximation theorems that were developed in the section. Then we note that the linearity of  $E$  and the estimate gives

$$\begin{aligned}
\|Eu_m - Eu_l\|_{W^{1,p}(\mathbb{R}^n)} &= \|E(u_m - u_l)\|_{W^{1,p}(\mathbb{R}^n)} \\
&\leq C\|u_m - u_l\|_{W^{1,p}(\Omega)}
\end{aligned}$$

which shows that  $\{Eu_m\}_{m=1}^\infty$  is a Cauchy sequence. This Cauchy sequence converges to  $\bar{u}$ , which we have defined as  $Eu$ . It should be noted that this estimate was independent of the sequence  $\{u_m\}_{m=1}^\infty$ .

The conclusions of the theorem follow directly from the above work.  $\square$

In the above theorem, we assumed that the boundary was  $C^1$ . If the boundary were, or rather we required that it to be,  $C^2$ , then the proof would follow in the same manner but the choice of the reflection would be different. Regardless, without proof, the following bound can be derived.

$$\|Eu\|_{W^{2,p}(\mathbb{R}^n)} \leq \|u\|_{W^{2,p}(\Omega)}.$$

For the case where  $k > 2$ , the above proof serves as an outline, though the reflection will have to be carefully constructed. I encourage the reader, if they understand how the coefficients in the higher-order reflection were chosen, to attempt the proof.

### 3.4 Traces.

We move to discuss the possibility of assigning values to  $u \in W^{1,p}(\Omega)$  for points on  $\partial\Omega$ . The reason we want to assign values to the boundary is because later on we will study boundary-value problems, which constitute the majority of real-world problems. The issue is that the function  $u$  can be discontinuous at the boundary, since we only assumed that  $u \in W^{1,p}(\Omega)$  and not necessarily on the closure of  $\Omega$ . The notion of the *trace operator* allows us to assign values on the boundary, without introducing discontinuities.

**Theorem 3.6.** *Assume that  $\Omega$  is bounded and  $\partial\Omega$  is  $C^1$ . Then there exists a bounded linear operator*

$$T : W^{1,p}(\Omega) \rightarrow L^p(\partial\Omega)$$

*such that the following are satisfied:*

1.  $Tu = u|_{\partial\Omega}$  if  $u \in W^{1,p}(\Omega) \cap C(\bar{\Omega})$
2.  $\|Tu\|_{L^p(\partial\Omega)} \leq C\|u\|_{W^{1,p}(\Omega)}$  where  $C = C(p, \Omega)$  is independent of  $u$ .

*If such a linear operator exists, then we call  $Tu$  the trace of  $u$  on  $\partial\Omega$ .*

*Proof.* The first property will be proved in proving the second property. From elementary functional analysis, we know that if  $T$  is a bounded linear operator, then  $T$  must be a continuous mapping. Thus, we will focus on showing that  $T$  is a continuous mapping, and derive the desired result in the process.

Borrowing upon the ideas from the previous proof, we suppose that if  $\partial\Omega$  is not flat on  $\{x_n = 0\}$ , then we flatten it out using an injective mapping  $\phi$ , with inverse  $\psi$ . If  $\partial\Omega$  is already flat on  $\{x_n = 0\}$ , then we do not have to perform the flattening procedure.

Assume that  $u \in C^1(\bar{\Omega})$ . Pick up some  $\hat{x} \in \partial\Omega$ . Let  $B = B(\hat{x}, r)$  and  $\hat{B} = B(\hat{x}, r/2)$  be the ball of radius  $r$  and  $r/2$ , respectively, centered around  $\hat{x}$ . As with the previous proof, let us define

$$\begin{cases} B^+ := B \cap \{x_n \geq 0\} \subset \Omega \\ B^- := B \cap \{x_n < 0\} \subset \mathbb{R}^n \setminus \Omega \end{cases}$$

and

$$\begin{cases} \widehat{B}^+ := \widehat{B} \cap \{x_n \geq 0\} \subset B^+ \subset \Omega \\ \widehat{B}^+ := \widehat{B} \cap \{x_n < 0\} \subset \mathbb{R}^n \setminus \Omega. \end{cases}$$

Let us denote, by  $\Gamma$ , the set

$$\Gamma := \left\{ x \mid x \in \widehat{B}^+ \cap \{x_n = 0\} \right\}.$$

Let us pick up some  $\xi \in C_c^\infty(B)$  where  $\xi \geq 0$  and  $\xi \equiv 1$  on  $\widehat{B}^+$ . Let us denote  $x' = (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$ . Then, we observe that

$$\begin{aligned} \int_{\Gamma} |u|^p \, dx' &\leq \int_{\{x_n=0\}} \xi |u|^p \, dx' \\ &= - \int_{B^+} (\xi |u|^p)_{x_n} \, dx \\ &= - \int_{B^+} \xi_{x_n} |u| + p |u|^{p-1} \operatorname{sgn}(u) u_{x_n} \xi \, dx \\ &= I + J \end{aligned}$$

where  $\operatorname{sgn}(u)$  is the sign of  $u$ . The first step comes from using the positivity of  $\xi$  in  $B^+$ . The second step comes from the fundamental theorem of calculus, and noting that the function  $\xi$  has compact support in  $B$ . The third step is attained by an explicit expansion of the previous step.

Since  $\xi_{x_n}$  is a compactly-supported and smooth function, it attains a maximum inside  $B^+$ . Thus,

$$- \int_{B^+} \xi_{x_n} |u|^p \, dx \leq - \max_x (\xi_{x_n}) \int_{B^+} |u|^p \, dx = C_0 \int_{B^+} |u|^p \, dx.$$

For the second integral, we observe that

$$\begin{aligned} - \int_{B^+} p |u|^{p-1} \operatorname{sgn}(u) u_{x_n} \xi \, dx &\leq - \max_x (\xi) \operatorname{sgn}(u) \int_{B^+} p |u|^{p-1} u_{x_n} \, dx \\ &\leq C_1 \int_{B^+} p |u|^{p-1} |u_{x_n}| \, dx \\ &\leq C_1 \int_{B^+} p |u|^{p-1} |Du| \, dx \end{aligned}$$

where  $Du$  denotes the full gradient. The first step comes from noting that  $\xi$  is a smooth, continuous function. The second step comes from the fact that  $u_{x_n} \leq |u_{x_n}|$ . The third step comes from noting that  $|u_{x_n}| \leq |Du|$ . Next, let  $q = \frac{p}{p-1}$  so that  $\frac{1}{q} + \frac{1}{p} = 1$ . Then, employing Young's inequality, we observe that

$$p |u|^{p-1} |Du| \leq \frac{p |u|^{q(p-1)}}{q} + \frac{p |Du|^p}{p} = (p-1) |u|^p + |Du|^p.$$

Thus, we observe that

$$- \int_{B^+} p |u|^{p-1} \operatorname{sgn}(u) u_{x_n} \xi \, dx \leq C_1 \int_{B^+} (p-1) |u|^p + |Du|^p \, dx.$$

Combining the above calculations, we can see that

$$\begin{aligned} \int_{\Gamma} |u|^p \, dx' &\leq C_0 \int_{B^+} |u|^p \, dx + C_1 \int_{B^+} (p-1) |u|^p + |Du|^p \, dx \\ &= (C_0 + C_1(p-1)) \int_{B^+} |u|^p \, dx + C_1 \int_{B^+} |Du|^p \, dx. \end{aligned}$$

Borrowing again from ideas in the previous proof, we note that since  $\partial\Omega$  is a compact set, there exists finitely many points  $\hat{x}_i \in \partial\Omega$  and open subsets  $\Gamma_i$ , for  $i = 1, \dots, N$ , such that the boundary can be recovered from taking the union of these subsets. For each  $i$ , using the above result, we conclude that

$$\|u\|_{L^p(\Gamma_i)} \leq C \|u\|_{W^{1,p}(\Omega)}.$$

We can then define  $Tu := u|_{\partial\Omega}$  so that in the above bound, we have

$$\|Tu\|_{L^p(\Gamma_i)} \leq C \|u\|_{W^{1,p}(\Omega)}$$

where  $C$  is a constant that does not depend on  $u$ . Note that the mapping  $T$  is a linear mapping.

