# SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

## An introduction to Sobolev spaces

av

**Linus Lidman Bergqvist**

2014 - No 28

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

# An introduction to Sobolev spaces

Linus Lidman Bergqvist

Självständigt arbete i matematik 15 högskolepoäng, Grundnivå

Handledare: Joel Andersson

2014

# An introduction to Sobolev spaces

Linus Lidman Bergqvist

**Abstract**

In this thesis we begin by looking at the wave equation with boundary conditions. We find that certain solutions are discarded due to our $C^2$-requirement, although they by any means could be considered to be solutions to our boundary value problem. But if these functions are considered to be solutions to a partial differential equation, in what sense are they differentiable? And in what function space will they lie?

To answer the first question, we introduce the notion of distributions, which can be regarded as a generalization of the concept of a function. For example, the Dirac delta function is not an ordinary function, but it is a distribution. For distributions, we then introduce the notion of weak, or distributional derivative, which is our desired generalization of the usual derivative.

To answer the second question we define Sobolev spaces, which are spaces of functions that are sufficiently many times differentiable in the weak sense and whose derivatives all belong to some $L^p$-space. We first define Sobolev spaces for non-negative integers $k$, which means that the functions must be $k$ times differentiable in the weak sense. We then extend our definition of Sobolev spaces to arbitrary real numbers. We also define Sobolev spaces for functions defined on the boundary of some open subset of $\mathbb{R}^n$. This can not be done in exactly the same way as for other arbitrary bounded subsets of $\mathbb{R}^n$ since the boundary has volume measure 0, and thus all integrals in our usual Sobolev norm become 0.

We then derive some results concerning Sobolev spaces. First we define the restriction to the boundary of a function in a Sobolev space, which is not trivial since functions in Sobolev spaces are generally only defined up to a set of measure zero, and thus a function in a Sobolev space can be completely redefined on the boundary without affecting it as an object in a Sobolev space. This restriction map is essential since Sobolev spaces are closely related to partial differential equations with boundary conditions.

We then continue by proving that Sobolev spaces are continuously embedded in certain $L^p$-spaces and Hölder spaces.

Finally, we apply our results regarding Sobolev spaces and distributions to prove a theorem regarding existence and uniqueness of solutions to elliptic boundary value problems.

# Contents

# Chapter 1

# Introduction

## 1.1 Motivational example, the wave equation

Let us begin by finding all solutions to the $1+1$-dimensional wave equation

$$c^2 u_{xx} = u_{tt},$$

where $c$ is a constant that can be interpreted as the speed with which the wave propagates. We begin by finding solutions in the open half-plane $t > 0$. The idea here is to introduce new variables, $\eta$ and $\xi$ defined by

$$\eta = x - ct \text{ and } \xi = x + ct.$$

By simply applying the chain rule, the wave equation becomes

$$c^2 \cdot 4 \frac{\partial u}{\partial \xi \partial \eta} = 0 \iff \frac{\partial}{\partial \xi} \frac{\partial u}{\partial \eta} = 0.$$

By integrating this equation step by step we first see that $u_\eta$ is constant in $\xi$ and is thus a function of only $\eta$, say $h(\eta)$. Let $\phi$ be an antiderivative of $h$, then by integrating one more time we see that

$$u = \phi(\eta) + \psi(\xi) = \phi(x - ct) + \psi(x + ct).$$

In the above expression, $\psi$ and $\phi$ are more or less arbitrary functions. But if we require $u$ to be $C^2$, then both $\phi$ and $\psi$ must be $C^2$.

Let us now move on to solve an initial value problem for the wave equation in one dimension. Let $f(x)$ and $g(x)$ be two known functions on $\mathbb{R}$. We want to find all functions $u$ satisfying

$$c^2 u_{xx} = u_{tt}, \quad x \in \mathbb{R}, t > 0, \tag{1.1}$$

$$u(x, 0) = f(x), \quad x \in \mathbb{R}, \tag{1.2}$$

$$u_t(x, 0) = g(x), \quad x \in \mathbb{R}. \tag{1.3}$$

The initial conditions mean that we know what the solution looks like at $t = 0$ and also its rate of change. By our above solution to the wave equation without initial conditions, we know what form our solution must have. So our objective is to determine $\phi$ and $\psi$ for which

$$f(x) = u(x, 0) = \phi(x) + \psi(x), \; g(x) = u_t(x, 0) = -c\phi'(x) + c\psi'(x).$$

By the fundamental theorem of calculus, $G(x) = \int_0^x g(y)dy$ is an antiderivative of $g(x)$. Thus the second equation can be integrated, which yields

$$-\phi(x) + \psi(x) = \frac{1}{c}G(x) + K$$

where $K$ is an arbitrary constant. By combining this with the first formula, we get

$$\phi(x) = \frac{1}{2}\left(f(x) - \frac{1}{c}G(x) - K\right), \; \psi(x) = \frac{1}{2}\left(f(x) + \frac{1}{c}G(x) + K\right)$$

Thus our solution to the initial value problem is given by

$$\begin{aligned} u(x,t) &= \phi(x - ct) + \psi(x + ct) \\ &= \frac{1}{2}(f(x - ct) - \frac{1}{c}G(x - ct) - K + f(x + ct) + \frac{1}{c}G(x + ct) + K) \\ &= \frac{f(x - ct) + f(x + ct)}{2} + \frac{G(x + ct) - G(x - ct)}{2c} \\ &= \frac{f(x - ct) + f(x + ct)}{2} + \frac{1}{2c}\int_{x-ct}^{x+ct} g(y)dy. \end{aligned}$$

This result is known as *d′Alembert′s* formula.

We will now solve a specific initial value problem. It is in itself not of any special interest, but with the solution we get, we can make an important point about solutions to partial differential equations in general.

The initial value problem will be the following

$$u_{xx} = u_{tt}, \quad x > 0, \quad t > 0, \tag{1.4}$$

$$u(x, 0) = 2x \text{ for } x > 0, \tag{1.5}$$

$$u_t(x, 0) = 1 \text{ for } x > 0, \tag{1.6}$$

$$u(0, t) = 2t \text{ for } t > 0. \tag{1.7}$$

Since the first quadrant is convex – and thus the lines $x-t = c$ and $x+t = c$, for some constant $c$, can run unbroken through the entire area – all solutions will be of the form

$$u(x, t) = \phi(x - t) + \psi(x + t), \quad x > 0, \quad t > 0.$$

If however, the lines $x - t = c$ and $x + t = c$ would not run through the first quadrant unbroken, then the above expression would not be the general solution, since in that case $\phi(x-t)$ could be replaced by $\phi_1(x-t)$ in some area and by $\phi_2(x-t)$ in some other area where the two areas can not be directly connected by a straight line. The same argument also applies to the term $\psi(x+t)$.

We will try to find out what the functions $\phi$ and $\psi$ look like. If $t = 0$ we get $2x = u(x, 0) = \phi(x) + \psi(x)$ and $1 = u_t(x, 0) = -\phi'(x) + \psi'(x)$, and for $x = 0$, we have $2t = \phi(-t) + \psi(t)$. Since the name of the variable doesn't matter, we call the lone variable in each equation $s$. The three equations now become

$$2s = \phi(s) + \psi(s) \tag{1.8}$$

$$1 = -\phi'(s) + \psi'(s) \tag{1.9}$$

$$2s = \phi(-s) + \psi(s) \tag{1.10}$$

By integrating the second condition, we get

$$-\phi(s) + \psi(s) = s + C,$$

and by using this together with our first equation, we get

$$\phi(s) = \frac{1}{2}s - \frac{1}{2}C, \quad \psi(s) = \frac{3}{2}s + \frac{1}{2}C \text{ for } s > 0.$$

By the third equation, we have that

$$\phi(s) = -\frac{1}{2}s - \frac{1}{2}C \text{ for } s < 0.$$

We can now put our solution together, we get that

$$u(x,t) = \phi(x-t) + \psi(x+t) = \frac{1}{2}(x-t) + \frac{3}{2}(x+t) = 2x + t \text{ if } x > t > 0,$$

$$u(x,t) = \phi(x-t) + \psi(x+t) = \frac{1}{2}(t-x) + \frac{3}{2}(x+t) = x + 2t \text{ if } 0 < x < t.$$

By looking at the limit as $x$ goes to $t$ (or $t$ goes to $x$), one sees that the solution is in fact continuous. However, it is not differentiable on the line $x = t$ and we therefore do not consider it to be a solution because of our $C^2$ requirement. However, it does not really seem to be any good reason as to why it should not be considered to be a solution. But if we consider it to be a solution of our partial differential equation, in what sense does it have a derivative? And if the function is not $C^2$, in what function space does it lie? To clarify in what sense the above limit is a solution to our PDE, we must go to the theory of distributions. But first, we will go through some useful function spaces and some important results from functional analysis.

For further reading, see [8].

## 1.2 Some useful function spaces and results from functional analysis

**Definition 1.2.1.** $C^k(\Omega)$, where $\Omega$ is an open subset of $\mathbb{R}^n$.

We define $C^k(\Omega)$ as the space of continuous functions on $\Omega$ all of whose derivatives up to order $k$ are also continuous. That is $u \in C^k(\Omega)$ if $\partial^\alpha u \in C(\Omega)$ for every $\alpha \in \mathbb{N}_0^n, |\alpha| \leq k$, where $\mathbb{N}_0 = \{0, 1, 2, 3...\}$. $\alpha \in \mathbb{N}_0^n$ means that $\alpha$ is a multi-index, that is an $n$-tuple whose elements are non-negative integers. The absolute value of a multi-index is defined as the sum of the indices. If $x$ is an $n$-dimensional vector, then $x^\alpha$ is $x_1^{\alpha_1}...x_n^{\alpha_n}$ and $\partial^\alpha$ is defined as $\partial_1^{\alpha_1}...\partial_n^{\alpha_n}$.

We define $C^\infty(\Omega)$ in the same manner.

**Definition 1.2.2.** $C_0^\infty(\Omega)$

We define $C_0^\infty(\Omega)$ as the space of infinitely many times differentiable functions with compact support. The support of a function $f$ is defined as the closure of the set of points at which $f \neq 0$. For functions defined on $\mathbb{R}^n$, having compact support is equivalent to having bounded support.

For example, the function

$$\begin{cases} 0 \text{ if } |x| \geq 1 \\ e^{-\frac{1}{1-|x|^2}} \text{ if } |x| < 1 \end{cases}$$

is such a function.

**Definition 1.2.3.** The Schwartz space, $S(\Omega)$

A function $f \in C^\infty(\mathbb{R}^n)$ is a Schwartz function if $f$ and all its derivatives tend to zero faster than any inverse power of $x$ as $|x| \to \infty$. That is, $f$ is a Schwartz function if there exists a constant C such that:

$$\sup_{x \in \mathbb{R}^n} |x^\alpha \partial^\beta f(x)| \leq C_{\alpha,\beta,f} \text{ where } \alpha, \beta \in \mathbb{N}_0^n.$$

We define the Schwartz space as the space of all Schwartz functions.

It can be proven that $C_0^\infty(\mathbb{R}^n)$ is dense in $S(\mathbb{R}^n)$ with respect to the standard topology on $S(\mathbb{R}^n)$. See [3] for definition.

We will now define $L^p$-spaces and state a few results about them.

**Definition 1.2.4.** $L^p$-norm, $||f||_{L^p(\Omega)}$ and the $L^p(\Omega)$-space

We define $||f||_{L^p(\Omega)}$ for $1 \leq p < \infty$ as $\left(\int_\Omega |f(x)|^p dx\right)^{\frac{1}{p}}$, where the integral is a Lebesgue integral.

For $p = \infty$, we define $||f||_{L^p(\Omega)}$ as $\operatorname{ess}\sup_{x\in\Omega} |f(x)|$.

Using this, we define $L^p(\Omega)$ to be the space

$$\{f : f \text{ is measurable and } ||f||_{L^p(\Omega)} < \infty\}$$

When $p = 2$, $||f||_{L^p(\Omega)}$ is a Hilbert space endowed with the inner product $\langle f, g\rangle = \int_\Omega f(x)\overline{g(x)}dx$, where the integral is once again a Lebesgue integral. We have that $\langle f, f\rangle = ||f||^2_{L^p(\Omega)}$.

It can be shown that $C_0^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for any $p \geq 1$. Since $C_0^\infty(\mathbb{R}^n) \subset S(\mathbb{R}^n)$, it follows that $S(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for $p \geq 1$.

In the definition of a norm, we require that $||v|| = 0$ if and only if $v$ is the zero-vector. However, if we look at the definition of the $L^p$-norm, we see that we can change the value of the function at any set if measure zero without changing its norm. Thus we do not have that the function which is constantly zero is the only function for which $||f||_{L^p(\Omega)} = 0$. To get around this, we say that two functions are equal in $L^p$ if they are equal almost everywhere. Thus $L^p$-spaces are spaces of equivalence classes of functions, where two functions are considered equivalent if the set on which they differ have measure zero, i.e. they are equal almost everywhere.

We will now prove a very useful inequality concerning $L^p$-spaces, namely Hölder's inequality.

**Lemma 1.2.1.** *If $a \geq 0$, $b \geq 0$ and $0 < \lambda < 1$, then*

$$a^\lambda b^{1-\lambda} \leq \lambda a + (1 - \lambda)b$$

*with equality if and only if $a = b$*

*Proof.* The statement is obvious if $b = 0$. If $b \neq 0$, we can divide both expressions by $b$. We now set $t = a/b$. Proving the statement has now been transformed to proving that $t^\lambda \leq \lambda t + (1 - \lambda)$ with equality if and only if $t = 1$. By methods of elementary calculus, it can easily be shown that $t^\lambda - \lambda t$ is strictly decreasing for $t > 1$ and strictly increasing for $t < 1$. It thus attains its maximum value, which is $1 - \lambda$ when $t = 1$. For $t = 1$, we have that $1^\lambda = 1 = \lambda + (1 - \lambda)$. As was to be shown. $\square$

**Theorem 1.2.2** (Hölder's inequality)**.** *Let $1 < p < \infty$ and choose $q$ to satisfy that $1/p + 1/q = 1$, i.e. $q = p/(p-1)$, we call $q$ the Hölder conjugate of $p$. If $f$ and $g$ are measurable on $\Omega$, then*

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

*In particular, we have that if $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$, then $fg \in L^1(\Omega)$. We have equality if and only if $a|f|^p = b|g|^q$ a.e. for some constants $a, b \neq 0$.*

*Proof.* The statement is obviously true if $||f||_{L^p(\Omega)} = 0$ or $||g||_{L^q(\Omega)} = 0$, since that would imply that $f$ or $g$ is 0 a.e., which of course implies that $||fg||_{L^1(\Omega)} = 0$. The statement is also trivial if

$$||f||_{L^p(\Omega)} = \infty \text{ or } ||g||_{L^q(\Omega)} = \infty.$$

We now observe that if the statement holds for two functions $f$ and $g$, then it also holds for every scalar multiple of $f$ and $g$, since if $f$ is replaced by $af$ and $g$ is replaced by $bg$, then both sides of (1) are changed by a factor $|ab|$. Because of this, it suffices to prove the statements for all functions $f$ and $g$ for which $||f||_{L^p(\Omega)} = ||g||_{L^q(\Omega)} = 1$ with equality if and only if $|f|^p = |g|^q$ a.e. To do so, we just apply the above lemma with $a = |f(x)|^p, b = |g(x)|^q$ and $\lambda = 1/p$. This yields

$$|f(x)g(x)| \leq |f(x)|^p/p + |g(x)|^q/q. \tag{2}$$

Integrating both sides yields

$$||fg||_{L^1(\Omega)} \leq p^{-1}\int |f|^p + q^{-1}\int |g|^q = 1/p + 1/q = 1 = ||f||_{L^p(\Omega)}||g||_{L^q(\Omega)}.$$

Equality holds if and only if it holds a.e. in (2), and by the above lemma, we have equality precisely when $|f|^p = |g|^q$ a.e. □

We will now prove another useful inequality concerning $L^p$-spaces. Namely Young's inequality.

**Theorem 1.2.3** (Young's inequality)**.** *Let $1 \leq p \leq \infty$. If $f \in L^1(\mathbb{R}^n)$ and $g \in L^p(\mathbb{R}^n)$, then $f * g \in L^p(\mathbb{R}^n)$ and*

$$||f * g||_{L^p(\mathbb{R}^n)} \leq ||f||_{L^1(\mathbb{R}^n)}||g||_{L^p(\mathbb{R}^n)}.$$

*Proof.* We first consider the case when $1 < p < \infty$. Let $q$ be the conjugate of $p$ in the Hölder's inequality sense. Using Hölder's inequality, we have

$$\begin{aligned}
|(f*g)(x)| = \left|\int f(x-y)g(y)d^n y\right| &\leq \int |f(x-y)|^{1/q}|f(x-y)|^{1/p}|g(y)|d^n y \\
&\leq \left(\int |f(x-y)|d^n y\right)^{1/q}\left(\int |f(x-y)||g(y)|^p d^n y\right)^{1/p} \\
&= ||f||_{L^1(\mathbb{R}^n)}^{1/q}\left(\int |f(x-y)||g(y)|^p d^n y\right)^{1/p}. \quad (1.11)
\end{aligned}$$

Thus, by looking at the $L^p$-norm of $(f*g)(x)$, we get

$$\begin{aligned}
||(f*g)||_{L^p(\mathbb{R}^n)}^p &\leq ||f||_{L^1(\mathbb{R}^n)}^{p/q}\int d^n x \int d^n y |f(x-y)||g(y)|^p \\
&= ||f||_{L^1(\mathbb{R}^n)}^{p/q}\int d^n y \int d^n x |f(x-y)||g(y)|^p \\
&= ||f||_{L^1(\mathbb{R}^n)}^{1+p/q}\int d^n y |g(y)|^p = ||f||_{L^1(\mathbb{R}^n)}^p ||g||_{L^p(\mathbb{R}^n)}^p
\end{aligned}$$

since integrals over $\mathbb{R}^n$ are invariant under translations.

