**Lectures on**

# Sobolev Spaces

**S. Kesavan**
The Institute of Mathematical Sciences,
Chennai.
e-mail: kesh@imsc.res.in

# 1 Distributions

In this section we will, very briefly, recall concepts from the theory of distributions that we will need in the sequel. For details, see Kesavan [1], Chapter 1.

Throughout these lectures, we will be working with an open set $\Omega \subset \mathbb{R}^N$.

Let us briefly motivate our study of distributions and Sobolev spaces. One of the important partial differential equations that we often study is the Laplace's equation:

$$-\Delta u \ = \ f \text{ in } \Omega$$

together with some appropriate boundary condition. It turns out that in elasticity and structural engineering, the importance of the solution $u$ stems from the fact that it minimizes, amongst 'admissible functions' $v$, the energy functional

$$J(v) \ = \ \frac{1}{2}\int_\Omega |\nabla v|^2 \ dx - \int_\Omega fv \ dx.$$

Now it can happen, in many applications, that the function $f$ is not continuous but just, say, in $L^2(\Omega)$. Then, for the second term in the above expression for $J$ to make sense, it follows that $v$ must also belong to $L^2(\Omega)$. The first term in $J$ will make sense if all the first partial derivatives of $u$, *i.e.* $\frac{\partial u}{\partial x_i}, 1 \leq i \leq N$ are all in $L^2(\Omega)$ as well.

But, when we are dealing with functions in $L^2(\Omega)$, what do we mean by its derivatives? This is where we need to generalize the notion of a function and its derivatives and interpret partial differential equations in the new set-up. The framework for this comes from the theory of distributions. Just as we can think of a real number as a linear operator on $\mathbb{R}$ acting by multiplication, we can consider certain functions as linear operators on some special space of functions and then generalize this.

**Definition 1.1** *Let $f : \Omega \to \mathbb{R}$ be a continuous function. Its* **support**, *denoted* $\text{supp}(f)$, *is the closure of the set where $f$ is non-zero. The function is said to be of* **compact support** *in $\Omega$ if the support is a compact set contained inside $\Omega$.* ■

**Definition 1.2** *The space of* **test functions** *in $\Omega$, denoted $\mathcal{D}(\Omega)$, is the space of all $\mathcal{C}^\infty$ functions defined on $\Omega$ which have compact support in $\Omega$.* ■

The space $\mathcal{D}(\Omega)$ is a very rich space and can be made into a locally convex topological vector space.

**Definition 1.3** *The dual of $\mathcal{D}(\Omega)$ is called the* **space of distributions**, *denoted $\mathcal{D}'(\Omega)$, on $\Omega$ and its elements are called* **distributions** *on $\Omega$.* ■

**Example 1.1** A function $u : \Omega \to \mathbb{R}$ is said to be **locally integrable** if $\int_K |u| \ dx < \infty$ for every compact subset $K$ of $\Omega$. A locally integrable function defines a distribution in the following way: if $\varphi \in \mathcal{D}(\Omega)$, then

$$u(\varphi) \ = \ \int_\Omega u\varphi \ dx$$

where we use the same symbol $u$ for the function as well as the distribution it generates, for, if $u$ and $v$ generate the same distribution, then it can be shown that $u = v$ a.e. Now every smooth function and every function in any of the spaces $L^p(\Omega)$, for $1 \leq p \leq \infty$, are all locally integrable and so they all can be considered as distributions. ■

**Example 1.2** Define

$$T(\varphi) \;=\; \varphi(0)$$

for all $\varphi \in \mathcal{D}(\mathbb{R})$. This defines a distribution on $\mathbb{R}$ and it can be shown that it cannot be obtained from any locally integrable function in the sense of Example 1.1 above. This is called the **Dirac distribution** supported at the origin and is denoted usually by the symbol $\delta$.■

**Example 1.3** Let $\mu$ be a measure on $\Omega$ such that $\mu(K) < \infty$ for every compact subset $K$ contained in $\Omega$. Then it defines a distribution given by

$$T_\mu(\varphi) \;=\; \int_\Omega \varphi \; d\mu$$

for every $\varphi \in \mathcal{D}(\Omega)$. The Dirac distribution mentioned above is just the distribution generated by the Dirac measure supported at the origin.