The inequality was derived, assuming that  $u \in C^1(\bar{\Omega})$ . In fact, this requirement might be too stringent, so we loosen it using the approximation theorems developed in the previous sections. Assume now that  $u \in W^{1,p}(\Omega)$ . Then pick up some  $u_m \in C^\infty(\bar{\Omega})$  such that  $u_m \rightarrow u$  in  $W^{k,p}(\Omega)$ . According to our results, we have

$$\begin{aligned} \|Tu_m - Tu_l\|_{L^p(\partial\Omega)} &= \|T(u_m - u_l)\|_{L^p(\partial\Omega)} \\ &\leq C\|u_m - u_l\|_{W^{1,p}(\Omega)}. \end{aligned}$$

This shows that  $\{Tu_m\}_{m=1}^\infty$  is a Cauchy sequence in  $L^p(\partial\Omega)$ . Thus, we define

$$Tu := \lim_{m \rightarrow \infty} Tu_m.$$

This is true, regardless of our choice of the  $u_m$  that we have chosen to approximate the function  $u$ . Hence, the mapping  $T$  is bounded, linear, and continuous.

To conclude this proof, we notice that if the function  $u$  is in the Sobolev space and it is also smooth on the boundary, then the  $u_m$  converge to  $u$  uniformly on  $\bar{\Omega}$ . Hence, we can say that  $Tu := u|_{\partial\Omega}$  if  $u \in W^{1,p}(\Omega) \cap C(\bar{\Omega})$ .  $\square$

Later on, we will study elliptic and parabolic boundary-value problems, with the homogeneous boundary condition. In order to understand the boundary condition, we need to first understand what it means for a function to have a zero trace. The following theorem sums this up.

**Theorem 3.7.** *Assume that  $\Omega$  is bounded and  $\partial\Omega$  is  $C^1$ . Furthermore, suppose that  $u \in W^{1,p}(\Omega)$ . Then*

$$Tu = 0 \iff u \in W_0^{1,p}(\Omega).$$

*That is to say,  $u$  must be in the closure of  $C_c^\infty(\Omega)$  in  $W^{1,p}(\Omega)$ , for the trace of  $u$  to be identically zero.*

*Proof.* Let us prove the forward statement. Suppose that  $u \in W_0^{1,p}(\Omega)$ . Recall that the space  $W_0^{1,p}(\Omega)$  is the closure of  $C_c^\infty(\Omega)$  in  $W^{1,p}(\Omega)$ . Thus, by definition, there exist functions  $u_m \in C_c^\infty(\Omega)$  such that  $u_m \rightarrow u$  in  $W^{1,p}(\Omega)$ . If  $Tu_m = 0$ , then by the continuity, linearity, and boundedness of  $T$ , we deduce that  $Tu = 0$  in  $\partial\Omega$ .

The other statement is a little harder to prove. Let us suppose that  $Tu = 0$  on  $\partial\Omega$ .  $\square$

### 3.5 Sobolev Embeddings.

We move to discuss embeddings of Sobolev spaces into other function spaces. These results will give rise to so-called Sobolev-type inequalities. These inequalities will establish estimates for arbitrary functions in the Sobolev space. These results will be utilized later, when we study second-order partial differential equations.

The motivation for finding embeddings of Sobolev spaces into other function spaces is so that if we have a function  $u \in W^{1,p}(\Omega)$ , then is there a way to automatically say that  $u$  belongs to another function space as well? The answer is yes for certain choices of  $p$ . One particular embedding we will consider is the embedding of  $W^{1,p}(\Omega)$  in  $L^q(\Omega)$ , for particular values of  $q$ . We desire to establish embeddings of Sobolev spaces into  $L^p$  spaces because there is a lot of theory surrounding  $L^p$  spaces. As such, if we can establish that a function  $u$  is in a Sobolev space, then it automatically belongs to an  $L^p$  space, hence can apply results from  $L^p$  theory.

Let  $1 \leq p < n$ . We can develop an estimate that bounds the norm of  $u$  in terms of  $Du$ , each with respect to different norms. In particular, we have that

$$\|u\|_{L^q(\mathbb{R}^n)} \leq C\|Du\|_{L^p(\mathbb{R}^n)} \tag{3.1}$$

for some constant  $C > 0$  that is independent of  $u \in C_c^\infty(\mathbb{R}^n)$ , and  $1 \leq q < \infty$ . The above estimate does not hold for arbitrary choices of  $q$ . In fact, the estimate only holds for particular values of  $q$ . Consider the following argument. Let us pick some function  $u \in C_c^\infty(\mathbb{R})$ , such that  $u \not\equiv 0$ , and define for  $\delta > 0$ , the scaled function

$$u_\delta(x) := u(\delta x).$$

Next, observe that

$$\begin{aligned} \int_{\mathbb{R}^n} |u_\delta|^q dx &= \int_{\mathbb{R}^n} |u(\delta x)|^q dx \\ &= \frac{1}{\delta^n} \int_{\mathbb{R}^n} |u(y)|^q dy \end{aligned}$$

where we have made a change of variables in the last step. Similarly, observe that

$$\begin{aligned} \int_{\mathbb{R}^n} |Du_\delta|^p dx &= \int_{\mathbb{R}^n} |Du(\delta x)|^p dx \\ &= \frac{\delta^p}{\delta^n} \int_{\mathbb{R}^n} |Du(y)|^p dy. \end{aligned}$$

Plugging these into (3.1), we find that

$$\frac{1}{\delta^{n/q}} \|u\|_{L^q(\mathbb{R}^n)} \leq C \frac{\delta}{\delta^{n/p}} \|Du\|_{L^p(\mathbb{R}^n)}$$

and thus

$$\|u\|_{L^q(\mathbb{R}^n)} \leq C \delta^{1 - \frac{n}{p} + \frac{n}{q}} \|Du\|_{L^p(\mathbb{R}^n)}. \quad (3.2)$$

If  $1 - \frac{n}{p} + \frac{n}{q} \neq 0$ , then we can obtain a contradiction by sending  $\delta$  either to 0 or  $\infty$  in (3.2). Thus, for (3.1) to hold, we require that

$$1 - \frac{n}{p} + \frac{n}{q} = 0$$

so that, with a little rearranging, we obtain the requirement that

$$q = \frac{np}{n-p}.$$

In the theory of Sobolev spaces, this is called the *Sobolev conjugate*. Thus, if an estimate like (3.1) holds, then it only holds for that specific choice of  $p$ . We now move to show the validity of this estimate, which is summarized in the following theorem.