This proves the theorem for the cases when $1 < p < \infty$. In the remaining cases, the theorem follows directly from

$$|(f*g)(x)| \leq \int |f(x-y)||g(y)|d^n y$$

if $p = 1$, and

$$|(f*g)(x)| \leq ||f||_{L^1(\mathbb{R}^n)}||g||_{L^\infty(\mathbb{R}^n)}$$

if $p = \infty$. This finishes the proof. □

**Theorem 1.2.4** (Log-convexity of $L^p$-norms.)**.** *Let $1 \leq p_0 < p_1 < \infty$, $\Omega$ be a bounded open subset of $\mathbb{R}^n$ and $f \in L^{p_0}(\Omega) \cap L^{p_1}(\Omega)$. Then $f \in L^p(\Omega)$ for every $p_0 \leq p \leq p_1$. Furthermore we have that*

$$||f||_{L^{p_\alpha}(\Omega)} \leq ||f||_{L^{p_0}(\Omega)}^{1-\alpha}||f||_{L^{p_1}(\Omega)}^{\alpha}.$$

*for all $0 \leq \alpha \leq 1$, where the exponent $p_\alpha$ is defined as $1/p_\alpha = (1-\alpha)/p_0 + \alpha/p_1$.*

*Proof.* By using Hölder's inequality with $|f|^{(1-\alpha)p_\alpha}, |f|^{\alpha p_\alpha}, p = p_0/(1-\alpha)p_\alpha$ and $q = p_1/(\alpha p_\alpha)$ we have that

$$\begin{aligned}
||f||_{L^{p_\alpha}(\Omega)}^{p_\alpha} = \int_\Omega |f|^{(1-\alpha)p_\alpha}|f|^{\alpha p_\alpha}dx &\leq ||f^{(1-\alpha)p_\alpha}||_{L^{p_0/(1-\alpha)p_\alpha}}||f^{\alpha p_\alpha}||_{L^{p_1/(\alpha p_\alpha)}} \\
&= ||f||_{L^{p_0}(\Omega)}^{(1-\alpha)p_\alpha}||f||_{L^{p_1}(\Omega)}^{\alpha}.
\end{aligned}$$

This finishes the proof. □

**Definition 1.2.5** (Hölder continuity). A function $f \in \Omega$, where $\Omega$ is an open subset of $\mathbb{R}^n$ is said to be Hölder continuous with exponent $0 < \alpha \leq 1$ if:

$$\sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|^\alpha} < C$$

for some constant $C$.

If $\alpha = 1$, we say that $f$ is Lipschitz-continuous, and if $\alpha = 0$, then it is simply bounded.

**Definition 1.2.6** (Hölder spaces $C^{k,\alpha}(\Omega)$). We define $C^{k,\alpha}(\Omega)$ as the space of continuous functions on $\Omega$ having continuous derivatives up to order $k$, for which the $k$:th partial derivatives are Hölder continuous with exponent $\alpha, 0 < \alpha \leq 1$.

The Hölder space $C^{k,\alpha}(\Omega)$ is endowed with the norm $||f||_{C^{k,\alpha}}$ defined as

$$||f||_{C^k(\Omega)} + \max_{|\beta|=k} |\partial^\beta f|_{C^{0,\alpha}}$$

where $\beta$ ranges over multi-indices and $||f||_{C^k(\Omega)} = \max_{|\beta| \leq k} \sup_{x \in \Omega} |\partial^\alpha f(x)|$.

**Definition 1.2.7** (Boundedness and continuity for linear operators). A linear operator between two normed vector spaces, $X$ and $Y$, is said to be a bounded linear operator if there exists some constant $M > 0$ such that

$$||Lv||_Y \leq M||v||_X$$

for all $v$ in $X$.

The linear operator L between $X$ and $Y$ is said to be continuous if

$$x_n \to x \text{ implies that } Tx_n \to Tx.$$

**Theorem 1.2.5** (Equivalence of boundedness and continuity for linear operators). *Let $L$ be a linear operator between two normed spaces $X$ and $Y$. Then $L$ is bounded if and only if it is a continuous linear operator.*

*Proof.* We will first prove that boundedness implies continuity and then that continuity implies boundedness.

Suppose $L$ is bounded. Then for all vectors $v$ and $h$ in $X$, we have

$$||L(v + h) - L(v)|| = ||L(h)|| \leq M||h||.$$

This shows that letting $h$ tend to zero means that $||L(v + h) - L(v)||$ tends to zero, thus proving continuity.

Now assume that $L$ is continuous. Since $L$ is continuous, it is of course continuous at the zero vector. It thus exists a $\delta > 0$ such that $||L(h) - L(0)|| \leq 1$ for all vectors $h$ in $X$ for which $||h|| \leq \delta$. Thus for all non-zero vectors $v$ in $X$ we have

$$||L(v)|| = \left\| \frac{||v||}{\delta} L\left(\frac{\delta v}{||v||}\right) \right\| = \frac{||v||}{\delta} \left\| L\left(\frac{\delta v}{||v||}\right) \right\| \leq \frac{||v||}{\delta} \cdot 1 = \frac{1}{\delta}||v||.$$

This proves that $L$ is bounded. Thereby proving the equivalence of the two statements. $\square$

We will now state and prove the bounded linear transformation theorem, often abbreviated as the B.L.T Theorem. Due to the way we will later on define our Sobolev spaces, this theorem will be of fundamental importance.

**Theorem 1.2.6** (The B.L.T Theorem)**.** *Let $T$ be a bounded linear transformation from a normed vector space $V_1$ to a complete normed vector space $V_2$. Then $T$ can be uniquely extended to a bounded linear transformation (with the same bound) $\overline{T}$ from the completion of $V_1$ to $V_2$.*

*That this completion does in fact exist follows since every normed vector space is also a metric space, with the metric being defined by $d(x, y) = ||x - y||$, and every metric space has a completion. This completion can for example be created by considering equivalence classes of Cauchy sequences.*

*Proof.* Let $y$ be an element of the completion $\overline{V_1}$ of $V_1$. There exists a sequence $\{x_n\}_{n=1}^{\infty} \in V_1$ such that $x_n \to y$. Now look at $||T(x_m) - T(x_n)||$.

We have that $||T(x_m) - T(x_n)|| = ||T(x_m - x_n)|| \leq C||x_m - x_n||$ since $T$ is assumed to be a bounded linear transformation. Since convergent sequences are Cauchy, we have that $||x_m - x_n||$ can be made arbitrarily small by choosing $m$ and $n$ to be large enough. Thus $||T(x_m) - T(x_n)||$ goes to zero and hence $\{T(x_n)\}$ is Cauchy in $V_2$. Since $V_2$ is complete, the limit of $\{T(x_n)\}$ exists.

We will now define the extension.

We define $\overline{T}(y)$ as $\lim_{n\to\infty} T(x_n)$. We now need to show that this extension is well defined. Let $\{x'_n\}$ be another sequence such that $x'_n \to y$. Then $||T(x'_n) - T(x_n)|| = ||T(x'_n - x_n)|| \leq C||x'_n - x_n||$ which tends to zero as $n$ tends to infinity. This implies that $\overline{T}(y)$ is well-defined.

Next we need to show that $\overline{T}$ is linear. Let $x, y \in \overline{V_1}$ and $a, b \in \mathbb{R}$, and let $\{x_n\}$ and $\{y_n\}$ be two sequences such that $x_n \to x$ as $n \to \infty$ and $y_n \to y$ as $n \to \infty$. Then $ax_n + by_n \to ax + by$ as $n \to \infty$. We have that

$\overline{T}(ax + by) = \lim_{n\to\infty} T(ax_n + by_n) = a \lim_{n\to\infty} T(x_n) + b \lim_{n\to\infty} T(y_n) = a\overline{T}(x) + b\overline{T}(y)$.

Hence $\overline{T}$ is linear. Next we need to show that $\overline{T}$ is bounded. Let $y \in \overline{V_1}$, we have

$||\overline{T}(y)|| = \lim_{n\to\infty} ||T(y_n)|| \leq \lim_{n\to\infty} C||y_n|| = C||y||$ since $T$ is bounded. Thus $\overline{T}$ is bounded. If $y \in V_1$, then we can choose $\{y_n\} = y$ for every $n$. It follows that $\overline{T}$ has the same bound as $T$ for $y \in V_1$.

Finally, we now show uniqueness of the extension $\overline{T}$. Assume that there is another extension, call it $T_2$. For any $y \in \overline{V_1}$ there is a sequence $\{y_n\}$ in $V_1$ such that $y_n \to y$ as $n \to \infty$.

$\overline{T} - T_2$ will be a bounded linear transformation, and hence continuous. Using this linear map on the sequence $\{y_n\}$, we see that $(\overline{T} - T_2)(y_n) = 0$ for every $n$. It thus follows that $\overline{T}(y) = T_2(y)$ by continuity.

This finishes the proof. $\square$

We will now state some important definitions and results from functional analysis.

**Definition 1.2.8.** We define a Banach space to be a complete normed vector space, where complete means that every every Cauchy sequence has a well defined limit in the space.

**Definition 1.2.9.** We define a Hilbert space to be a complete inner-product space with a norm adhering from the inner product. We denote by $\langle \cdot, \cdot \rangle$ the inner product of the Hilbert space.

**Definition 1.2.10.** Let $X$ be a vector space over some field $K$, which is either $\mathbb{R}$ or $\mathbb{C}$. A linear map from $X$ to $K$ is called a linear functional on $X$.

**Definition 1.2.11.** If $X$ is a normed vector space, the space $L(X, K)$ of bounded linear functionals on $X$ is called the dual space of $X$.

If $H$ is a Hilbert space, we denote its dual space by $H^*$.

We will need the following lemma in order to continue with our theorem.

**Lemma 1.2.7.** *If $M$ is a closed subspace of $H$ then $H = M \oplus M^\perp$. That is, each $x \in H$ can be uniquely expressed as $x = y + z$ where $y \in M$ and $z \in M^\perp$ and $M^\perp$ is defined as $\{x \in H : \langle x, y\rangle = 0$ for all $y \in M\}$. Furthermore, $y$ and $z$ are the unique elements of $M$ and $M^\perp$ whose distance (in norm) to $x$ is minimal.*

*Proof.* The proof can be found in [4]. □

**Theorem 1.2.8** (Riesz representation theorem). *If $f \in H^*$, there is a unique $y \in H$ such that $f(x) = \langle x, y \rangle$ for all $x \in X$.*

*Proof.* We begin by showing uniqueness. if $\langle x, y \rangle = \langle x, y' \rangle$ for all $x$, then by choosing $x = y - y'$, we conclude that $||y - y'||^2 = 0$, and thus $y = y'$.

We move on to proving existence. If $f$ is the zero functional, then obviously $y = 0$. If not, let $M = \{x \in H : f(x) = 0\}$. Then $M$ is a proper closed subspace of $X$, so $M^\perp$ is non-empty. Now use the previous lemma. Pick $z \in M^\perp$ with $||z|| = 1$. Let $u = f(x)z - f(z)x$. We have that $u \in M$, so

$$0 = \langle u, z \rangle = f(x)||z||^2 - f(z)\langle x, z \rangle = f(x) - \langle x, \overline{f(z)}z \rangle.$$

Thus $f(x) = \langle x, y \rangle$ where $y = \overline{f(z)}z$. □

For further reading on functional analysis, see [4] and [5].

## 1.3 Distributions

A function is defined as a relation between a set of inputs and a set of permissible outputs. Loosely speaking, a function is an object which takes one value and assigns to it another value. However, there are some objects which are very useful, and that resemble functions in how they interact with operations such as integration, but that are not really functions. The most famous example is probably the Dirac delta function. The Dirac delta function is defined as the object, $\delta(t)$, which has the following properties

$$\delta(t) \geq 0 \text{ for } -\infty < t < \infty, \tag{1.12}$$
$$\delta(t) = 0 \text{ for } t \neq 0, \tag{1.13}$$
$$\int_{-\infty}^{\infty} \delta(t)dt = 1. \tag{1.14}$$

Unfortunately, there is no conventional function – which takes one value and assigns to it another value – that posses the properties above. Since if condition (1.13) hold, then condition (1.14) can not hold. However, the object described above is very useful. It have been used throughout history in calculations, which have in some sense turned out to be correct. We want to describe a new class of object which in some sense resemble functions, but that are not as strictly defined as taking values and assigning other values to them. Loosely speaking, we want to define an object that instead of taking values at precise points, takes some sort of weighted average over intervals of positive length. To make our definition exact, we will first need to define something which we will call a test function.

The space of test functions is typically a space of functions which behave nicely in some sense. Typically you want to be able to integrate against them and you do not want to have any trouble with differentiating them. We therefore typically let a test function be an infinitely many times differentiable complex valued function defined on $\mathbb{R}^n$, i.e. $f : \mathbb{R}^n \to \mathbb{C}, f \in C^\infty(\mathbb{R}^n)$. Since we generally do not want any trouble integrating against them, we usually also want our test functions to vanish at infinity, therefore one often let the set of test functions be either $S(\mathbb{R}^n)$ or $C_0^\infty(\mathbb{R}^n)$.

We will now define our distributions in a more concrete way. A distribution is a mapping that assigns to each test function a complex value. So if $f$ is a distribution, we denote this value that $f$ assigns to some test function $\phi$ by $f[\phi]$. Of course, we get different classes of distributions depending on which space of test functions we choose. Since the Fourier transform is a continuous bijection of the Schwartz space onto itself, it is natural to let the space of test functions be the Schwartz space in connection with Fourier analysis. A tempered distribution $f$ is a mapping $f : S(\mathbb{R}^n) \to \mathbb{C}$ which has the following properties

(i) linearity: $f[c_1\phi_1 + c_2\phi_2] = c_1 f[\phi_1] + c_2 f[\phi_2]$

(ii) continuity: if $\phi_j \to \psi$ as $j \to \infty$, then $\lim\limits_{j\to\infty} f[\phi_j] \to f[\psi]$,

for all test functions $\phi_k \in S(\mathbb{R}^n)$ and all scalars $c_k$.

We denote the set of tempered distributions by $S'(\mathbb{R}^n)$.

We will now state some properties of tempered distribution. We say that two tempered distributions, $f$ and $g$, are equal if $f[\phi] = g[\phi]$ for all $\phi \in S(\mathbb{R}^n)$.

For $f, g \in S'(\mathbb{R}^n)$, $f + g$ is defined by $(f + g)[\phi] = f[\phi] + g[\phi]$, for all $\phi \in S'(\mathbb{R}^n)$. If $c$ is a scalar, then the distribution $cf$ is defined by $(cf)[\phi] = c \cdot f[\phi]$. Thus $S'(\mathbb{R}^n)$ is a linear space using these operations. We say that a distribution $f$ is zero on an open interval $(a, b)$ if $f[\phi] = 0$ for all $\phi \in S(\mathbb{R}^n)$ whose support is a subset of $(a, b)$. Two distributions $f$ and $g$ are equal on an open interval $(a, b)$ if their difference, $f - g$ is zero on $(a, b)$. If $f$ and $g$ are ordinary functions, then equality on an open interval $(a, b)$ means that $f = g$ everywhere on $(a, b)$ except possibly on a set of measure zero.

For a distribution that is an integrable function, $f$, we define $f[\phi]$ as $\displaystyle\int f\phi$.

Finally, we arrive at our generalization of derivative. Our motivation of the formulation adheres from integration by parts. If $f$ is a differentiable function and $\phi$ is a test function, where our set of test functions can be the Schwartz space, or smooth functions with compact support, or some other space of functions which vanishes at $\pm\infty$, then by integration by parts, we have

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

for every test function $\phi$. Now even if our notion of derivative only is valid for differentiable functions (by definition) we see that the above condition might be generalized. This inspires the following definition.

**Definition 1.3.1.** If $f \in S'(\mathbb{R})$, a new distribution $f'$ is defined by

$$f'[\phi] = -f[\phi']$$

for all $\phi$ in our space of test functions.

In the same manner, if $f \in S'(\mathbb{R}^n)$, we say that $v$ is the $\alpha$:th weak derivative of $f$ if

$$\int_{\mathbb{R}^n} f(x)\partial^\alpha\phi(x)dx = (-1)^{|\alpha|}\int_{\mathbb{R}^n} v(x)\phi(x)dx.$$

for every $\phi \in S(\mathbb{R}^n)$.

As stated above, if $f$ is an integrable function, then the derivative of $f$ using the above definition will coincide with the usual derivative of $f$ as distributions.

Using this definition, we can for example find the derivative of the Heaviside function $H$, defined as

$$\begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}.$$

By the above definition, we have

$$H'[\phi] = -H[\phi'] = -\int_0^\infty \phi'(x)dx = -[\phi]_{x=0}^\infty = -(0-\phi(0)) = \phi(0) = \delta[\phi].$$

Where the Dirac delta function as a distribution is defined by $\delta[\phi] = \phi(0)$. Thus the distributional derivative of the Heaviside function is the Dirac delta function.