**Example 1.4** Define the distribution $T$ on $\mathbb{R}$ by

$$T(\varphi) \;=\; \varphi'(0)$$

for every $\varphi \in \mathcal{D}(\mathbb{R})$. This distribution does not fall into any of the categories of the preceding examples and is an entirely new object. It is called the **dipole distribution**. ■

Assume that $u : \mathbb{R} \to \mathbb{R}$ is a smooth function. Then, so is its derivative $u'$, and both of them define distributions as in Example 1.1. Now, if $\varphi \in \mathcal{D}(\mathbb{R})$, then, by integration by parts, we have

$$\int_{\mathbb{R}} u'\varphi \; dx \;=\; -\int_{\mathbb{R}} u\varphi' \; dx.$$

Similarly,

$$\int_{\mathbb{R}} u''\varphi \; dx \;=\; -\int_{\mathbb{R}} u'\varphi' \; dx \;=\; \int_{\mathbb{R}} u\varphi'' \; dx$$

and so on. We can use this to define the derivative of any distribution $T$ on $\mathbb{R}$ as follows: if $T$ is a distribution on $\mathbb{R}$, then, for any positive integer $k$, define the distribution $\frac{d^kT}{dx^k}$ by

$$\frac{d^kT}{dx^k}(\varphi) \;=\; (-1)^k T(\varphi^{(k)})$$

for any $\varphi \in \mathcal{D}(\mathbb{R})$, where $\varphi^{(k)} = \frac{d^k\varphi}{dx^k}$. We can do this on any open subset $\Omega \subset \mathbb{R}^N$, for any space dimension $N$. To describe it we first establish some

useful notation.

**Notation**
A multi-index $\alpha$ is an $N$-tuple of non-negative integers. Thus,

$$\alpha = (\alpha_1, \cdots, \alpha_N)$$

where the $\alpha_i$ are all non-negative integers. We define

$$\begin{aligned} |\alpha| &= \sum_{i=1}^{N} \alpha_i; \\ x^\alpha &= x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_N^{\alpha_N} \end{aligned}$$

for $x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^N$;

$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_N^{\alpha_N}}.$$

**Definition 1.4** *If $T \in \mathcal{D}'(\Omega)$ is a distribution on an open set $\Omega \subset \mathbb{R}^N$, and if $\alpha$ is any multi-index, we can define the distribution $D^\alpha T$ by*

$$D^\alpha T(\varphi) = (-1)^{|\alpha|} T(D^\alpha \varphi)$$

*for all $\varphi \in \mathcal{D}(\Omega)$.* ■

Thus, by this trick of transfering the burden of differentiation onto elements of $\mathcal{D}(\Omega)$, every distribution becomes infinitely differentiable.

**Example 1.5** The dipole distribution (cf. Example 1.4) is nothing but $-\frac{d\delta}{dx}$, where $\delta$ is the Dirac distribution (cf. Example 1.2). ■

**Example 1.6** Consider the Heaviside function $H : \mathbb{R} \to \mathbb{R}$, defined by

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

This is locally integrable and so defines a distribution. If $\varphi \in \mathcal{D}(\mathbb{R})$, then

$$\frac{dH}{dx}(\varphi) = -\int_{-\infty}^{\infty} H(x)\varphi'(x)\, dx = -\int_0^{\infty} \varphi'(x)\, dx = \varphi(0) = \delta(\varphi).$$

Thus, eventhough $H$ is differentiable *almost everywhere* in the classical sense with the classical derivative being equal to zero a.e. (which is trivially a locally integrable function), the distribution derivative is not zero, but the Dirac distribution. ■

This leads us to the following question: if a locally integrable function $u$ admits a classical derivative (a.e.), denoted $u'$, which is also locally integrable, when can we say that the classical and distributional derivatives of $u$ are the

same? The answer lies in the integration by parts formula. If, for every $\varphi \in \mathcal{D}(\Omega)$ we have

$$\int u'\varphi \, dx \;=\; -\int u\varphi' \, dx,$$

then the classical derivative is also the distributional derivative. This obviously happens when $u$ is a smooth function. It is also true for absolutely continuous functions, cf. Kesavan [1].

In this context, we can also ask the following question. We know that if $u$ is a differentiable function on $\mathbb{R}$ such that $u' \equiv 0$, then $u$ is a constant function. Does the same hold for the distribution derivative? That is, if $T$ is a distribution so that $\frac{dT}{dx}$ is the zero distribution, then, is $T$ the distribution generated by a constant function? In other words, does there exist a constant $c \in \mathbb{R}$ such that, for every $\varphi \in \mathcal{D}(\mathbb{R})$, we have

$$T(\varphi) \;=\; c\int_{\mathbb{R}} \varphi \, dx?$$

We claim that this is indeed the case. If $\frac{dT}{dx} = 0$, then for every $\varphi \in \mathcal{D}(\mathbb{R})$, we have $T(\varphi') = 0$. Thus we need to know when an arbitrary member of $\mathcal{D}(\mathbb{R})$ can be expressed an the derivative of another member of the same space.