**Theorem 3.8.** (*Gagliardo-Nirenberg-Sobolev Inequality*). *Let  $1 \leq p \leq n$ . Then there exists a constant  $C = C(p, n)$  such that*

$$\|u\|_{L^{p^*}(\mathbb{R}^n)} \leq C \|Du\|_{L^p(\mathbb{R}^n)}$$

for all  $u \in C_c^1(\mathbb{R}^n)$ , where  $p^*$  is the Sobolev conjugate given by

$$p^* := \frac{np}{n-1}.$$

*Proof.* Suppose that  $p = 1$ . Then for each  $i = 1, \dots, n$  and  $x \in \mathbb{R}^n$ , we observe

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

which comes from the fundamental theorem of calculus. Notice that we have also used the fact that  $u$  has compact support, so that it vanishes near the boundary. If  $Du$  denotes the full derivative of  $u$ , then we can clearly see that

$$|u(x)| \leq \int_{-\infty}^{\infty} |Du(x_1, \dots, y_i, \dots, x_n)| dy_i$$

for each  $i$ . Thus if we apply the above inequality for each  $i$ , we find that

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

so that

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

For convenience, I will use the shorthand notation  $Du = Du(x_1, \dots, y_i, \dots, x_n)$ . Similarly, we will define for convenience,

$$\alpha = \frac{n}{n-1}, \quad \beta = \frac{1}{n-1}$$

Our goal is to turn the left hand side to look like the  $L^q(\mathbb{R}^n)$  norm. With this in mind, we integrate both sides with respect to  $x_1$ , to obtain

$$\int_{-\infty}^{\infty} |u(x)|^\alpha dx_1 \leq \int_{-\infty}^{\infty} \prod_{i=1}^n \left( \int_{-\infty}^{\infty} |Du(x_1, \dots, y_i, \dots, x_n)| dy_i \right)^\beta dx_1$$



$$\begin{aligned}
&= \left( \int_{-\infty}^{\infty} |Du| \, dy_1 \right)^{\beta} \left( \int_{-\infty}^{\infty} \prod_{i=2}^n \left[ \int_{-\infty}^{\infty} |Du| \, dy_i \right]^{\beta} dx_1 \right) \\
&\leq \left( \int_{-\infty}^{\infty} |Du| \, dy_1 \right)^{\beta} \left( \prod_{i=2}^n \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du| \, dx_1 \, dy_i \right]^{\beta} \right) \\
&= I_1 \left( \prod_{i=2}^n \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du| \, dx_1 \, dy_i \right]^{\beta} \right)
\end{aligned}$$

where the last inequality is obtained after applying the general Hölder's inequality. Continuing on, we integrate both sides with respect to  $x_2$ , to obtain

$$\begin{aligned}
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |u(x)|^{\alpha} \, dx_1 \, dx_2 &\leq \int_{-\infty}^{\infty} I_1 \prod_{i=2}^n \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du| \, dx_1 \, dy_i \right]^{\beta} dx_2 \\
&= \left( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du| \, dx_1 \, dy_2 \right)^{\beta} \int_{-\infty}^{\infty} I_1 \prod_{i=3}^n \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du| \, dx_1 \, dy_i \right]^{\beta} dx_2 \\
&\leq \left( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du| \, dx_1 \, dy_2 \right)^{\beta} \left( I_2 \prod_{i=3}^n \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du| \, dx_1 \, dy_i \, dx_2 \right]^{\beta} \right)
\end{aligned}$$

where  $I_2$  is given by

$$I_2 := \left( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |Du| \, dy_1 \, dx_2 \right)^{\beta}$$

The last step is again, due to an application of the generalized Hölder's inequality. Continuing in this fashion, we integrate with respect to  $x_3, \dots, x_n$ . Doing so, we observe that

$$\begin{aligned}
\int_{\mathbb{R}^n} |u(x)| \, dx &\leq \prod_{i=1}^n \left[ \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} |Du| \, dx_1 \dots dy_i \dots dx_n \right]^{\beta} \\
&= \left( \int_{\mathbb{R}^n} |Du| \, dx \right)^{\alpha}
\end{aligned}$$

which is the desired result for the case when  $p = 1$ . To prove the case where  $1 < p < n$ , pick some  $\gamma > 1$  and apply the same procedure as before, but this time to the function  $|u|^{\gamma}$ . We apply the estimate to the absolute value of  $u$  because in the end, we wish to keep the  $|Du|$  term. Utilizing the estimate for  $p = 1$ , we have

$$\begin{aligned}
\left( \int_{\mathbb{R}^n} |u|^{\gamma\alpha} \, dx \right)^{\frac{1}{\alpha}} &\leq \int_{\mathbb{R}^n} |D|u|^{\gamma}| \, dx \\
&= \gamma \int_{\mathbb{R}^n} |u|^{\gamma-1} |Du| \, dx \\
&\leq \gamma \left( \int_{\mathbb{R}^n} |u|^{(\gamma-1)\frac{p}{p-1}} \, dx \right)^{\frac{p-1}{p}} \left( \int_{\mathbb{R}^n} |Du|^p \, dx \right)^{\frac{1}{p}}
\end{aligned}$$

where the last inequality is due to an application of Hölder's inequality. The estimate can be further improved. The choice of  $\gamma$  is up to us, so let us pick  $\gamma$  such that

$$\gamma\alpha = (\gamma - 1)\frac{p}{p-1} \tag{3.3}$$

so that our choice of  $\gamma$  is

$$\gamma = \frac{p(n-1)}{n-p} > 1.$$

With this choice of  $\gamma$ , the estimate becomes

$$\left( \int_{\mathbb{R}^n} |u|^{\gamma\alpha} \, dx \right)^{\left[ \frac{1}{\alpha} - \frac{p-1}{p} \right]} \leq \gamma \left( \int_{\mathbb{R}^n} |Du|^p \, dx \right)^{\frac{1}{p}}.$$

Everything in the brackets can be reduced to

$$\frac{1}{\alpha} - \frac{p-1}{p} = \frac{np}{n-p} =: p^*.$$

Additionally, with this choice of  $\gamma$ , note that

$$\gamma\alpha = \left( \frac{p(n-1)}{n-p} \right) \left( \frac{n}{n-1} \right) = \frac{np}{n-p} =: p^*$$

Thus, the estimate becomes

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

which is indeed, the desired result.  $\square$

Using the GNS theorem, we will now show that the space  $W^{1,p}(\Omega) \subset L^{p^*}(\Omega)$ . That is, the space  $W^{1,p}(\Omega)$  is embedded in the space  $L^{p^*}(\Omega)$ .

**Theorem 3.9.** *Let  $\Omega$  be a bounded domain, the boundary be  $C^1$ ,  $1 \leq p < n$ , and  $u \in W^{1,p}(\Omega)$ . Then  $u \in L^{p^*}$  and*

$$\|u\|_{L^{p^*}(\Omega)} \leq C \|u\|_{W^{1,p}(\Omega)}.$$

*Proof.* Since the boundary is  $C^1$ , by the extension theorem discussed in section 3.3, there exists an extension  $Eu = \bar{u} \in W^{1,p}(\mathbb{R}^n)$  such that  $\bar{u} = u$  in  $\Omega$ ,  $\bar{u}$  has compact support, and we have the estimate

$$\|\bar{u}\|_{W^{1,p}(\mathbb{R}^n)} \leq C \|u\|_{W^{1,p}(\Omega)}.$$

Since  $\bar{u}$  has compact support, there exist functions  $u_m \in C_c^\infty(\mathbb{R}^n)$  ( $m = 1, \dots$ ) such that

$$u_m \rightarrow \bar{u} \text{ in } W^{1,p}(\mathbb{R}^n).$$

Now, by the GNS theorem, we have the estimate

$$\|u_m - u_l\|_{L^{p^*}(\mathbb{R}^n)} \leq C \|D(u_m - u_l)\|_{L^p(\mathbb{R}^n)} = C \|Du_m - Du_l\|_{L^p(\mathbb{R}^n)}$$

for each  $l, m \geq 1$ . The right hand-side tends to 0, and thus the above inequality implies that

$$u_m \rightarrow \bar{u} \text{ in } L^{p^*}(\mathbb{R}^n). \quad (3.4)$$

Additionally, we have

$$Du_m \rightarrow D\bar{u} \text{ in } L^p(\mathbb{R}^n). \quad (3.5)$$

To verify this claim, notice that by definition, if  $u_m \rightarrow \bar{u}$  in  $W^{1,p}(\Omega)$ , then

$$\|u_m - \bar{u}\|_{W^{1,p}(\mathbb{R}^n)} = \left( \|u_m - \bar{u}\|_{L^p(\mathbb{R}^n)} + \|Du_m - D\bar{u}\|_{L^p(\mathbb{R}^n)} \right)^{1/p}$$

tends to zero. Since

$$\|Du_m - D\bar{u}\|_{L^p(\mathbb{R}^n)} \leq \|u_m - \bar{u}\|_{W^{1,p}(\mathbb{R}^n)}$$

we have verified that  $Du_m \rightarrow D\bar{u}$  in  $L^p(\mathbb{R}^n)$ . Applying the GNS inequality once again, we find that

$$\|u_m\|_{L^{p^*}(\mathbb{R}^n)} \leq C \|Du_m\|_{L^p(\mathbb{R}^n)}.$$

Combining (3.4) and (3.5), we observe that

$$\|\bar{u}\|_{L^{p^*}(\mathbb{R}^n)} \leq C \|D\bar{u}\|_{L^p(\mathbb{R}^n)}.$$

By definition of  $\bar{u}$ , we arrive at the desired result. This completes the proof.  $\square$

With the above inequality, we have now shown the embedding of  $W^{1,p}$  into  $L^{p^*}$ , which we will denote  $W^{1,p} \hookrightarrow L^{p^*}$ . So if we have a function  $u \in W^{1,p}(\Omega)$ , then we automatically have that  $u \in L^{p^*}(\Omega)$ . One particular property of  $L^p$  spaces is that its dual space is easily characterized, while the dual of  $W^{1,p}(\Omega)$  is not. Thus, by proving this embedding exists, we can work with  $L^p$  spaces instead.