For further reading on distributions, see [8].

# Chapter 2

# Sobolev spaces

## 2.1 Definitions of Sobolev spaces

Recall our boundary value problem from chapter 1. We found a function, which reasonably should be considered to be a solution to our boundary value problem, but which was not differentiable and therefore was discarded due to our $C^2$-requirement. If we were to consider this function as a solution, a few questions needed to be answered. The first was in which sense it could be considered to be differentiable. This question was answered by the definition of weak derivative. But the second question still remains. In which function space does this function and other solutions to partial differential equations lie? We previously noted that the $C^k$-requirements were far too restrictive. It seems to require that the function is integrable in some sense, and thus that it belongs to some $L^p$-space. However, $L^p$-spaces gives no information about the behaviour of the functions weak derivatives. It seems reasonable to require some regularity on the derivatives too. A way of doing this is by demanding that the functions and all its derivatives up to some specified order should be in some $L^p$-space. This is the idea upon which the definitions of Sobolev spaces are built.

**Definition 2.1.1** ($H^l(\Omega), W^{l,p}(\Omega), H^l_0(\Omega)$, where $l \in \mathbb{N}_0, 1 \leq p < \infty$ and $\Omega$ is an open subset of $\mathbb{R}^n$)**.** We begin by defining

$$||u||_{l,\Omega} = \left( \sum_{|\alpha| \leq l} \int_\Omega |\partial^\alpha u(x)|^2 d^n x \right)^{1/2}, \text{ and}$$

$$||u||_{W^{l,p}(\Omega)} = \left( \sum_{|\alpha| \leq l} \int_\Omega |\partial^\alpha u(x)|^p d^n x \right)^{1/p}.$$

and

$$||u||_{W^{k,\infty}(\Omega)} = \max_{0\leq|\alpha|\leq k} ||\partial^\alpha u||_{L^\infty(\Omega)}, \quad p = \infty.$$

We now define $H^l(\Omega)$ as the completion of the inner-product space

$$\{u \in C^l(\Omega), ||u||_{l,\Omega} < \infty\}$$

equipped with the inner product

$$\langle u, v\rangle_{l,\Omega} = \sum_{|\alpha|\leq l} \int_\Omega \partial^\alpha u(x)\overline{\partial^\alpha v(x)} d^n x$$

Similarly, we define $H_0^l(\Omega)$ as the completion of $C_0^\infty(\Omega)$ with respect to the norm $||u||_{l,\Omega}$ adhering from the inner product $\langle u, v\rangle_{l,\Omega}$ and $W^{l,p}(\Omega)$ as the completion of the normed vector space $\{u \in C^l(\Omega), ||u||_{W^{l,p}(\Omega)} < \infty\}$ with respect to the norm $||u||_{W^{l,p}(\Omega)}$. $W_0^{l,p}$ is defined similarly to $H_0^l$. It is worth noting that $W^{l,2} = H^l$.

For the rest of this section, we will mainly be concerned with $H^l(\Omega)$ and $H_0^l(\Omega)$.

It follows from the definition of $||u||_{l,\Omega}$ that if $l' \geq l$, then $||u||_{l,\Omega} \leq ||u||_{l',\Omega}$ for all $u \in H^{l'}(\Omega)$. Every element of $H^l(\Omega)$ can be regarded as an equivalence class of Cauchy sequences. A sequence of $C^l(\Omega)$-functions, $\{u_i\}$ is Cauchy with respect to $||u||_{l,\Omega}$ if and only if $\{\partial^\alpha u_i\}$ is Cauchy for every $\alpha \in \mathbb{N}_0^n, |\alpha| \leq l$. We may thus regard $H^l(\Omega)$ as the set of all $u \in L^2(\Omega)$ for which all weak derivatives up to order $l$ are again in $L^2(\Omega)$. This combined with the fact that $C_0^\infty(\Omega) \subset C^{l'}(\Omega) \subset C^l(\Omega)$ yields that $H_0^{l'}(\Omega) \subset H^{l'}(\Omega) \subset H^l(\Omega) \subset L^2(\Omega)$.

We will now try to generalize the definition of $H^l(\Omega)$ and $H_0^l(\Omega)$ so that we no longer require $l$ to be a non-negative integer.

When $\Omega = \mathbb{R}^n$, using Plancherel's formula, we get that

$$||u||_{l,\mathbb{R}^n}^2 = \int_{\mathbb{R}^n} \sum_{|\alpha|\leq l} k^{2\alpha} |\widehat{u(k)}|^2 \frac{d^n k}{(2\pi)^n}.$$

The right hand side may differ by a multiple of $(2\pi)^n$ depending on your Fourier transform conventions.

We can now replace $||u||_{l,\mathbb{R}^n}$ by the norm

$$||u||_{F,l} = \left[\int_{\mathbb{R}^n} (1 + |k|^2)^l |\widehat{u(k)}|^2 \frac{d^n k}{(2\pi)^n}\right]^{1/2},$$

because there are constants $c$ and $C$ depending only on $n$ and $l$ such that

$$c(\sum_{|\alpha|\leq l} k^{2\alpha}) \leq (1+|k|^2)^l \leq C(\sum_{|\alpha|\leq l} k^{2\alpha}).$$

Therefore, the two norms are equivalent in the sense that a sequence of functions converge in one norms if and only if it converges in the other, and thus, the spaces defined as the completion with respect to one of the norm is the same as the space defined as the completion with respect to the other.

Even though the sum from the first norm only makes sense for $l \in \mathbb{N}_0$, the expression $(1+|k|^2)^l$ makes sense and is positive for all real $l$.

$\left[\int_{\mathbb{R}^n} (1+|k|^2)^l |\widehat{u(k)}|^2 \frac{d^n k}{(2\pi)^n}\right]^{1/2}$ makes sense for all functions $u \in L^2(\mathbb{R}^n)$ (though it may be $+\infty$). We use this in order to define the following.

**Definition 2.1.2** ($H^s(\mathbb{R}^n), s \in \mathbb{R}$)**.** The space $H^s(\mathbb{R}^n)$ is defined as the completion of:

$$\{u \in L^2(\mathbb{R}^n), ||u||_{F,s} < \infty\}$$

equipped with the inner product

$$\langle u, v\rangle_{F,s} = \int_{\mathbb{R}^n} (1+|k|^2)^s \hat{u}(k)\overline{\hat{v}(k)} \frac{d^n k}{(2\pi)^n}.$$

*Remark.* By the B.L.T. theorem, the fact that $H^s(\mathbb{R}^n)$ is defined as the completion of $L^2(\mathbb{R}^n)$ with respect to the norm $||u||_{F,l}$ and since the Schwartz space is dense in $L^2(\mathbb{R}^n)$, the following two statements follow:

1) The Fourier transform, $u \in S(\mathbb{R}^n) \to \hat{u} \in S(\mathbb{R}^n)$, has a unique extension to a bounded linear map

$$F_s : H^s(\mathbb{R}^n) \to \{g : \mathbb{R}^n \to \mathbb{C}^n, \quad g \text{ is measurable}, ||g||_{F,l} < \infty\}.$$

The extension is one-to-one, onto and inner product preserving if the target space is equipped with the inner product $\langle f, g\rangle_{F,s}$.

For convenience, we will persist in denoting $F_s u$ by $\hat{u}$

2) For an $n$-dimensional multi-index $\alpha$, there exists a unique extension of the linear map $u \in S(\mathbb{R}^n) \to \partial^\alpha u \in S(\mathbb{R}^n)$ to a bounded linear map $\partial^\alpha$ : $H^s(\mathbb{R}^n) \to H^{s-|\alpha|}(\mathbb{R}^n)$.

For convenience, we will persist in using the notation $\partial^\alpha$.

**Definition 2.1.3** **($H_0^s(\Omega), s \geq 0$).** If $s$ is an integer, we continue to use Definition 2.1.1. If $s$ is not an integer, we define $H_0^s(\Omega)$ to be the completion of $C_0^\infty(\Omega)$ under the norm $||\cdot||_{F,l}$.

We will now define $H^s(\Omega)$ for negative $s$. The motivation of our definition comes from noting that $H^s(\mathbb{R}^n)^* \cong H^{-s}(\mathbb{R}^n)$, which is proven in the following lemma.

**Lemma 2.1.1.** *Let $s \in \mathbb{R}$.*

*i) Let $u \in H^s(\mathbb{R}^n)$ and $v \in H^{-s}(\mathbb{R}^n)$. Then the map*

$$H^s(\mathbb{R}^n) \times H^{-s}(\mathbb{R}^n) \to \mathbb{C}$$

$$(u, v) \to \int \hat{u}(k)\overline{\hat{v}}(k)\frac{d^n k}{(2\pi)^n}$$

*is sesquilinear, which means that it is linear in the first argument and conjugate linear in the second. The map also obeys*

$$\left|\int \hat{u}(k)\overline{\hat{v}}(k)\frac{d^n k}{(2\pi)^n}\right| \leq ||u||_{F,s}||u||_{F,-s}$$

*ii) If $L \in H^s(\mathbb{R}^n)$, then there exists a $v \in H^{-s}(\mathbb{R}^n)$ such that*

$$Lu = \int \hat{u}(k)\overline{\hat{v}}(k)\frac{d^n k}{(2\pi)^n}$$

*also $||L|| = ||v||_{F,-s}$.*

*Proof.* i) The linearity and conjugate linearity follows immediately from the definition and linearity of the integral. The inequality can be shown using Cauchy Schwarz as follows

$$\begin{aligned}\int \hat{u}(k)\overline{\hat{v}}(k)\frac{d^n k}{(2\pi)^n} &\leq \int |(1+|k|^2)^{s/2}\hat{u}(k)||(1+|k|^2)^{-s/2}\overline{\hat{v}}(k)|\frac{d^n k}{(2\pi)^n} \\ &\leq ||u||_{F,s}||v||_{F,-s} \quad (2.1)\end{aligned}$$

ii) Let $L \in H^s(\mathbb{R})^*$. It follows from the Riesz representation theorem that there exists a $g \in H^s(\mathbb{R}^n)$ for which

$$Lu = \langle u, g\rangle_{F,s} = \int (1+|k|^2)^s \hat{u}(k)\overline{\hat{g}}(k)\frac{d^n k}{(2\pi)^n}.$$

Also $||L|| = ||g||_{F,s}$. If we now choose the $v \in H^{-s}(\mathbb{R}^n)$ which satisfies that $\hat{v}(k) = (1+|k|^2)^s \hat{g}(k)$, this $v$ fulfils the requirements of the lemma. $\square$

**Definition 2.1.4** ($H^s(\Omega), s < 0$)**.** If $s < 0$ we define $H^s(\Omega)$ as the dual space of $H_0^{-s}(\Omega)$, i.e. $H^s(\Omega) := H_0^{-s}(\Omega)^*$

We will also state a few definitions and a theorem which characterizes the space $H^{-1}(\Omega)$.

**Definition 2.1.5.** We will write $\langle \cdot, \cdot \rangle$ to denote the pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$.

**Definition 2.1.6.** $||f||_{H^{-1}(\Omega)}$ is defined by

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

**Theorem 2.1.2** (Characterization of $H^{-1}(\Omega)$)**.** *Let $\Omega$ be a bounded open subset of $\mathbb{R}^n$.*

*i) Let $f \in H^{-1}(\Omega)$, then there exists functions, $f^0, f^1, ..., f^n \in L^2(\Omega)$ such that*

$$\langle f, v\rangle = \int_\Omega f^0 v + \sum_{i=1}^n f^i v_{x_i} dx$$

*where $v \in H^1(\Omega)$,*

*ii) also*

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

*Proof.* The full proof can be found in [3]. $\square$

We now want to define the space $H^l(\partial\Omega)$ in a similar way as we defined $H^l(\Omega)$, but for functions that are only defined on the boundary $\partial\Omega$ of $\Omega$. However, we can not do this in the same way as before, since the $n$-dimensional volume measure of $\partial\Omega$ is 0, and our previous norm therefore is useless. So in order to proceed, we need to define a measure on the boundary. We will do so by using local coordinates.

**Definition 2.1.7.** A diffeomorphism is a differentiable map which is bijective, and whose inverse is also differentiable. If the map and its inverse are $k$ times differentiable with continuous derivatives, we say that it is a $C^k$-diffeomorphism.

**Definition 2.1.8.** a) $\Omega$ is said to have $C^k$ boundary if for each point $p \in \partial\Omega$, there is an open neighbourhood $b(p)$ around $p$ and a $C^k$ diffeomorphism $\phi_p : b(p) \to \mathbb{R}^n$ such that the following hold:

$$\phi_p(b(p) \cap \Omega) = \mathbb{R}^n_+$$

and

$$\phi_p(b(p) \cap \partial\Omega) = \mathbb{R}^{n-1}$$

b) $\Omega$ is said to have smooth boundary if each $\phi_p$ $(p \in \partial\Omega)$ of a) is a $C^\infty$-diffeomorphism.

From here on, we will assume that $\Omega$ is a bounded open subset of $\mathbb{R}^n$ with smooth boundary.

A system $(b(p), \phi_p), p \in \partial\Omega$ as in the previous definition is called a local coordinate system for $\partial\Omega$.

We will now define a few properties for functions defined on the boundary. A function defined on the boundary, $f : \partial\Omega \to \mathbb{C}$, is said to be $C^\infty$ if there exists a coordinate system $(b(p), \phi_p)$ such that $f \circ \phi_p^{-1} : \mathbb{R}^{n-1} \to \mathbb{C}$ is $C^\infty$. We write this as $f \in C^\infty(\partial\Omega)$.

We will now define $H^s(\partial\Omega), s \in \mathbb{R}$ for functions defined on the boundary.

**Definition 2.1.9** ($H^s(\partial\Omega), s \in \mathbb{R}$)**.** Since the boundary of $\Omega$ is compact, every open cover of $\partial\Omega$ has a finite subcover. More specifically, there exists points, $p_i \in \partial\Omega, i = 1, 2, 3, 4...., N$, such that $\partial\Omega \subset \bigcup_{i=1}^{N} b(p_i)$. Now one can choose functions $\chi_i \in C_0^\infty(b(p_i)), 1 \leq i \leq N$, taking values in $[0, 1]$ so that $\sum_{i=1}^{N} \chi_i = 1$ on some neighbourhood of $\partial\Omega$. This collection of smooth functions $\chi_i$ is called a partition of unity.

a) For smooth functions on the boundary, $f$, we define

$$||f||^2_{s,\partial\Omega} = \sum_{i=1}^{N} ||\chi_i f \circ \phi_{p_i}^{-1}||^2_{F,s,n-1}$$

b) We now define $H^s(\partial\Omega)$ as the completion of $C^\infty(\partial\Omega)$ under the norm $||\cdot||_{s,\partial\Omega}$. $H^s(\partial\Omega)$ is also equipped with the inner product

$$\langle f, g\rangle_{s,\partial\Omega} = \sum_{i=1}^{N} \langle (\chi_i f)\circ \phi_{p_i}^{-1}, (\chi_i g)\circ \phi_{p_i}^{-1}\rangle_{s,n-1}.$$

The definition of the Sobolev norm on the boundary seems to depend on how we choose the partition of unity used in the definition of the norm. We will now proceed by proving that all the norms defined by using different partitions of unity are in fact equivalent, and thus the completion with respect to the different norms define the same space.

But in order to prove that, we need to use the following lemma:

**Lemma 2.1.3.** *Let $\psi \in C_0^\infty(\mathbb{R}^m)$ and let $\phi : \mathbb{R}^m \to \mathbb{R}^m$ be a $C^\infty$ diffeomorphism. Then the map: $C_0^\infty(\mathbb{R}^m) \to C_0^\infty(\mathbb{R}^m)$, $f \mapsto (\psi f)\circ\phi^{-1}$ extends to a bounded linear map on $H^s(\mathbb{R}^m)$.*

*We also have that $||(\psi f)\circ\phi^{-1}||_{F,s,m} \leq C||f||_{F,s,m}$*

*The proof idea is to apply the B.L.T. Theorem. Using that $C_0^\infty(\mathbb{R}^m)$ is dense in $H^s(\mathbb{R}^m)$, we see that we can immediately apply the B.L.T. theorem if the map is indeed a bounded linear transformation.*

*The complete proof can be found in* [3].

**Theorem 2.1.4.** *All norms defined using different partitions of unity are equivalent.*

*Proof.* Let $(b(p), \phi_p)$ and $(b(\widetilde{p}), \phi_{\widetilde{p}})$ be two different coordinate systems and let $\chi_i \in C_0^\infty(b(p_i)), 1 \leq i \leq N$ and $\widetilde{\chi_i} \in C_0^\infty(\widetilde{b}(\widetilde{p_i})), 1 \leq i \leq \widetilde{N}$ be two partitions of unity. Then if $f \in C^\infty(\partial\Omega)$, we have that

$$\begin{aligned}(\chi_i f)\circ\phi_{p_i}^{-1} &= 1\cdot(\chi_i f)\circ\phi_{p_i}^{-1} = \sum_{j=1}^{\widetilde{N}} \widetilde{\chi_j}\cdot((\chi_i f)\circ\phi_{p_i}^{-1}) \\ &= \sum_{j=1}^{\widetilde{N}} (\widetilde{\chi_j}\chi_i f)\circ\phi_{p_i}^{-1} = \sum_{j=1}^{\widetilde{N}} (\widetilde{\chi_j}\chi_i f)\circ\widetilde{\phi}_{\widetilde{p_j}}^{-1}\circ\widetilde{\phi}_{\widetilde{p_j}}\circ\phi_{p_i}^{-1}\end{aligned}$$

We now see that $(\widetilde{\chi_j}\chi_i)\circ\phi_{p_i}^{-1}(x)$ will vanish except for when $\phi_{p_i}^{-1}(x) \in \widetilde{b}(\tilde{p_j})\cap b(p_i)$, which means that $x \in \phi_{p_i}(\widetilde{b}(\tilde{p_j})\cap b(p_i))$. And then $\phi_{p_i}^{-1}(x)$ will be in the domain of $\phi_{\widetilde{p_j}}$.