**Lemma 1.1** *Let $\varphi \in \mathcal{D}(\mathbb{R})$. Then, there exists $\psi \in \mathcal{D}(\mathbb{R})$ such that $\psi' = \varphi$, if, and only if, $\int_{\mathbb{R}} \varphi \, dx = 0$.*

**Proof:** If $\varphi = \psi'$ with $\psi$ of compact support, then clearly $\int_{\mathbb{R}} \varphi \, dx = 0$. Conversely, let $\text{supp}(\varphi) \subset [-a, a] \subset \mathbb{R}$. Define

$$\psi(x) \;=\; \begin{cases} 0, & \text{if} x \leq -a, \\ \int_{-a}^{x} \varphi(t) \, dt, & \text{if} x > -a. \end{cases}$$

Then, clearly $\psi$ is infinitely differentiable and, by the given condition, has support in $[-a, a]$ as well. Its derivative is obviously $\varphi$. ■

Now, let $\varphi_0 \in \mathcal{D}(\mathbb{R})$ be chosen such that $\int_{\mathbb{R}} \varphi_0(t) \, dt = 1$. Then, given any $\varphi \in \mathcal{D}(\mathbb{R})$, define

$$\varphi_1 \;=\; \varphi - \left(\int_{\mathbb{R}} \varphi \, dx\right) \varphi_0 \;\in\; \mathcal{D}(\mathbb{R}).$$

Then the integral of $\varphi_1$ over $\mathbb{R}$ vanishes. So $\varphi_1 = \psi'$ for some $\psi \in \mathcal{D}(\mathbb{R})$. If $\frac{dT}{dx} = 0$, then, by definition, we have that $T(\psi') = 0$ and so we get

$$T(\varphi) \;=\; \left(\int_{\mathbb{R}} \varphi \, dx\right) T(\varphi_0)$$

which establishes our claim with $c = T(\varphi_0)$. Note that whatever the function $\varphi_0$ we may choose in $\mathcal{D}(\mathbb{R})$ whose integral is unity, the value of $T(\varphi_0)$ is the same (why?) so that $c$ is well-defined.

We conclude this section by mentioning two very useful collections of smooth functions.

Consider the following function defined on $\mathbb{R}^N$.

$$\rho(x) \;=\; \begin{cases} e^{-\frac{1}{1-|x|^2}}, & \text{if} |x| < 1, \\ 0, & \text{if} |x| \geq 1, \end{cases}$$

where $|x|$ is the euclidean length of the vector $x \in \mathbb{R}^N$. It can be shown that this is a function in $\mathcal{D}(\mathbb{R}^N)$, with support in $B(0;1)$, the ball centred at the origin and of unit radius. Let

$$k \;=\; \int_{\mathbb{R}^N} \rho(x)\; dx.$$

**Definition 1.5** *The family of* **mollifiers** $\{\rho_\varepsilon\}_{\varepsilon>0}$ *is the collection of functions defined by*

$$\rho_\varepsilon(x) \;=\; k^{-1}\rho(x/\varepsilon). \;\blacksquare$$

Then it is easy to see that $\rho_\varepsilon \in \mathcal{D}(\mathbb{R}^N)$ with support $B(0;\varepsilon)$, the ball centered at the origin and of radius $\varepsilon$ and is such that

$$\int_{\mathbb{R}^N} \rho_\varepsilon(x)\; dx \;=\; \int_{B(0;\varepsilon)} \rho_\varepsilon(x)\; dx \;=\; 1$$

for all $\varepsilon > 0$.

**Definition 1.6** *Let* $\{U_i\}_{i=1}^m$ *be open sets whose union is* $U$*. Then a* $\mathcal{C}^\infty$**-partition of unity**, *subordinate to the collection* $\{U_i\}_{i=1}^m$ *is a collection* $\{\psi_i\}_{i=1}^m$ *of* $\mathcal{C}^\infty$ *functions defined on* $U$ *such that*
*(i)* $\text{supp}(\psi_i) \subset U_i$*, for all* $1 \leq i \leq m$*;*
*(ii)* $0 \leq \psi_i(x) \leq 1$*, for all* $x \in U$ *and for all* $1 \leq i \leq m$*;*
*(iii)* $\sum_{i=1}^m \psi_i(x) = 1$*, for all* $x \in U$*.* $\blacksquare$

# 2 Sobolev Spaces

Let $\Omega \subset \mathbb{R}^N$ be an open set and let $\partial\Omega$ denote its boundary.