Before we move on and discuss another Sobolev-type inequality, the Poincaré's inequality, we introduce a lemma.

**Lemma 3.1.** ( *$L^p$  Inclusion Lemma*). Let  $\Omega$  be bounded,  $1 \leq q \leq p^*$ , and  $u \in L^{p^*}(\Omega)$ . Then

$$\|u\|_{L^q(\Omega)} \leq \|u\|_{L^{p^*}(\Omega)}.$$

*Proof.* Observe that

$$\begin{aligned} \|u\|_{L^{p^*}(\Omega)} &\leq \|u\|_{L^{p^*}(\Omega)}^{p^*} \\ &= \int_{\Omega} |u|^{p^*} dx \\ &= \int_{\Omega} |u|^{p^*-1} |u| dx \\ &= \left( \int_{\Omega} |u|^{(p^*-1)\alpha} dx \right)^{1/\alpha} \left( \int_{\Omega} |u|^q dx \right)^{1/q} \end{aligned}$$

where  $\alpha$  is chosen so that  $\frac{1}{\alpha} + \frac{1}{q} = 1$ , so that we may apply Hölder's inequality. Thus, we conclude that

$$\|u\|_{L^q(\Omega)} \leq \|u\|_{L^{p^*}(\Omega)}$$

which is the desired result.  $\square$

We now move to discuss two Poincaré's inequalities, which will provide us with an estimate for  $u \in W_0^{1,p}(\Omega)$ .

**Theorem 3.10.** Let  $\Omega$  be a bounded domain, the boundary be  $C^1$ ,  $1 \leq p < n$ . Then for all  $1 \leq q \leq p^*$ , there exists a constant  $C$  independent of  $u \in W_0^{1,p}(\Omega)$  such that we have the estimate

$$\|u\|_{L^q(\Omega)} \leq C \|Du\|_{L^p(\Omega)}.$$

*Proof.* Since  $u \in W_0^{1,p}(\Omega)$ , by definition there exists functions  $u_m \in C_c^\infty(\Omega)$  ( $m = 1, \dots$ ) such that

$$u_m \rightarrow u \text{ in } W_0^{1,p}(\Omega).$$

For each of these  $u_m$ , let us define the extended function  $\bar{u}_m := Eu_m$  such that

$$\bar{u}_m := \begin{cases} u_m, & \text{if } x \in \Omega \\ 0, & \text{if } x \in \mathbb{R}^n \setminus \bar{\Omega}. \end{cases}$$

Notice that this extension is possible because  $u_m$  is a smooth and compactly supported function. As a result, this extension does not introduce any particular discontinuities. Applying the GNS theorem, we can see that

$$\|u_m - u_l\|_{L^{p^*}(\mathbb{R}^n)} \leq C \|Du_m - Du_l\|_{L^p(\mathbb{R}^n)}$$

so that, since  $u \in W_0^{1,p}(\Omega)$ ,

$$u_m \rightarrow u \text{ in } L^{p^*}(\mathbb{R}^n). \quad (3.6)$$

By the same argument presented in a previous proof, we can see that

$$Du_m \rightarrow Du \text{ in } L^p(\mathbb{R}^n). \quad (3.7)$$

By the GNS inequality, we have that

$$\|\bar{u}_m\|_{L^{p^*}(\mathbb{R}^n)} \leq C \|D\bar{u}_m\|_{L^p(\mathbb{R}^n)}.$$

Using the definition of the extension of  $u_m$ , the estimate reduces to

$$\|u_m\|_{L^{p^*}(\Omega)} \leq C \|Du_m\|_{L^p(\Omega)}.$$

Combining (3.6) and (3.7), we find that

$$\|u\|_{L^{p^*}(\Omega)} \leq C \|Du\|_{L^p(\Omega)}. \quad (3.8)$$

Combining (3.8) with lemma 3.1, we obtain that

$$\|u\|_{L^q(\Omega)} \leq C \|Du\|_{L^p(\Omega)}$$

for  $q \in [1, p^*]$ , which is the desired result.  $\square$

In view of the above result, we can see that the norm  $\|Du\|_{L^p(\Omega)}$  is equivalent to  $\|u\|_{W^{1,p}(\Omega)}$ .

**Lemma 3.2.** *If  $u \in W_0^{1,p}(\Omega)$  and  $\Omega$  is bounded, then the norm  $\|Du\|_{L^p(\Omega)}$  is equivalent to  $\|u\|_{W^{1,p}(\Omega)}$ .*

*Proof.* From theorem 3.9 and choosing  $q = p^*$ , we have

$$\|u\|_{L^{p^*}(\Omega)} \leq C_0 \|u\|_{W^{1,p}(\Omega)}.$$

From theorem 3.10, we have

$$\|u\|_{L^{p^*}(\Omega)} \leq C_1 \|Du\|_{L^p(\Omega)}.$$

In both the above, the constants  $C_0, C_1$  are arbitrary, positive, constants. Combining these two results, we can clearly see that the two norms are equivalent, given the assumptions.  $\square$

### 3.6 Compactness.

## 4 Second-Order Elliptic Equations.

We move to discuss second-order elliptic PDEs subject to boundary conditions. Second-order elliptic PDEs arise in many physical applications in engineering, biology, and physics. For example, Laplace's equation appears in many applications in electrostatics and Poisson's equation arises in many applications in electrodynamics.

Though Laplace and Poisson's equations are classic examples of second-order elliptic PDEs, in this section, we will use the tools developed in our discussion on Sobolev spaces to study the more general problem

$$\begin{cases} Lu = f, & \text{in } \Omega \\ u = 0, & \text{on } \partial\Omega \end{cases} \quad (4.1)$$

where  $u = u(x), u : \bar{\Omega} \rightarrow \mathbb{R}$  is the unknown function of interest. The function  $f : \Omega \rightarrow \mathbb{R}$  is given, and  $L$  is a second-order partial differential operator. The partial differential operator can take one of two forms:

$$Lu = - \sum_{i,j=1}^n \left( a^{ij}(x) u_{x_i} \right)_{x_j} + \sum_{i=1}^n b^i(x) u_{x_i} + c(x)u \quad (4.2)$$

or

$$Lu = - \sum_{i,j=1}^n a^{ij}(x) u_{x_i x_j} + \sum_{i=1}^n b^i(x) u_{x_i} + c(x)u \quad (4.3)$$

where the coefficient functions are known and  $a^{ij} = a^{ji}$ , which is called the symmetry condition.

Eq. (4.2) is known as the *divergence form* of the operator  $L$ . Eq. (4.3) is known as the *non-divergence form* of the operator  $L$ . Both forms have their advantages and disadvantages. The divergence form of  $L$  is useful for modern energy-based methods and the non-divergence form is useful for maximum principle-based methods, for discussing properties of solutions to PDEs.