We now apply Lemma 2.1.3 with $\psi = \chi_i\circ\widetilde{\phi}_{\widetilde{p_j}}^{-1}$, $\phi^{-1} = \widetilde{\phi}_{\widetilde{p_j}}\circ\phi_{p_i}^{-1}$, and $f$ replaced by $(\widetilde{\chi_j} f)\circ\widetilde{\phi}_{\widetilde{p_j}}^{-1}$.

We wish to prove that there exists some constant, $K$, such that $||f||^2_{s,\partial\Omega,\chi} \leq K \cdot ||f||^2_{s,\partial\Omega,\widetilde{\chi}}$, where $||f||_{s,\partial\Omega,\chi}$ and $||f||_{s,\partial\Omega,\widetilde{\chi}}$ are the two norms defined by using different partitions of unity. Since $\chi$ and $\widetilde{\chi}$ are arbitrary partitions of unity, this will finish the proof. We only need to show that that the inequality holds one way, since the partitions of unity are arbitrary, and thus by just changing the names we have proved the inequality the other way too.

$$||(\chi_i f) \circ \phi^{-1}||_{F,s,n-1} \leq \sum_{j=1}^{\widetilde{N}} ||(\widetilde{\chi_j}\chi_i f) \circ \widetilde{\phi}^{-1}_{\widetilde{p_j}} \circ \widetilde{\phi}_{\widetilde{p_j}} \circ \phi^{-1}||_{F,s,n-1}$$

$$\leq \sum_{j=1}^{\widetilde{N}} C||(\widetilde{\chi_j} f) \circ \widetilde{\phi}^{-1}_{\widetilde{p_j}}||_{F,s,n-1} \leq C \cdot \widetilde{N} \cdot \max_{1 \leq j \leq \widetilde{N}} ||(\widetilde{\chi_j} f) \circ \widetilde{\phi}^{-1}_{\widetilde{p_j}}||_{F,s,n-1}$$

Since we are dealing with a norm, which is non-negative , the same inequalities are still true if we square every expression. Thus by taking the sum over all indices $i$, we get:

$$\sum_{i=1}^{N} ||(\chi_i f) \circ \phi^{-1}||^2_{F,s,n-1} \leq C^2 \cdot \widetilde{N}^2 \cdot N \cdot \max_{1 \leq j \leq \widetilde{N}} ||(\widetilde{\chi_j} f) \circ \widetilde{\phi}^{-1}_{\widetilde{p_j}}||^2_{F,s,n-1}.$$

This finishes the proof. □

For further reading on Sobolev spaces, see [1] [2] and [3]. For further reading on Fourier analysis, see [7] and [8].

## 2.2 Restrictions to the boundary

We will now study restrictions to the boundary, and their connection to Sobolev spaces. Since a function in a Sobolev space is only defined up to a set of measure zero, and since the volume measure of the boundary is zero, any function in a Sobolev space can be completely redefined on the boundary and still be considered to be the same object in the Sobolev space. Thus we can not define an ordinary function restriction in a meaningful way for Sobolev spaces. What we want to do is to approximate our function by a sequence of other functions which behave nicely on the boundary, and then define the restriction map as the limit of the restriction map of this sequence.

**Definition 2.2.1** ($C^l(\bar{\Omega})$)**.** For an open subset $\Omega$ of $\mathbb{R}^n$, we define $C^l(\bar{\Omega})$ to be the set of all functions $u \in C^l(\mathbb{R}^n)$ for which the partial derivatives up to order $l$, $\partial^\alpha u, |\alpha| \leq l$, are bounded and uniformly continuous on $\Omega$. We write $C^\infty(\bar{\Omega}) = \cap_{l=0}^{\infty} C^l(\bar{\Omega})$.

**Theorem 2.2.1.** *Let $l \in \mathbb{N}$ and let $\Omega$ be an open subset of $\mathbb{R}^n$ with smooth boundary. Now define the restriction map:*

$$r : C^\infty(\bar{\Omega}) \to C^\infty(\partial\Omega)$$

$$u \to u|_{\partial\Omega}$$

*There exists a unique map*

$$R : H^l(\Omega) \to H^{l-1/2}(\partial\Omega)$$

*and constants $C_1, C_2$ such that the following hold:*

(1) *$R$ extends $r$. That means $Ru = ru$ for all $u$ in the domain of $r$, $C^\infty(\bar{\Omega})$.*

(2) *$R$ is bounded. That means $||Ru||_{l-1/2,\partial\Omega} \leq C_1||u||_{l,\Omega}$.*

(3) *$R$ is surjective.*

(4) *For every function $f \in H^{l-1/2}(\partial\Omega)$, there is a function $u \in H^l(\Omega)$ such that $Ru = f$ and $C_2||f||_{l-1/2,\partial\Omega} \geq ||u||_{l,\Omega}$, i.e, $R$ has a bounded right inverse.*

(5) *The kernel of $R$ is $H_0^1(\Omega) \cap H^l(\Omega)$.*

*Remark.* Theorem 2.2.1 is true for all real $l > 1/2$, with the exception that when $1/2 < l < 1$, $R$ has kernel $H_0^l(\Omega)$.

The proof of theorem 2.2.1 will be divided into parts. We will begin by showing that $r$ is bounded. Using this combined with a theorem that states that $C^\infty(\bar{\Omega})$ is dense in $H^l(\Omega)$, we can extend $r$ – by applying the B.L.T. theorem – to all of $H^l(\Omega)$. We call this extension $R$. Part (1), that $R$ extends $r$, and part (2), that $R$ is bounded will then follow immediately. The uniqueness also follows since the extension of the B.L.T. theorem is unique.

The boundedness of $r$ will first be proven for $\Omega = \mathbb{R}^n$ with $\partial\Omega = \{x_n = 0\}$. This result will then be extended to $\Omega = \mathbb{R}_+^n$ with $\partial\Omega = \{x_n = 0\}$ and finally this will be extended to any bounded open subset $\Omega$ of $\mathbb{R}^n$, as in the theorem.

We will then continue by showing that each $f \in H^{l-1/2}(\partial\Omega)$ can be extended to some $u \in H^l(\Omega)$ in a bounded way. This proves part (4), that $R$ has a bounded right inverse, which of course implies surjectivity, which is part (3). In the same manner as with the boundedness, this result will first be proven for $\mathbb{R}^n$ and then be extended to more general subsets.

We obviously have that $r$ vanishes on $C_0^\infty(\Omega)$, which by construction is dense in $H_0^l(\Omega)$. This means, that $r$ being bounded, implies that the kernel of $R$, which is the extension of $r$ by continuity, will contain $H_0^l(\Omega)$. But now note, that if $u \in H_0^1(\Omega) \cap H^l(\Omega)$, then there will exist a sequence of $C_0^\infty(\Omega)$-functions, $u_j$, that converge to $u$ with respect to the norm $||\cdot||_{1,\Omega}$. By part (2) of theorem 2.2.1 with $l = 1$, we have

$$||Ru||_{1/2,\partial\Omega} = ||Ru - Ru_j||_{1/2,\partial\Omega} \leq C||u - u_j||_{1,\Omega}.$$

Since $u_j$ converge to $u$ in the norm $||\cdot||_{1,\Omega}$, $Ru = 0$.

Thus the kernel of $R$ contains all of $H_0^1(\Omega) \cap H^l(\Omega)$. To prove part (5), it thus suffices to show that any function in the kernel of $R$ can be approximated in the $H^1(\Omega)$ norm by functions in $C_0^\infty(\Omega)$. In the same way as above, we begin by showing this for $\mathbb{R}^n$ and then we generalize this result.

We will begin by examining the case when $\Omega = \mathbb{R}^n$. Since $(x_1, x_2, ..., x_{n-1})$ will be used a lot, we will denote $(x_1, x_2, ..., x_{n-1})$ by $x'$ for convenience.

**Lemma 2.2.2** (Restriction from $\mathbb{R}^n$ to $\mathbb{R}^{n-1}$)**.** *Let $s$ be a real number which satisfies that $s > 1/2$. The linear transform*

$$\begin{aligned} r : C_0^\infty(\mathbb{R}^n) \to C_0^\infty(\mathbb{R}^{n-1}) \ &\textit{defined by} \\ (ru)(x', x_n) &= u(x', 0). \end{aligned}$$

*has a unique extension to a bounded linear map*

$$R : H^s(\mathbb{R}^n) \to H^{s-1/2}(\mathbb{R}^{n-1}).$$

*Proof.* The idea is to use the B.L.T. theorem. But in order to apply the B.L.T. theorem, we must show that $r$ is indeed bounded. This will be proven if we can show there exists a constant $C$, depending only on $s$, such that

$$||ru||_{F,s-1/2,n-1} \leq C||u||_{s,n}$$

for all $u \in C_0^\infty(\mathbb{R}^n)$. Once this is shown, the theorem will follow immediately by applying the B.L.T. theorem.

For $k \in \mathbb{R}^n$, we write $k = (k', k_n)$ with $k' = (k_1, k_2, ..., k_{n-1})$. By using the Fourier transform and the inverse Fourier transform, we get

$$\begin{aligned} \int (\widehat{ru})(k') e^{ik'x'} \frac{d^{n-1}k'}{(2\pi)^{n-1}} &= (ru)(x') = u(x', 0) \\ &= \iint \hat{u}(k', k_n) \frac{dk_n}{2\pi} e^{i(k' \cdot x')} \frac{d^{n-1}k'}{(2\pi)^{n-1}}. \end{aligned}$$

By comparing the first and the last integral, we see that $\widehat{ru}(k') = \displaystyle\int \hat{u}(k', k_n) \frac{dk_n}{2\pi}$.

Using this equality and Cauchy-Schwarz yields that

$$
\begin{aligned}
|(\widehat{ru})(k')|^2 &= \left| \int \hat{u}(k', k_n)(1 + k'^2 + k_n^2)^{s/2}(1 + k'^2 + k_n^2)^{-s/2} \frac{dk_n}{2\pi} \right|^2 \\
&\leq \left[ \int |\hat{u}(k', k_n)|^2 (1 + k'^2 + k_n^2)^s \frac{dk_n}{2\pi} \right] \left[ \int \frac{1}{(1 + k'^2 + k_n^2)^s} \frac{dk_n}{2\pi} \right]
\end{aligned}
$$

Setting $k_n = p\sqrt{1 + k'^2}$ in the second integral of the last expression yields

$$
\begin{aligned}
&\left[ \int |\hat{u}(k', k_n)|^2 (1 + k'^2 + k_n^2)^s \frac{dk_n}{2\pi} \right] \left[ \frac{1}{(1 + k'^2)^{s-1/2}} \int \frac{1}{(1 + p^2)^s} \frac{dp}{2\pi} \right] \\
&\qquad \leq \frac{C^2}{(1 + k'^2)^{s-1/2}} \left[ \int |\hat{u}(k', k_n)|^2 (1 + k'^2 + k_n^2)^s \frac{dk_n}{2\pi} \right],
\end{aligned}
$$

where the constant $C^2 = \displaystyle\int \frac{1}{(1 + p^2)^s} \frac{dp}{2\pi}$ depends only on $s$. It is also finite, since $s > 1/2$.

Using the previous inequality, we get that

$$
\begin{aligned}
||ru||^2_{F,s-1/2,n-1} &= \int (1 + k'^2)^{s-1/2} |\widehat{ru}(k')|^2 \frac{d^{n-1}k'}{(2\pi)^{n-1}} \\
\leq C^2 \int |\hat{u}(k', k_n)|^2 (1 + k'^2 + k_n^2)^s \frac{d^n k}{(2\pi)^n} &= (C||u||_{F,s,n})^2.
\end{aligned}
$$

By applying the B.L.T. theorem, this finishes the proof. □

**Lemma 2.2.3** (Extension from $\mathbb{R}^{n-1}$ to $\mathbb{R}^n$)**.** *For any real number $s > 1/2$, there exists a constant $C$, depending only on $n$ and $s$, such that for each $f \in H^{s-1/2}(\mathbb{R}^{n-1})$ there is a $u \in H^s(\mathbb{R}^n)$ for which*

*$Ru = f$ and $||u||_{F,s,n} \leq C||f||_{F,s-1/2,n-1}$, where $R$ is the map from the previous lemma.*

*This proves that the restriction map has a bounded right inverse. This of course also implies surjectivity.*

*Proof.* We claim that if $f \in S(\mathbb{R}^{n-1})$ , then the $u$ determined by

$$
\hat{u}(k', k_n) = \hat{f}(k') 2\sqrt{\pi} e^{-k_n^2/(1+|k'|^2)} / \sqrt{1 + |k'|^2}
$$

satisfies the desired conditions for that $f$.

Once this claim is proven, we will be able to extend the map $f \to u$ by the B.L.T. theorem to $H^{s-1/2}(\mathbb{R}^{n-1})$ and the theorem will follow.

We begin by verifying that $\hat{u}$ is indeed in $S(\mathbb{R}^n)$. Every derivative of

$$e^{-k_n^2/(1+|k'|^2)}/\sqrt{1+|k'|^2}$$

is a polynomial in $k$ multiplied by $e^{-k_n^2/(1+|k'|^2)}(1+|k'|^2)^{-l/2}$ for some natural number $l$. This will be bounded by a polynomial in $k'$ because of the factor $e^{-k_n^2/(1+|k'|^2)}$, and thus, as $f \in S(\mathbb{R}^{n-1}) \to \hat{f} \in S(\mathbb{R}^{n-1})$, $\hat{u} \in S(\mathbb{R}^n)$.

We will need to verify that $Ru = f$. This can equivalently be stated as verifying that $\hat{f}(k') = \mathcal{F}(u(k', 0))$, which by a result in Fourier analysis equals $\displaystyle\int \hat{u}(k', k_n)\frac{dk_n}{2\pi}$.

That this is true follows immediately from the definition of $\hat{u}(k', k_n)$ together with

$$2\sqrt{\pi}\int e^{-k_n^2/(1+|k'|^2)}\frac{dk_n}{2\pi\sqrt{1+|k'|^2}} = \frac{1}{\sqrt{\pi}}\int e^{-p^2}dp = 1.$$

Where we made the change of variables $k_n = p\sqrt{1+|k'|^2}$ in the calculations.

The verification of the boundedness is as follows:

$$||u||_{F,s,n}^2 = \int (1+|k|^2)^s |\hat{u}(k)|^2 \frac{d^n k}{(2\pi)^n}$$

$$= 4\pi \int (1+|k|^2)^s (1+|k'|^2)^{-1} e^{-2k_n^2/(1+|k'|^2)} |\hat{f}(k')|^2 \frac{d^{n-1}k'}{(2\pi)^{n-1}}\frac{dk_n}{2\pi}$$

$$= 2\int (1+p^2)^s e^{-2p^2}dp \int (1+|k'|^2)^{s-1+1/2}|\hat{f}(k')|^2 \frac{d^{n-1}k'}{(2\pi)^{n-1}}$$

$$= \int 2(1+p^2)^s e^{-2p^2}dp ||f||_{F,s-1/2,n-1}^2.$$

Setting $C = \displaystyle\int 2(1+p^2)^s e^{-2p^2}dp$ finishes the proof. □

We will now generalize these results to the case where $\Omega = \mathbb{R}^n_+$ and $\partial\Omega = \mathbb{R}^{n-1}$. In order to do so, we will need the following lemmas:

**Lemma 2.2.4.** *Let $l \in \mathbb{N}_0$. The set of restrictions of functions in $C_0^\infty(\mathbb{R}^n)$ to $\mathbb{R}^n_+$ is dense in $H^l(\mathbb{R}^n_+)$. In particular, $C^\infty(\overline{\mathbb{R}^n_+}) \cap H^l(\mathbb{R}^n_+)$ is dense in $H^l(\mathbb{R}^n_+)$.*

*Proof.* The proof can be found in [3]. □

and

**Lemma 2.2.5.** *Extension from $\mathbb{R}^n_+$ to $\mathbb{R}$.*

*Let $l \in \mathbb{N}_0$. There exists a bounded linear operator*

$$E_+ : H^l(\mathbb{R}^n_+) \to H^l(\mathbb{R}^n)$$

*such that $E_+u(x) = u(x)$ for all $u \in C^\infty(\overline{\mathbb{R}^n_+}) \cap H^l(\mathbb{R}^n_+), x \in \overline{\mathbb{R}^n_+}$.*

*Proof.* The proof can be found in [3]. □

That theorem 2.2.1 holds when $\Omega = \mathbb{R}^n_+$ and $\partial\Omega = \mathbb{R}^{n-1}$ can now be shown by using lemma 2.2.2 and lemma 2.2.3, boundedness and existence of bounded right inverse for $\Omega = \mathbb{R}^n$ and $\partial\Omega = \mathbb{R}^{n-1}$, by constructing $R = \tilde{R}E_+$ where $\tilde{R}$ is the restriction map from $\mathbb{R}^n$ to $\mathbb{R}^{n-1}$ of lemma 2.2.2 and $E_+$ is the extension map from $\mathbb{R}^n_+$ to $\mathbb{R}^n$ of lemma 2.2.5.