**Definition 4.1.** We say that the partial differential operator  $L$  is uniformly elliptic if there exists  $\theta > 0, \theta \in \mathbb{R}$  such that

$$\sum_{i,j=1}^n a^{ij}(x) \xi_i \xi_j \geq \theta |\xi|^2$$

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

For each  $x \in \Omega$ , let  $A(x)$  be the matrix where the  $ij$ -th element is  $a^{ij}(x)$ . Then the above condition for  $L$  to be a uniformly elliptic partial differential operator is equivalent to requiring that for each  $x \in \Omega$ , the matrix  $A$  is symmetric positive definite, and for any  $x \in \Omega$ , the smallest eigenvalue is greater than or equal to  $\theta$ .

For example, if we take  $a^{ij} = \delta_{ij}$ , where  $\delta_{ij}$  is the Kronecker delta, and  $b^i = 0$  and  $c \equiv 0$  for all  $x \in \Omega$ , then the partial differential operator  $L$  is the standard Laplacian,  $-\Delta$ . The problem then reduces to finding  $u$  such that

$$\begin{cases} -\Delta u = f, & \text{in } \Omega \\ u = 0, & \text{on } \partial\Omega \end{cases}$$

which is the Poisson equation subject to homogeneous Dirichlet boundary conditions. Though we have established representation formulas for this problem on specific domains, we must be careful to assume that these representation formulas hold even if our domain  $\Omega$  is that specific domain. We must instead work directly with the PDE and unveil its properties.

## 4.1 Motivation for Weak Solutions.

Our goal is now to define and construct weak solutions  $u$  to the boundary value PDE problem. Later, we will examine the smoothness, amongst other topological properties, of  $u$ .

Let us assume that the coefficient functions  $a^{ij}, b^i, c \in L^\infty(\Omega)$  for  $i, j = 1, \dots, n$ . Furthermore, let us assume that  $f \in L^2(\Omega)$ . To define the weak solution, let us assume that  $L$  has the divergence form. Let  $\phi \in C_c^\infty(\Omega)$  be a test function. Integrating the PDE  $Lu = f$  over  $\Omega$ , we obtain

$$-\int_{\Omega} \sum_{i,j=1}^n \left( a^{ij} u_{x_i} \right)_{x_j} \phi \, dx + \int_{\Omega} \sum_i b^i u_{x_i} \phi \, dx + \int_{\Omega} cu \phi \, dx = \int_{\Omega} f \phi \, dx.$$

The first term can be reduced using integration by parts, to obtain

$$\int_{\Omega} \sum_{i,j=1}^n a^{ij} u_{x_i} \phi_{x_j} \, dx + \int_{\Omega} \sum_i b^i u_{x_i} \phi \, dx + \int_{\Omega} cu \phi \, dx = \int_{\Omega} f \phi \, dx.$$

Note that since  $\phi$  is a test function, it vanishes on the boundary. Using the approximation tools discussed in section 3.1, we find that the above holds for functions  $\phi \in H_0^1(\Omega)$ . In particular, the only assumption we need to place on  $\phi$  was that it has a 1<sup>st</sup>-weak derivative. If we replace the very smooth function  $\phi$  with a  $\phi \in H_0^1(\Omega)$ , then we also require that  $u \in H_0^1(\Omega)$ . The space  $H_0^1(\Omega)$ , which is the closure of  $C_c^\infty(\Omega)$  in  $H^1(\Omega)$ , was chosen because it incorporates the required boundary condition,  $\phi = 0$  on  $\partial\Omega$ . Without this boundary condition, the step where we used integration by parts would not be valid. So we find that we never required  $\phi$  to be a very smooth function.

As the left-hand side of the previous inequality is used quite often, it is given a special name.

**Definition 4.2.** Let  $u, v \in H_0^1(\Omega)$ . The bilinear form  $B[\cdot, \cdot]$  associated with the elliptic operator  $L$  is defined as

$$B[u, \phi] = \int_{\Omega} \sum_{i,j=1}^n a^{ij} u_{x_i} \phi_{x_j} + \sum_i b^i u_{x_i} \phi + cu \phi \, dx.$$

Using this notation, we can define the weak solution to the boundary value problem.

**Definition 4.3.** We say that  $u \in H_0^1(\Omega)$  is a weak solution of the boundary value problem (4.1) if

$$B[u, \phi] = (f, \phi) = \int_{\Omega} f \phi \, dx$$

for all  $\phi \in H_0^1(\Omega)$ , where  $(\cdot, \cdot)$  denotes the inner product in  $L^2(\Omega)$ . Using the language of the calculus of variations, this is a variational formulation of the boundary-value problem.

More generally, if we consider the boundary-value problem

$$\begin{cases} Lu = f^0 - \sum_{i=1}^n f_{x_i}^i, & \text{in } \Omega \\ u = 0, & \text{on } \partial\Omega. \end{cases}$$

Hereinafter, we will only consider the boundary value problem where the boundary condition is the homogeneous Dirichlet boundary condition. It is often the case, especially in physical applications, that they boundary condition is non-homogeneous. Without loss of generality, we may assume that the boundary conditions is identically equal to 0. Consider the following argument. Let us consider the problem

$$\begin{cases} Lu = f, & \text{in } \Omega \\ u = g, & \text{on } \partial\Omega \end{cases} \quad (4.4)$$

where  $\partial\Omega$  is  $C^1$  and  $u \in H^1(\Omega)$  is a weak solution to the PDE. Using the language developed in our discussion of traces, the above requires that  $u = g$  on  $\partial\Omega$  in the trace sense. If this is true, then it must also be true that  $g$  is the trace of some function  $w \in H^1(\Omega)$ . From this, we deduce that  $\bar{u} = u - w$  is an element of  $H_0^1(\Omega)$  since  $u = 0$  in the trace sense. Therefore, if we solve the boundary-value problem

$$\begin{cases} L\bar{u} = \bar{f}, & \text{in } \Omega \\ \bar{u} = 0, & \text{on } \partial\Omega \end{cases}$$

where  $\bar{f} = f - Lw \in H^{-1}(\Omega)$ , we can recover the solution to (4.4) very easily.

## 4.2 Existence of Weak Solutions.

Motivated by our search for weak solutions, we step back and ask ourselves: what conditions, if any, must be satisfied for a weak solution to exist? To this end, we introduce a fundamental result from functional analysis that will be useful for studying the existence of weak solutions.

*Remark 4.1.* For the remainder of this section, let  $H$  be a Hilbert space endowed with the norm  $\|\cdot\|$  and inner product  $(\cdot, \cdot)$ .

Before we discuss the Lax-Milgram theorem, which will be useful for showing that the weak solution exists, we first make some remarks about the operator  $\mathcal{L}$ . Let us consider

$$B[u, v] = \int_{\Omega} \sum_{i,j=1}^n a_{ij} u_{x_i} v_{x_j} + \sum_{i=1}^n b_i u_{x_i} v + c u v \, dx \quad (4.5)$$

which is exactly the non-divergence form for the second-order elliptic operator. Then for  $u, v \in H_0^1(\Omega)$ , we observe that

$$|B[u, v]| \leq \int_{\Omega} \sum_{i,j=1}^n |a_{ij}| |u_{x_i}| |v_{x_j}| + \sum_{i=1}^n |b_i| |u_{x_i}| |v| + c u v \, dx$$

where we have used the triangle and Cauchy-Schwarz inequalities. Furthermore, we observe that

**Theorem 4.1.** *Let  $B : H \times H \rightarrow \mathbb{R}$  be a bilinear mapping, such that for positive constants  $\alpha, \beta$ , we have*