The statement regarding the kernel of $R$ will now be proven.

**Lemma 2.2.6.** *Let $R$ be the restriction map of theorem 2.2.1 with $\Omega = \mathbb{R}^n$ and $\partial\Omega = \mathbb{R}^{n-1}$, then the kernel of $R$ is $H^l(\mathbb{R}^n_+) \cap H^1_0(\mathbb{R}^n_+)$.*

*Proof.* We will begin by proving that $H^l(\mathbb{R}^n_+) \cap H^1_0(\mathbb{R}^n_+)$ is contained in the kernel.

Every function $u_j \in C^\infty_0(\mathbb{R}^n_+)$ will naturally be in the kernel of the restriction map. Now since $C^\infty_0(\mathbb{R}^n_+)$ is dense in $H^1_0(\mathbb{R}^n_+)$, for every function $u \in H^l(\mathbb{R}^n_+) \cap H^1_0(\mathbb{R}^n_+)$, there exists a sequence of functions, $\{u_j\}$, that converge to $u$ with respect to the norm $|| \cdot ||_{1,\mathbb{R}^n_+}$.

By part 2) of theorem 2.2.1, with $l = 1$, we get:

$$||Ru||_{F,1/2,n-1} = ||Ru - Ru_j||_{F,1/2,n-1} \leq C||u - u_j||_{1,\mathbb{R}^n_+}.$$

Since this holds for all $j$, and since $||u - u_j||_{1,\mathbb{R}^n_+} \to 0$ as $j \to \infty$, it follows that $Ru = 0$. Thus the kernel of $R$ contains all of $H^l(\mathbb{R}^n_+) \cap H^1_0(\mathbb{R}^n_+)$.

To prove that the kernel of $R$ is contained in $H^l(\mathbb{R}^n_+) \cap H^1_0(\mathbb{R}^n_+)$, we will show that any function in the kernel of $R$ can be approximated in the $H^1(\mathbb{R}^n_+)$ norm by functions in $C^\infty_0(\mathbb{R}^n_+)$

We now define a function $\phi$ which has the following properties:

1. $\phi \in C^\infty(\mathbb{R})$
2. $\phi(t) = 0$ if $|t| \leq \dfrac{1}{2}$
3. $\phi(t) = 1$ if $|t| > 1$

The proof will be carried out in two steps. In the first, we prove the following

If $u \in H^1(\mathbb{R}^n_+)$, we have some $\epsilon > 0$ and we define $u_\epsilon(x) = u(x', x_n)\phi(x_n/\epsilon)$, then

$$||u - u_\epsilon||^2_{1,\mathbb{R}^n} \leq 2\psi/\epsilon||Ru||^2_{L^2(\mathbb{R}^{n-1})} + 2(1+\psi)||u||^2_{1,\Omega_\epsilon}, \text{ where}$$

$$\Omega_\epsilon = \{x \in \mathbb{R}^n, x_n < \epsilon\}, \text{ and } \psi = \max\{\phi'(t)^2, 1/2 \leq t \leq 1\}.$$

Since multiplication by $\phi(x_n/\epsilon)$ is a bounded operator on $H^1(\mathbb{R}^n)$ and by the standard density argument, we can use the B.L.T. theorem, and it thus suffices to prove the claim for $u \in C^\infty(\overline{\mathbb{R}^n_+})$. For $u \in C^\infty(\overline{\mathbb{R}^n_+})$, we see that $u_\epsilon$ is in $H^1_0(\mathbb{R}^n_+)$, since by creating a sequence of functions, $\phi_j$, which have the following properties

1. $\phi_j \in C^\infty_0(\mathbb{R})$
2. $\phi_j(t) = 0$ if $|t| \leq \dfrac{1}{2}$ or $|t| \geq j + 1/2$
3. $\phi_j(t) = 1$ if $j > |t| > 1$,

we can regard $u_\epsilon$ as the limit of $\{u_{\epsilon,j}\}$, where $\{u_{\epsilon,j}\}$ consists of $u$ multiplied with $\phi_j$. Every function in $\{u_{\epsilon,j}\}$ will be in $C^\infty_0(\Omega)$ since both $u$ and $\phi_j$ are smooth, and thus so is their product. Also since $\phi_j$ has compact support, so does $u_{\epsilon,j}$. Once again using that multiplication by $\phi(x_n/\epsilon)$ is a bounded operator on $H^1(\mathbb{R}^n)$, $\{u_{\epsilon,j}\}$ converge in the norm, and since $H^1_0(\mathbb{R}^n_+)$ is complete, $u_\epsilon \in H^1_0(\Omega)$.

By recalling the definition of the Sobolev norm as the sum of $L^p$-norms, a bound for the sum of the $L^p$-norms for the function and its derivatives will also be a bound for the Sobolev norm. With this in mind, we note the following

Note that

$$||u - u_\epsilon||^2_{L^2(\mathbb{R}^n_+)} = ||(1 - \phi(x_n/\epsilon))u||^2_{L^2(\mathbb{R}^n_+)} \leq \int_{0<x_n<\epsilon} |u(x)|^2 d^n x \quad (1)$$

by simply looking at the definition of $1 - \phi(x_n/\epsilon)$ and observing that it equals 0 for $x_n \geq \epsilon$.

We will now try to find similar bounds for the derivatives. For $\dfrac{\partial}{\partial x_j}, 1 \leq j < n$, this is easy since $1 - \phi(x_n/\epsilon)$ is constant in $x_j$, and the derivative is thus zero. In a similar way as above, we get the following

For $1 \leq j \leq n-1$, we have that

$$\frac{\partial}{\partial x_j}[\phi(x_n/\epsilon)u(x)] = \phi(x_n/\epsilon)\frac{\partial u}{\partial x_j}(x).$$

Thus

$$\begin{aligned}\left\|\frac{\partial}{\partial x_j}(u - u_\epsilon)\right\|^2_{L^2(\mathbb{R}^n_+)} &= \left\|(1 - \phi(x_n/\epsilon))\frac{\partial u}{\partial x_j}(x)\right\|^2_{L^2(\mathbb{R}^n_+)} \\ &\leq \int_{0 \leq x_n \leq \epsilon} \left|\frac{\partial u}{\partial x_j}(x)\right|^2 d^n x. \end{aligned} \tag{2}$$

And finally, when $j = n$, we have that

$$\frac{\partial}{\partial x_n}[u(x) - \phi(x_n/\epsilon)u(x)] = [1 - \phi(x_n/\epsilon)]\frac{\partial u}{\partial x_n}(x) - \frac{1}{\epsilon}\phi'(x_n/\epsilon)u(x).$$

Using that $(a+b)^2 \leq 2a^2 + 2b^2$ on the above expression, we get that

$$\left\|\frac{\partial}{\partial x_n}[u - u_\epsilon]\right\|^2_{L^2(\mathbb{R}^n_+)} \leq 2\left\|[1 - \phi(x_n/\epsilon)]\frac{\partial u}{\partial x_n}(x)\right\|^2_{L^2(\mathbb{R}^n_+)} + 2\left\|\frac{1}{\epsilon}\phi'(x_n/\epsilon)u(x)\right\|^2_{L^2(\mathbb{R}^n_+)}.$$

In the same manner as before, we see that the first term is bounded by

$$2\int_{0 \leq x_n \leq \epsilon}\left|\frac{\partial u}{\partial x_n}(x)\right|^2 d^n x. \tag{3}$$

Since $\phi(t)$ is constant for $|t| \leq 1/2$ and for $|t| \geq 1$, we see that $\phi'(x_n/\epsilon)$ vanishes for $0 \leq x_n \leq \epsilon/2$ and for $x_n \geq \epsilon$. Because of this, we see that

$$\begin{aligned} 2\|\frac{1}{\epsilon}\phi'(x_n/\epsilon)u(x)\|^2_{L^2(\mathbb{R}^n_+)} &= \frac{2}{\epsilon^2}\int_{\frac{\epsilon}{2} < x_n < \epsilon}\phi'(x_n/\epsilon)^2|u(x)|^2 d^n x \\ = \frac{2}{\epsilon}\int_{\frac{1}{2} < y < 1}\phi'(y)^2|u(x', \epsilon y)|^2 d^{n-1}x' dy &\leq 2\frac{\psi}{\epsilon}\int_{1/2 < y < 1}|u(x', \epsilon y)|^2 d^{n-1}x dy \end{aligned}$$

Where we have made the change of variables $x_n = \epsilon y$ and where $\psi$ is defined in the same way as before.

By the fundamental theorem of calculus, we see that

$$\begin{aligned} u(x', \epsilon y) &= u(x', 0) + \int_0^{\epsilon y}\frac{\partial u}{\partial x_n}(x', x_n)dx_n \\ &= (Ru)(x') + \int_0^{\epsilon y}\frac{\partial u}{\partial x_n}(x', x_n)dx_n \end{aligned}$$

Now, once again using that $(a+b)^2 \leq 2a^2 + 2b^2$ together with Cauchy-Schwarz, we see that

$$|u(x', \epsilon y)|^2 \leq 2|(Ru)(x')|^2 + 2\epsilon \int_0^\epsilon |\frac{\partial u}{\partial x_n}(x', x_n)|^2 dx_n \text{ for all } y \in (0,1).$$

Thus

$$\begin{aligned} 2||\frac{1}{\epsilon}\phi'(x_n/\epsilon)u(x)||^2_{L^2(\mathbb{R}^n_+)} \\ \leq 2\frac{\psi}{\epsilon}\int |(Ru)(x')|^2 d^{n-1}x' + 2\psi \int_{0<x_n<\epsilon} |\frac{\partial u}{\partial x_n}(x)|^2 d^n x. \quad (4) \end{aligned}$$

Adding $(1), (2), (3)$ and $(4)$ finishes the proof of step 1.

Step 2 is simply to finish the proof of Lemma 2.2.6. This is done by applying the above result.

Let $u \in H^l(\mathbb{R}^n_+)$ be in the kernel of $R$. Using the result of step 1 and using that $u$ is by assumption in the kernel of $R$ – and thus the $2\psi/\epsilon||Ru||^2_{L^2\mathbb{R}^{n-1}}$ term from step 1 is zero – we see that:

$$\lim_{\epsilon\to 0} ||u - u_\epsilon||^2_{1,\mathbb{R}^n_+} \leq 2(1+\psi) \lim_{\epsilon \to 0} ||u||^2_{1,\Omega_\epsilon} = 0$$

This finishes the proof. □

This result can now be generalized to the case when $\Omega$ is an arbitrary open subset of $\mathbb{R}^n$. But in order to generalize the theorem, the following results about $H^l(\Omega)$ are needed.

**Lemma 2.2.7** (Extension from $\Omega$ to $\mathbb{R}^n$)**.** *Let $\Omega$ be a bounded open subset of $\mathbb{R}^n$ with smooth boundary, let $l \in \mathbb{N}_0$ and let $O$ be an open subset of $\mathbb{R}^n$ for which $\overline{\Omega} \cap \overline{O} = \emptyset$. Then there exists a bounded linear operator*

$$E : H^l(\Omega) \to H^l(\mathbb{R}^n) \textit{ for which}$$

$$Eu(x) = u(x) \textit{ a.e. for } x \in \Omega \textit{ and}$$
$$Eu(x) = 0 \textit{ a.e. for } x \in O.$$

*Proof.* The proof can be found in [3]. □

and

**Lemma 2.2.8.** *Let $\Omega$ be a bounded open subset of $\mathbb{R}^n$ with smooth boundary and let $l \in \mathbb{N}_0$. Then $C^\infty(\overline{\Omega})$ is dense in $H^l(\Omega)$.*

*Proof.* Let $u \in H^l(\Omega)$ and $Eu \in H^l(\mathbb{R}^n)$ be its extension to $\mathbb{R}^n$ given by the previous lemma with some $O$ satisfying the properties of $O$ in the previous lemma (for example $O = \{x \in \mathbb{R}^n, |x-y| > 1 \text{ for all } y \in \overline{\Omega}\}$). Since $C_0^\infty(\mathbb{R}^n)$ is dense in $S(\mathbb{R}^n)$, and $S(\mathbb{R}^n)$ is dense in $H^l(\mathbb{R}^n)$ by construction, we have that $C_0^\infty(\mathbb{R}^n)$ is dense in $H^l(\mathbb{R}^n)$. Because of this, there is a sequence of functions, $f_j \in C_0^\infty(\mathbb{R}^n)$ that converges to $Eu$ in $H^l(\mathbb{R}^n)$. Now let $P : H^l(\mathbb{R}^n) \to H^l(\Omega)$ be the operator that restricts functions on $\mathbb{R}^n$ to $\Omega$. We then have:

$$\lim_{j\to\infty} ||u - Pf_j||_{l,\Omega} = \lim_{j\to\infty} ||P(Eu - f_j)||_{l,\Omega} \leq \lim_{j\to\infty} ||Eu - f_j||_{F,l,n} = 0$$

This finishes the proof. $\square$

The general case of theorem 2.2.1 can now be proven by using the extension from $\Omega$ to $\mathbb{R}^n$ of lemma 2.2.8 together with the already proven case of theorem 2.2.1 when $\Omega = \mathbb{R}^n_+$.

The complete proof of theorem 2.2.1 can be found in [3].

For further reading, see [1], [2] and [3].

## 2.3 Embedding theorems for Sobolev spaces

We will now prove several embedding theorems for Sobolev spaces. In order to do so, we must first go through some important concepts and definitions, among these the definition of continuous embedding.

**Definition 2.3.1** (Continuous embedding)**.** Let $X$ and $Y$ be two normed vector spaces, endowed with their respective norms $|| \cdot ||_X$ and $|| \cdot ||_Y$, for which $X \subset Y$. We say that $X$ is continuously embedded in $Y$, which we will denote by $X \hookrightarrow Y$, if the inclusion operator $i : X \to Y, i(x) = x$ is a continuous map. This is equivalent to the statement that there exists some constant $C$ such that:

$$||x||_Y \leq C||x||_X, \forall x \in X.$$

In order to continue with the embedding theorems, we will need the following extension theorem.

**Theorem 2.3.1** (Extension from $W^{1,p}(\Omega)$ to $W^{1,p}(\mathbb{R}^n)$)**.** *Let $\Omega$ be an open subset of $\mathbb{R}^n$ with $C^1$ boundary $\partial\Omega$. For $1 \leq p \leq \infty$, there exists a continuous linear operator*

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

*with the properties*

1. *$Eu = u$ a.e. in $\Omega$ and*
2. *$||Eu||_{W^{1,p}(\mathbb{R}^n)} \leq C||u||_{W^{1,p}(\Omega)}$.*

*Proof.* The proof can be found in [6]. □

We also need to use the following density theorem:

**Theorem 2.3.2** ($C_0^\infty(\mathbb{R}^n)$ is dense in $W^{1,p}(\mathbb{R}^n)$)**.**

*Proof.* The proof can be found in [1]. □

We will now begin by showing the embedding theorems for the case when $\Omega = \mathbb{R}^n$. This result will later be used to generalize the theorem to more general subsets.

**Case 1** ($p < n$)**.**

**Theorem 2.3.3** (Sobolev-Gagliardo-Nirenberg-theorem)**.** *For $1 \leq p < n$. Let $p*$ be defined by*

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

*Then $W^{1,p}(\mathbb{R}^n) \hookrightarrow L^{p*}(\mathbb{R}^n)$, which will be shown by showing that there exists a constant $C = C(n,p)$ such that*

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

Before we start with the proof, we will tell a bit about $p*$.

$p*$ is called the Sobolev exponent corresponding to $p$ and it is in fact the only exponent for which the above inequality can hold. This can be shown by the following scaling argument.

Since $C_0^\infty(\mathbb{R}^n)$ is dense in $W^{1,p}(\mathbb{R}^n)$, we may assume that $u \in C_0^\infty(\mathbb{R}^n)$ instead. Given $u \in C_0^\infty(\mathbb{R}^n)$, define $u_\lambda(x) = u(\lambda x)$ This will be our scaled function for $\lambda > 0$.

If we apply the Sobolev-Gagliardo-Nirenberg-inequality to our function $u_\lambda(x)$, we get

$$||u_\lambda||_{L^{p*}(\mathbb{R}^n)} \leq C||\nabla u_\lambda||_{L^p(\mathbb{R}^n)}.$$

By making a change of variables on both sides, we get

$$\frac{1}{\lambda^{n/p*}}||u||_{L^{p*}(\mathbb{R}^n)} \leq C\frac{\lambda}{\lambda^{n/p}}||\nabla u||_{L^p(\mathbb{R}^n)}$$

which is equivalent to

$$||u||_{L^{p*}(\mathbb{R}^n)} \leq C\lambda^{1-n/p+n/p*}||\nabla u||_{L^p(\mathbb{R}^n)}$$

If this is to be true for all $\lambda > 0$ the exponent $(1 - n/p + n/p*)$ must be 0. From this, it follows that $p*$ must equal $np/(n - p)$.

In order to continue with the proof of theorem 2.3.3, we must first prove the following lemma.