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

*for  $u, v \in H$  and*

$$\beta \|u\| \leq B[u, v]$$

*for  $u \in H$ . If  $f : H \rightarrow \mathbb{R}$  is a bounded linear functional on  $H$ , then there exists a unique element  $u \in H$  such that*

$$B[u, v] = \langle f, v \rangle$$

*for all  $v \in H$ .*

*Proof.* From the assumption, we know that the mapping  $x \mapsto B[u, v]$  is a bounded linear mapping in  $H$ . Thus, by the Riesz representation theorem, there exists a unique  $w \in H$  such that

$$B[u, v] = (w, v).$$

Whence the above holds we will write  $Au = w$ , for  $A : H \rightarrow H$ . Then the above becomes

$$B[u, v] = (Au, v). \quad (4.6)$$

We propose that the operator  $A$  is a bounded linear operator. To see this, we first note that for  $\lambda_1, \lambda_2 \in \mathbb{R}, u_1, u_2, v \in H$  we have that

$$\begin{aligned} (A(\lambda_1 u_1 + \lambda_2 u_2), v) &= B[\lambda_1 u_1 + \lambda_2 u_2, v] \\ &= \lambda_1 B[u_1, v] + \lambda_2 B[u_2, v] \\ &= \lambda_1 (Au_1, v) + \lambda_2 (Au_2, v) \\ &= (\lambda_1 Au_1 + \lambda_2 Au_2, v) \end{aligned}$$

thus concluding that the operator  $A$  is indeed linear. The last step comes from the linearity, in the first argument, of the inner product. Additionally, we note that

$$\|Au\|^2 = (Au, Au) = B[u, Au] \leq \alpha \|u\| \|Au\|$$

and thus we conclude that

$$\|Au\| \leq \alpha \|u\|$$

thus concluding that the operator  $A$  is indeed bounded. These two arguments confirm that  $A$  is a bounded linear operator.

Our next goal is to show that  $A$  is a one-to-one function and that the range of  $A$  is closed. Notice that

$$\beta \|u\|^2 \leq B[u, u] = (Au, u) \leq \|Au\| \|u\|$$

by the Cauchy-Schwarz lemma, and thus  $\beta \|u\| \leq \|Au\| \leq \alpha \|u\|$ . This confirms that the operator  $A$  is one-to-one and that the range of  $A$  is closed.  $\square$

## 5 Second-Order Parabolic Equations.

In this section, we discuss a class of PDEs that involve time as a variable. These PDEs are often referred to as *evolution equations*, because the solutions evolve over time. Our goal, as an analyst, is to study the behavior of the solution over time. In particular, we will study second-order parabolic PDEs subject to appropriately defined initial and boundary conditions. Much of the development of the theory surrounding these PDEs will mimic our treatment of second-order elliptic equations. In particular, we study the existence and uniqueness of weak solutions as well as other interesting properties.

For this section, let  $U$  to be open, bounded subsets of  $\mathbb{R}^n$  and as before, set  $U_T = U \times (0, T]$  for some fixed  $T > 0$ . We will study the initial/boundary value problem

$$\begin{cases} u_t + Lu = f, & \text{in } U_T \\ u = 0, & \text{on } \partial U \times [0, T] \\ u = g, & \text{on } U \times \{t = 0\}, \end{cases} \quad (5.1)$$

where  $f : U_T \rightarrow \mathbb{R}$  and  $g : U \rightarrow \mathbb{R}$  are given, and  $u = u(x, t)$ ,  $u : U_T \rightarrow \mathbb{R}$  is the unknown function of interest. The operator  $L$  denotes, for each time  $t$ , the partial differential operator which has either the divergence form

$$Lu = - \sum_{i,j=1}^n (a^{ij}(x, t) u_{x_i})_{x_j} + \sum_{i=1}^n b^i(x, t) u_{x_i} + c(x, t) u \quad (5.2)$$

or the non-divergence form

$$Lu = - \sum_{i,j=1}^n a^{ij}(x, t) u_{x_i x_j} + \sum_{i=1}^n b^i(x, t) u_{x_i} + c(x, t) u \quad (5.3)$$

for given coefficients  $a^{ij}, b^i, c$  for each  $i, j = 1, \dots, n$ . As with second-order elliptic equations, both forms have their advantages and disadvantages.

As with second-order elliptic equations, we will consider the boundary value problem where the boundary problem is the homogeneous Dirichlet boundary condition. If the boundary condition is not homogeneous, we can force it to be homogeneous using the same methods used in our discussion of second-order elliptic equations.

**Definition 5.1.** We say that the partial differential operator  $\frac{\partial}{\partial t} + L$  is uniformly parabolic if there exists a constant  $\theta > 0$  such that

$$\sum_{i,j=1}^n a^{ij}(x, t) \xi_i \xi_j \geq \theta |\xi|^2$$

for all  $(x, t) \in U_T$  and  $\xi \in \mathbb{R}^n$ . Note that for each fixed time  $t$ , the operator  $L$  is uniformly parabolic in the spatial variable  $x$ .

As discussed before, let  $x \in \Omega, 0 \leq t \leq T$  be some fixed time, and  $A(x)$  be the matrix where the  $ij$ -th element is  $a^{ij}(x, t)$ . Then the above condition for  $L$  to be uniformly parabolic in  $x$  is the equivalent to requiring that for each  $x \in \Omega$  the matrix  $A$  is symmetric positive definite, and for any  $x \in \Omega$  the smallest eigenvalue is greater than or equal to  $\theta$ .

If we take  $a^{ij} = \delta_{ij}, b^i = c = 0$  and any arbitrary function  $f$ , then the operator  $L$  is the standard Laplacian  $-\Delta$  and the PDE becomes the standard non-homogeneous heat equation! We will later see that many solutions to (5.1) have many properties that the heat kernel is equipped with. For other examples of second-order parabolic PDE, look at the Fokker-Planck, Kolmogorov, and Liouville's equation.

Again, the representation formulas developed in a first course of study in partial differential equations will not particularly work

### 5.1 Weak Solutions.

## 6 Appendix.

### 6.1 Results from Functional Analysis.

This section is dedicated to a review of some of the fundamental results from functional analysis that are used substantially throughout the paper.

**Lemma 6.1.** (*Young's Inequality*). Let  $1 < p, q < \infty$  and  $\frac{1}{p} + \frac{1}{q} = 1$ . Then

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

for  $a, b > 0$ .

*Proof.* We use the fact that the mapping  $x \rightarrow e^x$  is a convex map. Therefore,

$$ab = e^{\log a + \log b} = e^{\frac{1}{p} \log a^p + \frac{1}{q} \log b^q} \leq \frac{1}{p} e^{\log a^p} + \frac{1}{q} e^{\log b^q} = \frac{a^p}{p} + \frac{b^q}{q}$$

which is the desired result.  $\square$

**Lemma 6.2.** (*Hölder's Inequality*). Let  $1 \leq p, q \leq \infty$ ,  $\frac{1}{p} + \frac{1}{q} = 1$ . Then if  $u \in L^p(\Omega)$ ,  $v \in L^q(\Omega)$ , we have

$$\int_{\Omega} |uv| \, dx \leq \|u\|_{L^p(\Omega)} + \|v\|_{L^q(\Omega)}$$

which is the desired result.