**Lemma 2.3.4.** $||\Pi_{i=1}^n f_i^{\frac{1}{n-1}}||_{L^1(\mathbb{R}^n)} \leq (\Pi_{i=1}^n ||f_i||_{L^1(\mathbb{R}^{n-1})})^{\frac{1}{n-1}}$

*Where $f_i$ is a function of $(x_1, x_2, ..., x_{i-1}, x_{i+1}, ..., x_n)$ and $f_i \geq 0$.*

*Proof.* The proof will be carried out by induction on n.

Let $x_i' = (x_1, x_2, ..., x_{i-1}, x_{i+1}, ..., x_n)$. We begin with the case where $n = 2$. Using Fubini's theorem, we get

$$\begin{aligned}\int_{\mathbb{R}^2} \Pi_{i=1}^2 f_i(x_i')dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1(x_2)f_2(x_1)dx_1dx_2 \\ = \Pi_{i=1}^2 \int_{\mathbb{R}} f_i(x)dx \leq \Pi_{i=1}^2 ||f_i||_{L^1(\mathbb{R})}.\end{aligned}$$

Thus the initial case is proven. We continue by showing that if the statement holds for some number $n$, then it also holds for the next number, $n+1$. By keeping $x_{n+1}$ fixed, by using Hölder's inequality with $p = n$ and $q = \dfrac{n}{n-1}$ and finally by using the induction hypothesis for $n$, we get

$$\begin{aligned}\int_{\mathbb{R}^n} \Pi_{i=1}^{n+1}(f_i(x_i'))^{\frac{1}{n}}dx \leq ||f_{n+1}||_{L^1(\mathbb{R}^n)}^{\frac{1}{n}} \left(\int_{\mathbb{R}^n} \Pi_{i=1}^n (f_i(x_i'))^{\frac{1}{n-1}}dx\right)^{\frac{n-1}{n}} \\ \leq ||f_{n+1}||_{L^1(\mathbb{R}^n)}^{\frac{1}{n}} \Pi_{i=1}^n \left(\int_{\mathbb{R}^{n-1}} |f_i(y, x_{n+1})|dy\right)^{\frac{1}{n}}. \qquad (1)\end{aligned}$$

By integrating the product on the right hand side and using Hölder's inequality again, this time with $1 = \dfrac{1}{n} + ... + \dfrac{1}{n}$, we get

$$\int_{-\infty}^{\infty} \Pi_{i=1}^n \left(\int_{\mathbb{R}^{n-1}} |f_i(y, x_n)|dy\right)^{\frac{1}{n}} dx_{n+1} \leq \Pi_{i=1}^n ||f_i||_{L^1(\mathbb{R}^n)}^{\frac{1}{n}}.$$

By integrating inequality (1) and using the last inequality, we see that

$$\int_{\mathbb{R}^{n+1}} (\Pi_{i=1}^{n+1} f_i(x_i'))^{\frac{1}{n}} dx_1...dx_{n+1} \leq (\Pi_{i=1}^{n+1} ||f_i||_{L^1(\mathbb{R}^n)})^{\frac{1}{n}}.$$

Thus the induction step is proven. The whole lemma now follows by induction. □

We can now proceed with the proof of theorem 2.3.3

*Proof.* of theorem 2.3.3:

As in the scaling argument, we may assume that $u$ is in $C_0^\infty(\mathbb{R}^n)$ since $C_0^\infty(\mathbb{R}^n)$ is dense in $W^{1,p}(\mathbb{R}^n)$. We will split the proof up in two cases, when $p = 1$ and when $1 < p < n$.

*i)* $p = 1$: Let $i \in \{1, ..., n\}$ be fixed. Since $u \in C_0^\infty(\mathbb{R}^n)$ by assumption, we have

$$\begin{aligned} |u(x)| = |u(x_1, ..., x_n) - \lim_{s\to-\infty} u(x_1, ..., x_{i-1}, s, x_{i+1}, ..., x_n)| \\ = \left| \int_{-\infty}^{x_i} \partial_i u(x_1, ..., x_{i-1}, s, x_{i+1}, ..., x_n) ds \right| \\ \leq \int_{\mathbb{R}} |\partial_i u(x_1, ..., x_{i-1}, s, x_{i+1}, ..., x_n)| ds \end{aligned}$$

We define

$$f_i(x_i') = \int_{\mathbb{R}} |\partial_i u(x_1, ..., x_{i-1}, s, x_{i+1}, ..., x_n)| ds,$$

where once again $x_i' = (x_1, ..., x_{i-1}, x_{i+1}, ..., x_n)$.

Using the last inequality for every integer $1 \leq i \leq n$ and multiplying, we get

$$|u(x)|^{\frac{n}{n-1}} \leq \Pi_{i=1}^n (f_i(x_i'))^{\frac{1}{n-1}}.$$

By integrating both sides in the above inequality, using lemma 2.3.4 and raising both sides to the power $(n-1)/n$ yields

$$||u||_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \leq (\Pi_{i=1}^n ||f_i||_{L^1(\mathbb{R}^{n-1})})^{\frac{1}{n}}.$$

By replacing the $f_i$ in the above expression by its original definition, using the inequality of arithmetic and geometric means we get

$$(\Pi_{i=1}^n ||f_i||_{L^1(\mathbb{R}^{n-1})})^{\frac{1}{n}} = (\Pi_{i=1}^n ||\partial_i u||_{L^1(\mathbb{R}^n)})^{\frac{1}{n}} \leq \frac{1}{n} \sum_{i=1}^n ||\partial_i u||_{L^1(\mathbb{R}^n)}$$

and finally we get that

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n ||\partial_i u||_{L^1(\mathbb{R}^n)} = \frac{1}{n} \int_{\mathbb{R}^n} \sum_{i=1}^n |\partial_i u| dx \\ \leq \frac{1}{\sqrt{n}} \int_{\mathbb{R}^n} \left( \sum_{i=1}^n |\partial_i u|^2 \right)^{1/2} dx = \frac{1}{\sqrt{n}} ||\nabla u||_{L^1(\mathbb{R}^n)}. \quad (1) \end{aligned}$$

where the inequality holds since

$$\frac{1}{n}\sum_{i=1}^{n}|\partial_i u| \leq \frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n}|\partial_i u|^2\right)^{1/2}$$

which in turn can be shown to hold since

$$\left(\sum_{i=1}^{n}|\partial_i u|\right)^2 \leq n\sum_{i=1}^{n}|\partial_i u|^2$$

which is ultimately true since

$$2|\partial_i u||\partial_j u| \leq |\partial_i u|^2 + |\partial_j u|^2$$

by Cauchy Schwarz.

Setting $C(n,p) = 1/\sqrt{n}$ finishes the proof for the case $p = 1$.

ii) $1 < p < n$:

Let $v = |u|^\gamma$ for some $\gamma \geq 1$ which will be determined later, and let $q$ be the conjugate of $p$. That is conjugate in the Hölder's inequality sense.

By raising both sides of $v = |u|^\gamma$ by the power $n/(n-1)$ and integrating, we get

$$||u||^{\gamma}_{L^{\frac{n\gamma}{n-1}}(\mathbb{R}^n)} = ||v||_{L^{\frac{n}{n-1}}(\mathbb{R}^n)}.$$

and by using the result from the case $p = 1$, we see that

$$\begin{aligned} ||v||_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \leq \frac{1}{\sqrt{n}}||\nabla v||_{L^1(\mathbb{R}^n)} &= \frac{\gamma}{\sqrt{n}}\int_{\mathbb{R}^n}|\nabla u||u|^{\gamma-1}dx \\ &\leq \frac{\gamma}{\sqrt{n}}||\nabla u||_{L^p(\mathbb{R}^n)}||u||^{\gamma-1}_{L^{(\gamma-1)q}(\mathbb{R}^n)}. \end{aligned} \tag{2}$$

Where we have used Hölder's inequality in the last step. We now make our choice of $\gamma$. We choose $\gamma$ to be equal to $\dfrac{p(n-1)}{n-p}$. Thus our above estimate is

$$||u||_{L^{p*}(\mathbb{R}^n)} \leq \frac{\gamma}{\sqrt{n}}||\nabla u||_{L^p(\mathbb{R}^n)}.$$

Setting $C(n,p) = \dfrac{\gamma}{\sqrt{n}} = \dfrac{p(n-1)}{(n-p)\sqrt{n}}$ finishes the proof. □

To continue with proving more embedding theorems, we will need the following embedding theorem for $L^p$-spaces.

**Theorem 2.3.5.** *Let $1 \leq p < q \leq \infty$, Then $L^q(\Omega) \hookrightarrow L^p(\Omega)$ if the measure of $\Omega$ is finite.*

**Theorem 2.3.6** ($W^{1,p}(\Omega) \hookrightarrow L^q(\Omega)$)**.** *Let $\Omega$ be a bounded open subset of $\mathbb{R}^n$ with $C^1$ boundary, and let $1 \leq p < n$, with the corresponding Sobolev exponent $p*$. Then $W^{1,p}(\Omega) \hookrightarrow L^q(\Omega)$ for $1 \leq q \leq p*$.*

*Proof.* Let $u$ be a function on $W^{1,p}(\Omega)$. By theorem 2.3.1, there exists a map, $E$, such that $Eu = v, v \in W^{1,p}(\mathbb{R}^n)$ and $v|_\Omega = u$. Using this together with theorem 2.3.3, we get that

$$||u||_{L^{p*}(\Omega)} \leq ||v||_{L^{p*}(\mathbb{R}^n)} \leq C_1||v||_{W^{1,p}(\mathbb{R}^n)} \leq C_2||u||_{W^{1,p}(\Omega)}.$$

Since $\Omega$ was assumed to be bounded, we can use the embedding theorem for $L^p$-spaces of theorem 2.3.5, thus

$$L^{p*}(\Omega) \hookrightarrow L^q(\Omega), 1 \leq q \leq p* .$$

Putting these two statements together finishes the proof. □

The above embedding theorem can in fact be generalized. It can be shown that the embedding is in fact a compact embedding. We use the following definitions to specify what that means and we then explicitly state the compact embedding property as a theorem.

**Definition 2.3.2** (Compact embedding)**.** Let $X$ and $Y$ be two complete normed vector spaces, where $X$ is embedded in $Y$. We say that $X$ is compactly embedded in $Y$, which we denote by $\subset\subset$, if the inclusion operator $i : X \to Y$ is a compact operator. That means that that the images of bounded sets in $X$ are precompact sets in $Y$. Which in turn means that the closure of the images of bounded sets in $X$ are compact sets in $Y$.

**Theorem 2.3.7** (Rellich-Kondrachov)**.** *Let $\Omega$ be a bounded open subset of $\mathbb{R}^n$ with $C^1$-boundary $\partial\Omega$. Now let $1 \leq p < n$ and denote by $p*$ the Sobolev exponent of $p$. Then $W^{1,p}(\Omega) \subset\subset L^q(\Omega)$ for $1 \leq q < p*$.*

*Proof.* The proof can be found in [2]. □

**Case 2** ($n = p$)**.**

We begin with the case when $\Omega = \mathbb{R}^n$ Before we begin with our proof, we will show a small lemma for positive real numbers.

**Lemma 2.3.8** ($n \cdot a \cdot b^{n-1} \leq (n+1) \cdot b^n + a^n$)**.** *Let $a$ and $b$ be positive real numbers and let $n$ be a positive integer, then*

$$n \cdot a \cdot b^{n-1} \leq (n+1) \cdot b^n + a^n.$$

*Proof.* The proof will be carried out in two steps, when $b \geq a$ and when $a > b$. We begin with the case where $b \geq a$.

If $b \geq a$, then $(n+1) \cdot b \cdot b^{n-1} + a^n \geq (n+1) \cdot b \cdot b^{n-1} \geq n \cdot a \cdot b^{n-1}$. Which shows that the inequality holds in the first case.

When $a > b$, we will prove the inequality by studying the derivatives. We have already shown that the inequality holds for $a = b$. It follows that if $(n+1) \cdot b^n + a^n$ grows faster than $n \cdot a \cdot b^{n-1}$ as a function of $a$ for all $a > b$, the theorem will hold. The derivative of the first expression with respect to $a$ is $n \cdot a^{n-1}$ and the derivative of the second expression with respect to $a$ is $n \cdot b^{n-1}$. Since we assume that $a > b$, the inequality for the derivatives holds, and thus the theorem follows. $\square$

**Theorem 2.3.9.** *$W^{1,p}(\mathbb{R}^n) \hookrightarrow L^q(\mathbb{R}^n)$ for any $p \leq q < \infty$, which means that $n \leq q < \infty$ since we assume that $p = n$.*

*Proof.* As before, we use the standard density argument to justify that we only consider $u \in C_0^\infty(\mathbb{R}^n)$. We use inequality (2) from the proof of part ii) of theorem 2.3.3. This gives us

$$||u||^\gamma_{L^{\frac{n\gamma}{n-1}}(\mathbb{R}^n)} \leq \gamma ||\nabla u||^{\gamma-1}_{L^{\frac{n(\gamma-1)}{n-1}}(\mathbb{R}^n)}.$$

If we now choose gamma to be equal to $n$ and use lemma 2.3.8 with $a = ||\nabla u||_{L^n(\mathbb{R}^n)}$ and $b = ||u||_{L^n(\mathbb{R}^n)}$, the above inequality becomes

$$||u||^n_{L^{\frac{n^2}{n-1}}(\mathbb{R}^n)} \leq n||\nabla u||_{L^n(\mathbb{R}^n)} ||u||^{n-1}_{L^{\frac{n(n-1)}{n-1}}(\mathbb{R}^n)}.$$
$$\leq ||\nabla u||^n_{L^n(\mathbb{R}^n)} + (n+1)||u||^n_{L^n(\mathbb{R}^n)} \leq C||u||^n_{W^{1,n}(\mathbb{R}^n)}$$

If we now choose $\gamma$ to be $n+1$, then in the same manner as above, we get that

$$||u||^{n+1}_{L^{\frac{n(n+1)}{n-1}}(\mathbb{R}^n)} \leq ||\nabla u||^{n+1}_{L^n(\mathbb{R}^n)} + n||u||^{n+1}_{L^{\frac{n^2}{n-1}}(\mathbb{R}^n)}.$$

We now notice that the first term on the right hand side is smaller than $||u||^{n+1}_{W^{1,n}(\mathbb{R}^n)}$ And by the inequality attained from the case when $\gamma = n$, we see that the second term is bounded by $C||u||^{n+1}_{W^{1,n}(\mathbb{R}^n)}$ for some constant $C$. Adding this together we see that

$$||u||^{n+1}_{L^{\frac{n(n+1)}{n-1}}(\mathbb{R}^n} \leq ||u||^{n+1}_{W^{1,n+1}(\mathbb{R}^n)}.$$

By repeating this process for $\gamma = n+2, n+3, n+4, ..., n+k, n+k+1$ and by using that the inequality holds for the previous case, we can create a sequence of numbers $p_k = \dfrac{(n+k)n}{n-1}$ which tends to infinity and for which

$$||u||_{L^{p_k}(\mathbb{R}^n)} \leq ||u||_{W^{1,n}(\mathbb{R}^n)}.$$

Now for every $q \geq n$ where $\frac{1}{q} = \frac{\alpha}{n} + \frac{1-\alpha}{p_k}$ for $\alpha = \alpha_k \in [0,1]$ for some $p_k \geq q$, we can use the log convexity of $L^p$-norms to attain

$$||u||_{L^q(\mathbb{R}^n)} \leq ||u||^{\alpha}_{L^n(\mathbb{R}^n)} ||u||^{1-\alpha}_{L^{p_k}(\mathbb{R}^n)}$$
$$\leq ||u||^{\alpha}_{L^n(\mathbb{R}^n)} C ||u||^{1-\alpha}_{W^{1,n}(\mathbb{R}^n)} \leq C ||u||^{\alpha+1-\alpha}_{W^{1,n}(\mathbb{R}^n)} = C ||u||_{W^{1,n}(\mathbb{R}^n)}.$$

This finishes the proof. □

**Theorem 2.3.10.** *Let $\Omega$ be a bounded open subset of $\mathbb{R}^n$ with $C^1$-boundary. Then $W^{1,p}(\Omega) \subset\subset L^q(\Omega)$ for all $1 \leq q < \infty$.*

*Proof.* Since closed and bounded subsets of $\mathbb{R}^n$ are compact in $\mathbb{R}^n$, we have that $\Omega$ is precompact in $\mathbb{R}^n$. By the standard embedding of $L^p$-spaces (theorem 2.3.5) $W^{1,n}(\Omega) \hookrightarrow W^{1,p}(\Omega)$, for every $p < n$. Thus the statement of the theorem is a consequence of theorem 2.3.6. □

**Case 3** $(n < p < \infty)$**.**

We will now prove some embedding theorems for Sobolev spaces into Hölder spaces. In order to do so, we need the following inequality.

**Theorem 2.3.11** (Morrey's Inequality)**.** *Let $n$ and $p$ be two constants such that $\frac{n}{p} < 1$, i.e. $n < p$. Then there is a constant $C_{n,p}$ depending only on $n$ and $p$ such that*

$$\textit{i)}\ |u(y) - u(x)| \leq C_{n,p} |y-x|^{1-\frac{n}{p}} ||\nabla u||_{L^p(B_{2|y-x|(x)})}$$
$$\leq C_{n,p} |y-x|^{1-\frac{n}{p}} ||u||_{W^{1,p}(B_{2|y-x|(x)})}$$
$$\textit{ii)}\ |u(x)| \leq C_{n,p}[||\nabla u||_{L^p(B_2(x))} + ||u||_{L^p(B_1(x))}] \leq C_{n,p} ||u||_{W^{1,p}(B_2(x))}$$

*for every $u \in W^{1,p}(\mathbb{R}^n)$. In this theorem, $B_r(x)$ is the open ball of radius $r$ centered at the point $x$. Note that this implies that*

$$||u||_{C^{0,1-n/p}(\mathbb{R}^n)} \leq C_{n,p} ||u||_{W^{1,p}(\mathbb{R}^n)}.$$

*by dividing both sides of inequality i) by $|y-x|^{1-\frac{n}{p}}$, adding this to inequality ii), using that for every $u \in W^{1,p}(\mathbb{R}^n)$, $||u||_{W^{1,p}(\Omega)} \leq ||u||_{W^{1,p}(\mathbb{R}^n)}$, where $\Omega$ is a bounded open subset of $\mathbb{R}^n$ and finally by recalling the definition of the Hölder norm.*

*Proof.* To make the proof easier to follow, it will be divided up into three steps. The first one is the following:

Let $r > 0$, then

$$\int_{B_r(x)} |u(y) - u(x)| d^n y \leq \frac{r^n}{n} \int_{B_r(x)} \frac{|\nabla u(y)|}{|y-x|^{n-1}} d^n y.$$

To show this, we examine the first integral by switching to spherical coordinates centered at $x$. We write $y = x + t\omega$ with $0 \leq t \leq r$ and $\omega$ running over the unit sphere $S^{n-1}$. We denote the usual surface area measure on $S^{n-1}$ by $d\sigma(\omega)$. Using the fundamental theorem of calculus, we see that:

$$|u(x+t\omega) - u(x)| \leq \left| \int_0^t \frac{d}{ds} u(x+s\omega) ds \right| \leq \int_0^t |\nabla u(x+s\omega)| ds.$$

By using this on the first integral, we have

$$\begin{aligned}
\int_{B_r(x)} |u(y) - u(x)| d^n y &= \int_0^r dt \int_{S^{n-1}} d\sigma(\omega) t^{n-1} |u(x+t\omega) - u(x)| \\
&\leq \int_0^r \int_{S^{n-1}} d\sigma(\omega) t^{n-1} \int_0^t ds |\nabla u(x+s\omega)| \\
&= \int_{S^{n-1}} d\sigma(\omega) \int_0^r ds \int_s^r dt\ t^{n-1} |\nabla u(x+s\omega)| \\
&= \int_{S^{n-1}} d\sigma(\omega) \int_0^r ds \frac{r^n - s^n}{n} |\nabla u(x+s\omega)| \\
\leq \frac{r^n}{n} \int_{S^{n-1}} d\sigma(\omega) \int_0^r ds |\nabla u(x+s\omega)| &= \frac{r^n}{n} \int_{B_r(x)} \frac{|\nabla u(y)|}{|y-x|^{n-1}} d^n y.
\end{aligned}$$

Where the last step can be seen to hold by going back to an integral over $B_r(x)$ where the function $|\nabla u(x+s\omega)|$ becomes $|\nabla u(y)|$ and the denominator $|y-x|^{n-1}$ is the Jacobian determinant. This proves the first step.

We will now prove the bound on $|u(y) - u(x)|$ where $x, y \in \mathbb{R}^n$.

Set $r = |y - x|$. By the triangle inequality, we have that:

$$|u(y) - u(x)| \leq |u(y) - u(w)| + |u(w) - u(x)|,$$

for every $w$.

We will now take the average over $w$ in $B_r(x) \cap B_r(y)$, which, has volume $C_n r^n$ for some constant $C_n$ which depends only on the dimension $n$. We thus get

$$
\begin{aligned}
&|u(y)-u(x)| \\
&\leq \frac{1}{C_n r^n}\int_{B_r(x)\cap B_r(y)} |u(y)-u(w)| d^n w + \frac{1}{C_n r^n}\int_{B_r(x)\cap B_r(y)} |u(w)-u(x)| d^n w \\
&\quad \leq \frac{1}{C_n r^n}\int_{B_r(y)} |u(y)-u(w)| d^n w + \frac{1}{C_n r^n}\int_{B_r(x)} |u(w)-u(x)| d^n w \\
&\qquad \leq \frac{1}{C_n n}\int_{B_r(y)} \frac{|\nabla u(w)|}{|y-w|^{n-1}} d^n w + \frac{1}{C_n n}\int_{B_r(x)} \frac{|\nabla u(w)|}{|x-w|^{n-1}} d^n w
\end{aligned}
$$

Where the second step is true since $B_r(x) \cap B_r(y) \subset B_r(x)$ and $B_r(x) \cap B_r(y) \subset B_r(y)$.

By using Hölder's inequality, we get that

$$
\int_{B_r(x)} \frac{|\nabla u(w)|}{|w-x|^{n-1}} d^n w \leq \left[\int_{B_r(x)} |\nabla u(w)|^p d^n w\right]^{1/p} \left[\int_{B_r(x)} \frac{1}{|w-x|^{\frac{(n-1)p}{(p-1}}} d^n w\right]^{\frac{p-1}{p}}.
$$

We see that the second integral in this expression converges if

$$
\frac{(n-1)p}{p-1} < n \iff \frac{n-1}{n} < \frac{p-1}{p} = 1 - \frac{1}{p} \iff n < p.
$$

Since this is true by assumption, the second integral converges. By switching to spherical coordinates, we now have that

$$
\int_{B_r(x)} \frac{1}{|w-x|^{\frac{(n-1)p}{p-1}}} d^n w = \int_0^r dt \int_{S^{n-1}} d\sigma(\omega) \frac{t^{n-1}}{t^{\frac{(n-1)p}{p-1}}} = \Omega_{n-1} \frac{p-1}{p-n} r^{\frac{p-n}{p-1}}
$$

where $\Omega_{n-1}$ is the surface area of $S^{n-1}$.

Using this inequality in the expression acquired by using Hölder's inequality, and putting that inequality into the already derived inequality for $|u(y) - u(x)|$, and finally using that $B_r(y) \subset B_{2r}(x)$ – thus justifying putting everything under the same integral – we get

$$
|u(y)-u(x)| \leq C_{n,p} r^{1-\frac{n}{p}} \left[\int_{B_{2r}(x)} |\nabla u(w)|^p d^n w\right]^{\frac{1}{p}}.
$$

Which is the desired bound for $|u(y) - u(x)|$.

Note that this means that $||u||_{C^{0,k}(\mathbb{R}^n)} \leq C||u||_{W^{1,p}(\mathbb{R}^n)}$ where $k = 1 - n/p$.

We will finally prove the bound for $|u(x)|$.

Using the triangle inequality, we see that $|u(x)| \leq |u(y) - u(x)| + |u(y)|$. Averaging this expression over $y \in B_1(x)$ and denoting the volume of the unit sphere by $V_1$, we see that

$$\begin{aligned}|u(x)| &\leq \frac{1}{V_1}\int_{B_1(x)} |u(y) - u(x)| d^n y + \frac{1}{V_1}\int_{B_1(x)} |u(y)| d^n y \\ &\leq C_{n,p} ||\nabla u||_{L^p(B_2(x))} + \frac{1}{V_1^{1/p}} ||u||_{L^p(B_1(x))}\end{aligned}$$

by using the previously proven inequality for $|u(y) - u(x)|$ and the fact that

$$\begin{aligned}\frac{1}{V_1}\int_{B_1(x)} |u(y)| d^n y &= \int_{B_1(x)} \left(\left(\frac{1}{V_1}|u(y)|\right)^p\right)^{1/p} d^n y \\ \leq \left[\frac{1}{V_1}\int_{B_1(x)} |u(y)|^p d^n y\right]^{1/p} &= \frac{1}{V_1^{1/p}} ||u||_{L^p(B_1(x))}.\end{aligned}$$

This finishes the proof. □

Intuitively, Morrey's inequality means that existence and certain regularity of the weak derivatives implies Hölder continuity for the function, after possibly having been redefined at some set of measure zero.

**Definition 2.3.3.** A function $\tilde{u}$ is called a version of $u$ if $\tilde{u} = u$ a.e.

**Theorem 2.3.12** ($W^{1,p}(\Omega) \hookrightarrow C^{0,\alpha}(\bar{\Omega})$)**.** *Let $\Omega$ be a bounded open subset of $\mathbb{R}^n$ with $C^1$-boundary and let $n < p < \infty$. Then $W^{1,p}(\Omega) \hookrightarrow C^{0,\alpha}(\bar{\Omega})$ with $\alpha = 1 - \dfrac{n}{p}$. This means that $u$ has a version $\tilde{u} \in C^{0,\alpha}(\bar{\Omega})$ and the estimate*

$$||\tilde{u}||_{C^{0,\alpha}(\bar{\Omega})} \leq C ||u||_{W^{1,p}(\Omega)}$$

*holds.*

*Note that the above inequality is very similar to Morrey's inequality. Because of this, the proof idea will be to extend $u$ from a function on $\Omega$ to a function on $\mathbb{R}^n$ using the extension of theorem 2.3.1 and then use Morrey's inequality.*

*Proof.* Let $u \in W^{1,p}(\Omega)$. Note that $\Omega$ satisfies the conditions for the extension of theorem 2.3.1 by assumption. We can thus extend $u$ to $v$ by $v = Eu \in W^{1,p}(\mathbb{R}^n)$. Since $C_0^\infty(\mathbb{R}^n)$ is dense in $W^{1,p}(\mathbb{R}^n)$, there exists a sequence $\{u_i\}_{i \in \mathbb{N}} \in C_0^\infty(\mathbb{R}^n)$ for which

$$\lim_{i \to \infty} ||u_i - v||_{W^{1,p}(\mathbb{R}^n)} = 0$$

By Morrey's inequality, we have that $||u_i - u_j||_{C^{0,\alpha}(\mathbb{R}^n)} \leq C ||u_i - u_j||_{W^{1,p}(\mathbb{R}^n)}$ for all $i, j \geq 1$. By letting $j \to \infty$, we see that there exists a function $\tilde{u} \in$

$C^{0,\alpha}(\mathbb{R}^n)$ for which $\lim_{i\to\infty}||u_i - \tilde{u}||_{C^{0,\alpha}(\mathbb{R}^n)} = 0$ since $C^{0,\alpha}(\mathbb{R}^n)$ is complete, and therefore the limit of functions in the space still lie in the space.

Since $\lim_{i\to\infty}||u_i-\tilde{u}||_{C^{0,\alpha}(\mathbb{R}^n)} = 0$, $\lim_{i\to\infty}||u_i-Eu||_{W^{1,p}(\mathbb{R}^n)} = 0$ and $Eu = u$ a.e. in $\Omega$, we see that $\tilde{u} = u$ a.e. in $\Omega$, so $\tilde{u}$ is a version of $u$ which satisfies

$$||\tilde{u}||_{C^{0,\alpha}(\mathbb{R}^n)} \leq C||v||_{W^{1,p}(\mathbb{R}^n)}.$$

By applying part ii) of theorem 2.3.1 which states that $E$ is continuous, the above inequality proves the embedding. □

By using some properties of Hölder spaces together with some results about compact operators, the above theorem can be generalized.

**Theorem 2.3.13.** *Let $\Omega$ be a bounded open subset of $\mathbb{R}^n$ with $C^1$-boundary and let $n < p < \infty$. Then $W^{1,p}(\Omega) \hookrightarrow C^{0,\alpha}(\bar{\Omega})$ for all $0 < \alpha < \alpha* = 1-n/p$.*

*Proof.* The proof can be found in [6]. □

We will now collect all Sobolev embeddings from previous theorems into one theorem, thus formulating the more general Sobolev embedding theorem.

**Theorem 2.3.14** (Sobolev embeddings)**.**

Let $\Omega$ be a bounded open subset of $\mathbb{R}^n$ with $C^1$-boundary, and $1 \leq p < \infty$. Then we have the following embeddings

1. $kp < n : W^{k,p}(\Omega) \hookrightarrow L^q(\Omega)$ for all $1 \leq q \leq p_k$, $\dfrac{1}{p_k} = \dfrac{1}{p} - \dfrac{k}{n}$. Also, the embedding is compact for all $q < p*$.
2. $kp = n : W^{k,p}(\Omega) \hookrightarrow L^q(\Omega)$ for all $1 \leq q < \infty$. And the embedding is compact for all $q$.
3. $kp > n : W^{k,p}(\Omega) \hookrightarrow C^{l,\alpha}(\bar{\Omega})$ for $l = [k - n/p] \in \mathbb{N}_0$ and $0 < \alpha \leq \alpha_0 = k - l - n/p$. The embedding is compact for $\alpha < \alpha_0$.

These statements follow from the previous embedding theorems by inductively bounding the norm of $W^{k,p}(\Omega)$ by the norm of $W^{k-1,p}(\Omega)$ using the previous embedding theorems until finally $k = 1$, after which we are done.

The previous theorem about embedding theorems can be strengthened. For example, it can be shown that Sobolev spaces are not only compactly embedded into $L^p$-spaces, but that in fact if $mp < n, n - mp < k < n$ where $m$ is a non-negative integers, $k$ is a non-negative integer satisfying that $1 \leq k \leq n$ and $p$ satisfies that $1 \leq p < \infty$ then

$$W^{j+m,p}(\Omega) \to W^{j,q}(\Omega^k),$$

where $j$ is a non-negative integer.

As a special case, when $k = 1$, we have

$$W^{j+m,p}(\Omega) \to W^{j,q}(\Omega).$$

In particular, when $j = 0$

$$W^{m,p}(\Omega) \to L^q(\Omega).$$

Our $C^1$-condition on the boundary is also more restrictive than what is necessary. In fact, it suffices that $\Omega$ satisfies the cone condition, which means that there exists some finite cone $K$ such that each $x \in \Omega$ is a vertex of a finite cone $K_x$ contained in $\Omega$ and congruent to $K$.

The full proofs of these theorems can be found in [1].

However, the proofs of the more general Sobolev embedding theorems are much longer, but they add quite little to the interesting proof ideas and concepts used for proving the embedding theorems.

# Chapter 3

# Elliptic boundary value problems

In this chapter, we will use the methods and theorems of previous sections to investigate the solvability of elliptic partial differential equations.

We will mainly be concerned with studying boundary value problems of the form

$$\begin{cases} Lu = f \text{ in } \Omega \\ u = 0 \text{ on } \partial\Omega \end{cases}$$

Where $\Omega$ is a bounded open subset of $\mathbb{R}^n$, $u : \overline{\Omega} \to R$ is the unknown function and where $f : \Omega \to \mathbb{R}$ is given and $L$ denotes a second-order partial differential operator.

The justification of this limitation will come right after we have made some fundamental definitions.

We will study second-order partial differential operators which will be on one of the following forms

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

or

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

If we let $a(x)$ be a $n \times n$-matrix valued function of $x$ and we let $b(x)$ be an $n$-dimensional column vector valued function of $x$, the first expression can be written as

$$Lu = \nabla \cdot (a\nabla u) + b^T \nabla u + cu$$

Because of this, we say that the partial differential equation $Lu = f$ is in divergence form if $Lu$ is on the first form. Consequently we say that it is in non-divergence form it is written in the other form.

If $a^{i,j}(x)(i,j = 1,...,n)$ are $C^1$-functions, then an operator in divergence form can also be written in non-divergence form and vice versa. From here on, we will also assume that the $n \times n$-matrix $a(x)$ is symmetric, i.e. $a^{i,j} = a^{j,i}(i,j = 1,...,n)$.

We will now define what we mean by an elliptic partial differential operator

**Definition 3.0.4.** The partial differential operator $L$ is said to be (uniformly) elliptic if there exists a constant, $K > 0$, such that

$$\sum_{i,j=1}^{n} a^{i,j}\eta_i\eta_j \geq K|\eta|^2$$

for a.e. $x \in \Omega$ and for every $\eta \in \mathbb{R}$.

Ellipticity thus means that for each fixed point $x$, the symmetric $n \times n$-matrix $a(x)$ is positive definite with smallest eigenvalue greater than or equal to $K$.

We will look at the weak solutions of our elliptic boundary value problem. To clarify what that means, we must first define what a weak solution means.

Let us consider our first boundary value problem where $L$ is written on divergence form. We will assume that all $a^{i,j}(x), b^i(x)$ and $c(x)$ are bounded. Furthermore we assume that $f \in L^2(\Omega)$.

If we for the moment assume that $u$ is indeed a smooth solution, then we can multiply both sides of the partial differential equation $Lu = f$ by a smooth test function $v \in C_0^\infty(\Omega)$ and then integrate over $\Omega$. By integrating the first term on the left hand side by parts – which is easy since the partial differential operator is in divergence form – and use that there will be no boundary terms since $v = 0$ on $\partial\Omega$, we get

$$\int_\Omega \sum_{i,j=1}^{n} a^{i,j}u_{x_i}v_{x_j} + \sum_{i=1}^{n} b^i u_{x_i} v + cuvdx = \int_\Omega fvdx.$$

Since $C_0^\infty(\Omega)$ is dense in $H_0^1(\Omega)$, we see that the same identity holds if the smooth function $v$ is replaced by any $v \in H_0^1(\Omega)$. Also, the identity still makes sense if only $u \in H_0^1(\Omega)$. We choose $H_0^1(\Omega)$ instead of $H^1(\Omega)$ to guarantee that the boundary condition $u = 0$ on $\partial\Omega$ holds.