*Proof.* Using the absolute homogeneity property of norms let us set, without loss of generality,  $\|u\| = \|v\| = 1$ . Then, by the Young's inequality, we observe that

$$\int_{\Omega} |uv| \, dx \leq \frac{1}{p} \int_{\Omega} |u|^p \, dx + \frac{1}{q} \int_{\Omega} |v|^q \, dx = 1 = \|u\|_{L^p(\Omega)} \leq \|u\|_{L^p(\Omega)} \|v\|_{L^q(\Omega)}.$$

$\square$

**Lemma 6.3.** (*General Hölder's Inequality*). Let  $1 \leq p_1, \dots, p_n \leq \infty$  be integers such that  $\sum_{i=1}^n \frac{1}{p_i} = 1$ . Furthermore, assume that for each  $k = 1, \dots, n$ ,  $u_k \in L^{p_k}(\Omega)$ . Then

$$\int_{\Omega} |u_1 \cdots u_n| \, dx \leq \prod_{k=1}^n \|u_k\|_{L^{p_k}(\Omega)}.$$

*Proof.* We proceed by induction.  $\square$

**Lemma 6.4.** (*Minkowski's Inequality*). Let  $1 \leq p \leq \infty$  and  $u, v \in L^p(\Omega)$ . Then

$$\|u + v\|_{L^p(\Omega)} \leq \|u\|_{L^p(\Omega)} + \|v\|_{L^p(\Omega)}.$$

*Proof.* To prove this, we use a common trick used to prove these types of inequalities.

$$\begin{aligned} \|u + v\|_{L^p(\Omega)}^p &= \int_{\Omega} |u + v|^p \, dx \\ &\leq \int_{\Omega} |u + v|^{p-1} (|u| + |v|) \, dx \end{aligned}$$

which comes from an application of the triangle inequality. Then, an application of Hölder's inequality gives,

$$\begin{aligned} \|u + v\|_{L^p(\Omega)}^p &\leq \left( \int_{\Omega} |u + v|^p \, dx \right)^{\frac{p-1}{p}} \left( \left( \int_{\Omega} |u|^p \, dx \right)^{\frac{1}{p}} + \left( \int_{\Omega} |v|^p \, dx \right)^{\frac{1}{p}} \right) \\ &= \|u + v\|_{L^p(\Omega)}^{p-1} (\|u\|_{L^p(\Omega)} + \|v\|_{L^p(\Omega)}) \end{aligned}$$

from which the desired result follows.  $\square$

**Definition 6.1.** Let  $V$  be a vector space defined over the real or complex numbers. Let  $\|\cdot\|_X$  and  $\|\cdot\|_Y$  be two norms. We say that the two norms are equivalent if, for  $x \in V$ , there exist positive constants  $C_0, C_1$  such that

$$C_0 \|x\|_X \leq \|x\|_Y \leq C_1 \|x\|_X.$$



## 6.2 The Space $H^{-1}$ .

This section is dedicated to studying properties of the Dual space of  $H_0^1(\Omega)$ . Dual spaces are studied for their nice properties, and are a crucial part of functional analysis, however they are hard to find. For this particular space, the dual space can be constructed using a clever application of results from functional analysis. These spaces are important in studying operators on a vector space.

**Definition 6.2.** We let  $H^{-1}(\Omega)$  denote the dual space of  $H_0^1(\Omega)$ .

By the definition of a dual space,  $f$  belongs to the dual space if  $f \in H_0^1(\Omega)$  is a bounded linear functional.

**Definition 6.3.** Let  $x \in X, x^* \in X^*$ . Then  $x^*(x) = \langle x, x^* \rangle$  denotes the pairing between  $x$  and  $x^*$ .

As we have already shown in our treatment of Sobolev spaces, the space  $W^{k,2}(\Omega)$  is a Banach space. Banach spaces are always endowed with a norm, which for  $H^{-1}(\Omega)$  is defined in the following manner.

**Definition 6.4.** Let  $f \in H^{-1}(\Omega)$ . Then the norm with respect to this space is defined by

$$\|f\|_{H^{-1}(\Omega)} := \sup \left\{ \langle f, u \rangle : u \in H_0^1(\Omega), \|u\|_{H_0^1} \leq 1 \right\}.$$

It was also previously mentioned (but not proved) that the space  $H_0^1(\Omega)$  is a Hilbert space under the following inner product.

**Definition 6.5.** Let  $u, v \in H_0^1(\Omega)$ . Then the inner product  $(u, v)$  is given by

$$(u, v) := \int_{\Omega} Du \cdot Dv + uv \, dx.$$

This definition of the inner product will be crucial to show some fundamental results about inhabitants of  $H^{-1}(\Omega)$ .

**Theorem 6.1.** Let  $f \in H^{-1}(\Omega)$ . Then there exist functions  $f^0, f^1, \dots, f^n \in L^2(\Omega)$  such that the pairing is given by

$$\langle f, u \rangle = \int_{\Omega} f^0 u + \sum_{i=1}^n f^i u_{x_i} \, dx \quad (6.1)$$

for some  $u \in H_0^1(\Omega)$ .

*Proof.* Applying the Riesz representation theorem, we deduce that there is a unique function  $u \in H_0^1(\Omega)$  such that for all  $v \in H_0^1(\Omega)$ ,

$$\langle f, v \rangle = \int_{\Omega} Du \cdot Dv + uv \, dx.$$

This establishes that

$$\begin{cases} f^0 = u, \\ f^i = u_{x_i} \quad (i = 1, \dots, n) \end{cases}$$

from which the desired result follows. □

Using the above result, we can provide a much better estimate of the norm of  $f \in H^{-1}(\Omega)$  with respect to  $H^{-1}(\Omega)$ . This is summarized in the following theorem.

**Theorem 6.2.** Let  $f \in H^{-1}(\Omega)$ . Then we have the estimate

$$\|f\|_{H^{-1}(\Omega)} = \inf \left\{ \left( \int_{\Omega} \sum_{i=0}^n |f^i|^2 \, dx \right)^{1/2} : f \text{ satisfies (6.1) for } f^0, \dots, f^n \in L^2(\Omega) \right\}$$

### 6.3 Mollification Theory.

Recall that if  $\Omega \subset \mathbb{R}$  is an open set, then  $\Omega_\epsilon = \{x \in \Omega : \text{dist}(x, \partial\Omega) > \epsilon\}$ .

**Definition 6.6.** Define  $\eta \in C^\infty(\Omega)$  in the following manner:

$$\eta(x) := \begin{cases} C \exp\left(\frac{1}{|x|^2-2}\right), & |x| < 1 \\ 0, & |x| \geq 1. \end{cases}$$

where the constant  $C$  is chosen so that

$$\int_{\mathbb{R}^n} \eta(x) dx = 1.$$

This is called the standard mollifier.

Using this definition of  $\eta$ , we can define variants of the standard mollifier.

**Definition 6.7.** Let  $\epsilon > 0$ . We call

$$\eta_\epsilon(x) := \frac{1}{\epsilon^n} \eta\left(\frac{x}{\epsilon}\right)$$

and  $\epsilon$  variant of the standard mollifier. Furthermore, the support of the function  $\eta_\epsilon$  is contained in  $B(0, \epsilon)$ .

We can also take the derivative with respect to  $x_i$  for  $i = 1, \dots, n$ . It is given by

$$\frac{\partial \eta_\epsilon}{\partial x_i}(x) = \frac{1}{\epsilon^{n+1}} \frac{\partial \eta}{\partial x_i}\left(\frac{x}{\epsilon}\right).$$

**Definition 6.8.** Let  $f : \Omega \rightarrow \mathbb{R}, f \in L_{\text{loc}}(\Omega)$ . Define

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

as the mollification of  $f$  by  $\eta_\epsilon$ . Note that  $f * \eta_\epsilon$  is given by

$$f * \eta_\epsilon = \int_{\Omega} \eta_\epsilon(x-y) f(y) dy = \int_{B(0, \epsilon)} \eta_\epsilon(y) f(x-y) dy$$

for all  $x \in \Omega_\epsilon$ .

Intuitively, if we are given a function that is irregular, we can construct an approximation to the original function, with the sharp points smoothed out, by convoluting it with a mollifier. This intuition provides the motivation for our discussion on constructing dense subsets of  $W^{k,p}(\Omega)$ . The tool of mollifiers will allow for us to approximate functions in  $W^{k,p}(\Omega)$  by functions in  $C^\infty(\Omega)$ .

We now explore properties of mollifiers.

**Theorem 6.3.** Let  $f : \Omega \rightarrow \mathbb{R}, f \in L_{\text{loc}}(\Omega)$ . Then,

1.  $f^\epsilon \in C^\infty(\Omega_\epsilon)$ .
2.  $f^\epsilon \rightarrow f$  almost everywhere as  $\epsilon \rightarrow 0$ .
3. If  $1 \leq p < \infty$  and  $f \in L_{\text{loc}}^p(\Omega)$ , then  $f^\epsilon \rightarrow f$  in  $L_{\text{loc}}^p(\Omega)$ .