**Definition 3.0.5.** i) The bilinear form $B[\cdot,\cdot]$ associated with the divergence form elliptic partial differential operator L is given by

$$B[u,v] := \int_\Omega \sum_{i,j=1}^{n} a^{i,j} u_{x_i} v_{x_j} + \sum_{i=1}^{n} b^i u_{x_i} v + cuv dx$$

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

ii) We say that $u \in H_0^1(\Omega)$ is a solution of our boundary value problem if

$$B[u,v] = \langle f, v\rangle_{L^2(\Omega)}$$

for all $v \in H_0^1(\Omega)$.

With the characterization of $H^{-1}(\Omega)$ from the first section of chapter 2 in mind, we can look at the boundary value problem

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

Where $f^i \in L^2(\Omega)$. We note that the right hand term $f = f^0 - \sum_{i=1}^{n} f^i_{x_i}$ in fact belongs to the dual space $H^{-1}(\Omega)$ of $H_0^1(\Omega)$.

**Definition 3.0.6.** We say that $u \in H_0^1(\Omega)$ is a weak solution of (1) if

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

for all $v \in H_0^1(\Omega)$, where $\langle\cdot,\cdot\rangle$ is the pairing of $H^{-1}(\Omega)$ and $H_0^1(\Omega)$.

By using integration by parts on the right hand side and by using that $v$ vanishes on the boundary, we see that

$$\langle f, v\rangle = \int_\Omega f^0 v + \sum_{i=1}^{n} f^i v_{x_i} dx.$$

We will now justify why we limit ourselves to considering boundary value problems on which the value on the boundary is zero. The justification is due to the possibility to transform boundary value problems with a prescribed non-zero boundary into problems which are zero on the boundary.

Suppose that $\partial\Omega$ is $C^1$ and that $u \in H_0^1(\Omega)$ is a weak solution to the boundary value problem

$$\begin{cases} Lu = f \text{ in } \Omega \\ u = g \text{ on } \partial\Omega \end{cases}.$$

This means that the restriction to the boundary of $u$ – in the sense of section 2 of chapter 2 – is $g$, but also, the bilinear form identity needs to hold for all $v \in H_0^1(\Omega)$. If this should be possible, then $g$ has to be the restriction to the boundary of some $H^1(\Omega)$-function, call it $w$. But then $\tilde{u}$ defined by $\tilde{u} := u - w$ belongs to $H_0^1(\Omega)$ and is a weak solution to the boundary problem

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

where $\tilde{f}$ is defined by $\tilde{f} := f - Lw \in H^{-1}(\Omega)$.

Thus solving boundary problems with non-zero boundary can be reduced to solving boundary problems with zero boundary and finding a $H^1$-function which has $g$ as its restriction to the boundary.

We will now prove a general theorem concerning solvability and uniqueness of solutions to expressions on bilinear form under certain conditions. Our approach will later on be to show that these conditions do indeed hold for our special bilinear form attained from our elliptic partial linear operator.

**Theorem 3.0.15** (Lax-Milgram Theorem)**.** *For this proof, assume that $H$ is a real Hilbert space, with norm denoted by $||\cdot||$ and inner product denoted by $(\cdot,\cdot)$. We continue to denote by $\langle\cdot,\cdot\rangle$ the pairing of $H$ with its dual space.*

*Now assume that $B : H \times H \to \mathbb{R}$ is a bilinear mapping for which there exists constants $a, b > 0$ such that*

$$|B[u,v]| \leq a||u||||v||, u, v \in H$$

*and*

$$b||u||^2 \leq B[u,u], u \in H.$$

*We sometimes say that that a bilinear map $B$ is coercive if the second property holds.*

*If the above conditions hold, there exists a unique element $u \in H$ such that*

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

*for all $v \in H$.*

*Proof.* The proof idea is to use the Riesz representation theorem in two steps. First we have that for each fixed element $u \in H$, the mapping $v \to B[u, v]$ is a bounded linear functional on $H$. We can thus apply the Riesz representation theorem to see that there exists a unique element $w \in H$ for which

$$B[u, v] = (w, v), v \in H$$

It is obvious that $w$ corresponds to $u$ in some way, so let us write $Au = w$, which yields

$$B[u, v] = (Au, v), v \in H.$$

The idea here is to use the Riesz representation theorem again, this time using that

$$\langle f, v \rangle = (w_2, v)$$

for some element $w_2 \in H$. Now if we can show that $A$ is actually a bounded linear operator which is one-to-one and whose range is all of $H$, then we can choose our $u$ such that $Au = w_2$, thus yielding

$$B(u, v) = (Au, v) = (w_2, v) = \langle f, v \rangle.$$

We will therefore prove that $A$ behaves nicely as an operator. We begin by proving that it is a bounded linear operator. If $\lambda_1, \lambda_2 \in \mathbb{R}$ and $u_1, u_2 \in H$, then for each $v \in H$ we have that

$$(A(\lambda_1 u_1 + \lambda_2 u_2), v) = B[\lambda_1 u_1 + \lambda_2 u_2, v]$$

$$= \lambda_1 B[u_1, v] + \lambda_2 B[u_2, v] = \lambda_1 (Au_1, v) + \lambda_2 (Au_2, v) = (\lambda_1 Au_1 + \lambda_2 Au_2, v).$$

This equality holds for each $v \in H$, and so $A$ is linear.

Also, $||Au||^2 = (Au, Au) = B[u, Au] \leq |B[u, Au]| \leq a||u||||Au||$.

Thus $||Au|| \leq a||u||$ and so $A$ is bounded.

Next we will prove that $A$ is one-to-one and that the range of $A$, $R(A)$, is closed in $H$. To prove this, we use

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

Thus $b||u|| \leq ||Au||$ This proves that $A$ is one-to-one since if $Au = Av$, then $b||u - v|| \leq ||Au - Av|| = 0$ and thus $u = v$. This also proves that $R(A)$ is closed since if we have a convergent sequence in $R(A)$, then every element in that sequence can be written as $A(u_j)$ (since the sequence is in the range of $A$), also, this sequence is a Cauchy sequence since it is convergent. By our inequality, $Au_j$ being Cauchy implies that $u_j$ is Cauchy, and since $H$ is complete, $u_j$ has a limit in $H$, call it u. Denote by $y$ the limit of $Au_j$, then $||y - Au|| \leq ||y - Au_j|| + ||Au_j - Au|| \leq ||y - Au_j|| + a||u_j - u||$ since $A$ is bounded. Thus $||y - Au||$ can be made arbitrarily small by choosing $j$ to be sufficiently large. It follows that $Au_j$ will converge to $Au$, which is in $R(A)$, and since $A$ is one-to-one, this limit is unique.

We will now prove that the range of $A$ is all of $H$. This will be shown by contradiction. If the range of $A$ is not all of $H$, then since $R(A)$ is closed, there would exist some non-zero element in $R(A)^{\perp}$, call it $w$. But for this element, $0 = (Aw, w) = B[w, w] \geq b||w||^2$. Since $b$ is assumed to be greater than zero, this implies that $||w|| = 0$ for $w \neq 0$, which is a contradiction. Thus the range of $A$ is all of $H$.

We can now put the pieces together as we previously said we would do. By the Riesz representation theorem there exists a $w_2 \in H$ for which $\langle f, v\rangle = (w_2, v)$ for all $v \in H$. Then by using that $A$ is a bijective map from $H$ onto itself, we can choose $u$ such that $Au = w_2$. For this $u$ we have

$$B[u, v] = (Au, v) = (w, v) = \langle f, v\rangle$$

for all $v \in H$.

It remains to prove uniqueness. Assume that $u$ and $\tilde{u}$ are two elements of $H$ for which $B[u, v] = B[\tilde{u}, v] = \langle f, v\rangle$. Then their difference, $u - \tilde{u}$ satisfies that $B[u - \tilde{u}, v] = 0$. But if we now set $v = u - \tilde{u}$, we get that $b||u - \tilde{u}||^2 \leq B[u - \tilde{u}, u - \tilde{u}] = 0$. Thus $u = \tilde{u}$.

This finishes the proof. $\square$

It is worth noting that if $B$ was symmetrical, i.e. $B[u, v] = B[v, u]$, then $B[u, v]$ would define a new inner product on $H$, on which we could immediately apply the Riesz representation to attain our proof. However, this is not nearly as useful. The major importance of the Lax-Milgram theorem is that it does not require our bilinear form to be symmetric. Since what we will use this theorem for is to prove uniqueness and existence of weak solutions to boundary value problems by showing the the requirements for the theorem hold for our bilinear form adhering from an elliptic partial differential operator. In these cases, we can not in general assume that the bilinear form is symmetric.

We will now return to our specific bilinear form

$$B[u, v] := \int_\Omega \sum_{i,j=1}^{n} a^{i,j} u_{x_i} v_{x_j} + \sum_{i=1}^{n} b^i u_{x_i} v + cuvdx$$

for $u, v \in H_0^1(\Omega)$. As stated above, we will try to verify the conditions for the Lax-Milgram theorem. This can be done if we allow one minor alteration.

But before we continue, we will need the following lemma

**Lemma 3.0.16.** $ab \leq \epsilon a^2 + \dfrac{b^2}{4\epsilon}, \epsilon > 0$

*Proof.* $0 \leq ((2\epsilon)^{1/2} a - \dfrac{b}{(2\epsilon)^{1/2}})^2 = 2\epsilon a^2 - 2ab + \dfrac{b^2}{2\epsilon}$.

The statement follows by adding $2ab$ to both sides and dividing by 2. □

We will also need the following inequality due to Poincaré.

**Theorem 3.0.17** (Poincaré's inequlity)**.** *Let $\Omega$ be a bounded open subset of $\mathbb{R}^n$. Let $u \in W_0^{1,p}(\Omega)$ for some $1 \leq p < n$. Then the following inequality hold*

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

*for each $1 \leq q \leq p^*$, where $p^*$ is the Sobolev exponent associated with $p$ and where the constant $C$ depends only on $p, q, n$ and $\Omega$.*

*In particular, since $p^* > p$, we have that*

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

*Proof.* The Sobolev exponent gives a hint of where we're going. Recall the Sobolev-Gagliardo-Nirenberg inequality which states that

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

where $p^*$ satisfies that $1/p^* = 1/p - 1/n$.

Now, since $u \in W_0^{1,p}(\Omega)$, there exists functions, $u_m \in C_0^\infty(\Omega)$, $(m = 1, 2, ..)$ which converge to $u$ in $W^{1,p}(\Omega)$ (by construction of $W^{1,p}(\Omega)$). We now extend each function $u_m$ to $\mathbb{R}^n$ by defining it to be zero on $\mathbb{R}^n - \overline{\Omega}$. We now apply the Sobolev-Gagliardo-Nirenberg inequality to the extended functions which yields that $||u||_{L^{p^*}(\Omega)} \leq C||\nabla u||_{L^p(\Omega)}$ since $\Omega$ is bounded and we thus have no problems with convergence. Since furthermore, we have that $||u||_{L^q(\Omega)} \leq C||u||_{L^{p^*}(\Omega)}$ if $1 \leq q \leq p^*$, this proves the theorem by putting the two inequalities together. □

**Theorem 3.0.18** (Energy estimates)**.** *There exists constants $a, b > 0$ and $c \geq 0$ such that*

$$|B[u, v]| \leq a||u||_{H_0^1(\Omega)}||v||_{H_0^1(\Omega)}$$

*and*

$$b||u||^2_{H_0^1(\Omega)} \leq B[u, u] + c||u||^2_{L^2(\Omega)}$$

*for all $u, v \in H_0^1\Omega)$.*

*Proof.* We can make a rough estimate to attain

$$|B[u,v]| \leq \sum_{i,j=1}^{n} ||a^{i,j}||_{L^\infty(\Omega)} \int_\Omega |\nabla u||\nabla v| dx + \sum_{i=1}^{n} ||b^i||_{L^\infty(\Omega)} \int_\Omega |\nabla u||v| dx + ||c||_{L^\infty(\Omega)} \int_\Omega |u||v| dx \leq a||u||_{H_0^1(\Omega)} ||v||_{H_0^1(\Omega)}$$

for some choice of big $a$.

Now recall the ellipticity condition that we have on $L$, namely that there exists some constant $K$ such that

$$\sum_{i,j=1}^{n} a^{i,j}(x)\eta_i\eta_j \geq K|\eta|^2$$

for a.e $x \in \Omega$ and for all $\eta \in \mathbb{R}^n$.

By integrating this inequality, we see that

$$K \int_\Omega |\nabla u|^2 dx \leq \int_\Omega \sum_{i,j=1}^{n} a^{i,j} u_{x_1} u_{x_j} dx = B[u,u] - \int_\Omega \sum_{i=1}^{n} b^i u_{x_1} u - cu^2 dx$$
$$\leq B[u,u] + \sum_{i=1}^{n} ||b^i||_{L^\infty(\Omega)} \int_\Omega |\nabla u||u| dx + ||c||_{L^\infty(\Omega)} \int_\Omega u^2 dx$$

We can now use lemma 3.0.16 and integrate to see that

$$\int_\Omega |\nabla u||u| dx \leq \epsilon \int_\Omega |\nabla u|^2 dx + \frac{1}{4\epsilon} \int_\Omega u^2 dx, \quad \epsilon > 0$$

By choosing $\epsilon > 0$ small enough to satisfy

$$\epsilon \sum_{i=1}^{n} ||b^i||_{L^\infty(\Omega)} < \frac{K}{2}$$

and then inserting this expression into the previous inequality yields

$$\frac{K}{2} \int_\Omega |\nabla u|^2 dx \leq B[u,u] + C \int_\Omega u^2 dx$$

where $C$ is chosen to be large enough (a more exact number can be given in terms of $\epsilon$ but it is not of any importance).

By now using Poincaré's inequality for $p = 2$, i.e.

$$||u||_{L^2(\Omega)} \leq C||\nabla u||_{L^2(\Omega)}$$

it follows that

$$b||u||^2_{H^1_0(\Omega)} \leq C_2||\nabla u||^2_{L^2(\Omega)} \leq B[u,u] + c||u||^2_{L^2(\Omega)}$$

for appropriate choices of $b > 0$ and $c \geq 0$.

This finishes the proof. □

If $c = 0$, in the above energy estimates, then $B[\cdot,\cdot]$ satisfies the conditions for the Lax-Milgram theorem. However, if $c > 0$, which it very well might be, then $B[\cdot,\cdot]$ does almost, but not really, satisfy the conditions. The following existence theorem will deal with that possibility. Note that if $c = 0$ in the following theorem, we have existence of a unique solution to the boundary value problem.

**Theorem 3.0.19.** *There is a number $c \geq 0$ such that for each $\mu \geq c$ and each function $f \in L^2(\Omega)$, there exists a unique weak solution, $u \in H^1_0(\Omega)$ of the boundary value problem*

$$\begin{cases} Lu + \mu u = f \ \ in\ \Omega \\ u = 0 \ \ on\ \partial\Omega \end{cases}$$

*Proof.* Take $c$ from the previous theorem and let $\mu \geq c$. Now define a new bilinear form by

$$B_\mu[u,v] = B[u,v] + \mu(u,v), (u, v \in H^1_0(\Omega)$$

This bilinear form corresponds to the operator $L_\mu u = Lu + \mu u$ and satisfies the hypothesis of the Lax-Milgram theorem. Now fix $f \in L^2(\Omega)$ and define $\langle f, v\rangle$ by $\langle f, v\rangle := (f, v)_{L^2(\Omega)}$. This is a bounded linear functional on $L^2(\Omega)$, and is thus a bounded linear functional on $H^1_0(\Omega)$ (since $H^1_0(\Omega) \subset L^2(\Omega)$). We now apply the Lax-Milgram theorem to find a unique function $u \in H^1_0(\Omega)$ which satisfies

$$B_\mu[u,v] = \langle f, v\rangle$$

for all $v \in H^1_0(\Omega)$. This $u$ is consequently the unique weak solution of the boundary value problem above. □

**Example 3.0.1.** *Consider the boundary value problem*

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

*We have that $B[u,v] = \displaystyle\int_\Omega \nabla u \nabla v dx$ by using integration by parts.*

*By Poincaré's inequality we have*

$$||u||^2_{H^1_0(\Omega)} \leq b(||u||^2_{L^2(\Omega)} + ||\nabla u||^2_{L^2(\Omega)})$$
$$\leq C||\nabla u||^2_{L^2(\Omega)} = C\int_\Omega |\nabla u|^2 dx = CB[u,u] \quad (3.1)$$

*By dividing both sides by* $C$*, we see that the previous theorem holds with* $c = 0$*. Thus, there exists a unique weak solution to the equation.*

For further reading on elliptic partial differential equations, see [2].

# Chapter 4

# Bibliography

[1] Robert A. Adams, Sobolev Spaces. Academic Press, 1975.

[2] Lawrence C. Evans, Partial Differential Equations. American Mathematical Society, 2nd edition, 2010.

[3] Joel Feldman, Mikko Salo, Gunther Uhlmann, The Calderón Problem An Introduction to Inverse Problems. 2012.

[4] Gerald B. Folland Real Analysis Modern Techniques and Their Applications. Wiley-Interscience, 2nd edition, 1999.

[5] Avner Friedman, Foundations of Modern Analysis. Dover Publications, Inc. 1982.

[6] Fabian Portmann, Sobolev Spaces. 2011

[7] Elias M. Stein, Rami Shakarchi, Fourier Analysis. Princeton University Press, 2003.

[8] Anders Vretblad Fourier Analysis and Its Applications. Springer 2003.