*Proof.* (1). Fix  $x \in \Omega_\epsilon, i = 1, \dots, n$  and pick  $h$  small enough so that  $x + he_i \in \Omega_\epsilon$ . Here,  $e_i$  denotes the unit vector where all elements are zero except for the  $i$ -th element. Our goal is to show that the derivative exists, for each  $i$ . By the definition of the derivative, we write

$$\begin{aligned} \frac{f^\epsilon(x + he_i) - f^\epsilon(x)}{h} &= \int_{\Omega} \frac{1}{h} [\eta_\epsilon(x + he_i - y) - \eta_\epsilon(x - y)] f(y) dy \\ &= \frac{1}{\epsilon^n} \int_{\Omega} \frac{1}{h} \left[ \eta\left(\frac{x + he_i - y}{\epsilon}\right) - \eta\left(\frac{x - y}{\epsilon}\right) \right] f(y) dy. \end{aligned}$$

Since

$$\frac{1}{h} \left[ \eta\left(\frac{x + he_i - y}{\epsilon}\right) - \eta\left(\frac{x - y}{\epsilon}\right) \right] \rightarrow \frac{1}{\epsilon} \frac{\partial \eta}{\partial x_i}\left(\frac{x - y}{\epsilon}\right) \text{ as } h \rightarrow 0$$

uniformly on  $\Omega$ , we know that the derivative,  $\frac{\partial f^\epsilon}{\partial x_i}$  exists. Furthermore, it is given by

$$\begin{aligned}\frac{\partial f^\epsilon}{\partial x_i}(x) &= \lim_{h \rightarrow 0} \frac{1}{\epsilon^n} \int_{\Omega} \frac{1}{h} \left[ \eta \left( \frac{x + h e_i - y}{\epsilon} \right) - \eta \left( \frac{x - y}{\epsilon} \right) \right] f(y) dy \\ &= \frac{1}{\epsilon^n} \int_{\Omega} \lim_{h \rightarrow 0} \frac{1}{h} \left[ \eta \left( \frac{x + h e_i - y}{\epsilon} \right) - \eta \left( \frac{x - y}{\epsilon} \right) \right] f(y) dy \\ &= \frac{1}{\epsilon^{n+1}} \int_{\Omega} \frac{\partial \eta}{\partial x_i} \left( \frac{x - y}{\epsilon} \right) f(y) dy \\ &= \int_{\Omega} \frac{\partial \eta_\epsilon}{\partial x_i}(x - y) f(y) dy.\end{aligned}$$

The limit can be taken inside the integral sign, since the convergence is uniform on  $\Omega$ .

By a similar argument, we can show that  $D^\alpha f^\epsilon$  exists and is given by

$$D^\alpha f^\epsilon(x) = \int_{\Omega} D^\alpha \eta_\epsilon(x - y) f(y) dy$$

for all  $x \in \Omega_\epsilon$  and multiindex  $\alpha$ . This proves (1).

(2). Fix a point  $x \in \Omega$ . Note that

$$\begin{aligned}|f^\epsilon(x) - f(x)| &= \left| \int_{\Omega} \eta_\epsilon(x - y) f(y) dy - \int_{\Omega} \eta_\epsilon(x - y) f(x) dy \right| \\ &= \left| \int_{\Omega} \eta_\epsilon(x - y) [f(y) - f(x)] dy \right| \\ &= \left| \int_{B(x, \epsilon)} \eta_\epsilon(x - y) [f(y) - f(x)] dy \right| \\ &\leq \int_{B(x, \epsilon)} |\eta_\epsilon(x - y)| |f(y) - f(x)| dy \\ &= \frac{1}{\epsilon^n} \int_{B(x, \epsilon)} \left| \eta \left( \frac{x - y}{\epsilon} \right) \right| |f(y) - f(x)| dy \\ &\leq C \int_{B(x, \epsilon)} |f(y) - f(x)| dy \rightarrow 0\end{aligned}$$

as  $\epsilon \rightarrow 0$ , almost everywhere by the Lebesgue differentiation theorem. The constant  $C$  comes from the fact that  $\eta_\epsilon$  is a compactly support and continuous function. The barred integral denotes the average over the ball. By the squeeze theorem, we conclude that

$$|f^\epsilon(x) - f(x)| \rightarrow 0 \text{ as } \epsilon \rightarrow 0$$

and thus  $f^\epsilon \rightarrow f$ , as  $\epsilon \rightarrow 0$ , which is the desired result.

(3). Choose an open set  $\omega \subset\subset \Omega$  and another open set  $\omega'$  such that  $\omega \subset\subset \omega' \subset\subset \Omega$ . Borrowing an idea from functional analysis, we assert that for sufficiently small  $\epsilon > 0$ , the following is true

$$\|f^\epsilon\|_{L^p(\omega)} \leq \|f\|_{L^p(\omega')}.$$

To justify this assertion, consider the following argument. Fix some  $x \in \omega$  and note that For  $1 \leq p < \infty$ , we have

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

Applying Hölder's inequality, we further obtain

$$|f^\epsilon(x)| \leq \left( \int_{B(x, \epsilon)} \eta_\epsilon(x - y) dy \right) \left( \int_{B(x, \epsilon)} \eta_\epsilon(x - y) |f(y)|^p dy \right)^{1/p}.$$

The integral of the mollifier over  $B(x, \epsilon)$ , with  $x$  fixed is identically 1 by definition. After raising both sides of the above inequality by  $p$ , we obtain the inequality

$$|f^\epsilon(x)|^p \leq \int_{B(x, \epsilon)} \eta_\epsilon(x-y) |f(y)|^p dy.$$

The left hand side looks almost like the definition of the  $L^p$  norm. Indeed, integrating both sides with respect to  $x$ , we obtain

$$\int_\omega |f^\epsilon(x)|^p dx \leq \int_\omega \left( \int_{B(x, \epsilon)} \eta_\epsilon(x-y) |f(y)|^p dy \right) dx$$

We exchange the order of integration and note that for  $y \in \omega'$ , we have

$$\int_\omega |f^\epsilon(x)|^p dx \leq \int_{\omega'} |f(y)|^p \left( \int_{B(y, \epsilon)} \eta_\epsilon(x-y) dx \right) dy = \int_{\omega'} |f(y)|^p dy$$

where again we have used the fact that the integral of the mollifier over the ball, with  $y$  fixed, is identically 1. This is the desired result.

Next, pick some  $g \in C(\omega')$  such that for some  $\delta > 0$ , we have

$$\|f - g\|_{L^p(\omega')} < \delta.$$

Finally, we have

$$\begin{aligned} \|f^\epsilon - f\|_{L^p(\omega)} &= \|f^\epsilon + g^\epsilon - g^\epsilon + g - g - f\|_{L^p(\omega)} \\ &\leq \|f^\epsilon - g^\epsilon\|_{L^p(\omega)} + \|g - f\|_{L^p(\omega)} + \|g^\epsilon - g\|_{L^p(\omega)} \\ &\leq 2\|f - g\|_{L^p(\omega')} + \|g^\epsilon - g\|_{L^p(\omega)} \\ &< 2\delta + \|g^\epsilon - g\|_{L^p(\omega)}. \end{aligned}$$

The first inequality is due to Minkowski. The second inequality is due to the result

$$\|f^\epsilon\|_{L^p(\omega)} \leq \|f\|_{L^p(\omega')}$$

which was proved above. Finally, note that since  $\omega$  is a compact set of  $\Omega$ , we have that  $g^\epsilon \rightarrow g$ , as  $\epsilon \rightarrow 0$ , uniformly in  $L^p(\omega)$ . Taking the limit supremum, we have

$$\limsup_{\epsilon \rightarrow 0} \|f^\epsilon - f\|_{L^p(\omega)} = 0.$$

Thus we have that  $f^\epsilon \rightarrow f$  in  $L^p_{\text{loc}}(\Omega)$ . □

## 7 References.