**Definition 2.1** *Let* $m$ *be a positive integer and let* $1 \leq p \leq \infty$*. The* **Sobolev Space** $W^{m,p}(\Omega)$ *is defined by*

$$W^{m,p}(\Omega) \;=\; \{u \in L^p(\Omega) \mid D^\alpha u \in L^p(\Omega),\ \textit{for all } |\alpha| \leq m\}. \;\blacksquare$$

The space $W^{m,p}(\Omega)$ is a vector space contained in $L^p(\Omega)$ and we endow it with the norm $\|.\|_{m,p,\Omega}$ defined as follows.

$$\|u\|_{m,p,\Omega} \;=\; \left(\sum_{|\alpha|\leq m} \|D^\alpha u\|^p_{L^p(\Omega)}\right)^{\frac{1}{p}}$$

if $1 \leq p < \infty$ and

$$\|u\|_{m,\infty,\Omega} \;=\; \max_{|\alpha|\leq m} \|D^\alpha u\|_{L^\infty(\Omega)}.$$

**Notations and conventions**

- When $p = 2$, we will write $H^m(\Omega)$ instead of $W^{m,2}(\Omega)$. The corresponding norm $\|.\|_{m,2,\Omega}$ will be written as $\|.\|_{m.,\Omega}$ and it is generated by the inner-product

$$(u, v)_{m,\Omega} \;=\; \sum_{|\alpha|\leq m} \int_\Omega D^\alpha u D^\alpha v \; dx.$$

- We define a *semi-norm* om $W^{m,p}(\Omega)$ by

$$|u|_{m,p,\Omega} \;=\; \left( \sum_{|\alpha|=m} \|D^\alpha u\|^p_{L^p(\Omega)} \right)^{\frac{1}{p}}$$

  when $1 \leq p < \infty$ and

$$|u|_{m,\infty,\Omega} \;=\; \max_{|\alpha|=m} \|D^\alpha u\|_{L^\infty(\Omega)}.$$

- For consistency, we define $W^{0,p}(\Omega)$ to be $L^p(\Omega)$. Henceforth the norm in $L^p(\Omega)$ will be denoted as $|.|_{0,p,\Omega}$ when $p \neq 2$ and by $|.|_{0,\Omega}$ when $p = 2$.

It follows from the definition of these spaces that the mapping

$$u \mapsto \left( u, \frac{\partial u}{\partial x_1}, \cdots, \frac{\partial u}{\partial x_N} \right)$$

is an isometry from $W^{1,p}(\Omega)$ onto a subspace of $(L^p(\Omega))^{N+1}$.

**Theorem 2.1** *The space $W^{1,p}(\Omega)$ is complete. It is reflexive if $1 < p < \infty$ and separable if $1 \leq p < \infty$. In particular, $H^1(\Omega)$ is a separable Hilbert space.*

**Proof:** Let $\{u_n\}$ be a Cauchy sequence in $W^{1,p}(\Omega)$. Then $\{u_n\}$, $\{\frac{\partial u_n}{\partial x_i}\}, 1 \leq i \leq N$ are all Cauchy sequences in $L^p(\Omega)$. Thus, there exist functions $u$ and $v_i, 1 \leq i \leq N$ in $L^p(\Omega)$ such that $u_n \to u$ and $\frac{\partial u_n}{\partial x_i} \to v_i, 1 \leq i \leq N$ in $L^p(\Omega)$. Now, let $\varphi \in \mathcal{D}(\Omega)$. Then

$$\int_\Omega \frac{\partial u_n}{\partial x_i} \varphi \; dx \;=\; -\int_\Omega u_n \frac{\partial \varphi}{\partial x_i} \; dx$$

by the definition of the distribution derivative. Since $\mathcal{D}(\Omega) \subset L^q(\Omega)$ for any $1 \leq q \leq \infty$, in particular, it is true when $q$ is the conjugate exponent of $p$. Hence, we can pass to the limit in the above relation to get

$$\int_\Omega v_i \varphi \; dx \;=\; -\int_\Omega u \frac{\partial \varphi}{\partial x_i} \; dx$$

for each $1 \leq i \leq N$. This shows that

$$\frac{\partial u}{\partial x_i} \;=\; v_i \in L^p(\Omega), \; 1 \leq i \leq N$$

and so it follows that $u \in W^{1,p}(\Omega)$ and that $u_n \to u$ in $W^{1,p}(\Omega)$. This proves the completeness of the spaces $W^{1,p}(\Omega)$.

The image of $W^{1,p}(\Omega)$ in $(L^p(\Omega))^{N+1}$ via the isometry described above is thus a closed subspace and so it inherits the reflexivity and separability properties of $(L^p(\Omega))^{N+1}$. This completes the proof of the theorem. ■

**Remark 2.1** Let $\{u_n\}$ be a sequence in $W^{1,p}(\Omega),\ 1 < p < \infty$. Let $u_n \to u$ in $L^p(\Omega)$ and let $\{\frac{\partial u_n}{\partial x_i}\}$ be bounded in $L^p(\Omega)$ for all $1 \leq i \leq N$. Then, since $L^p(\Omega)$ is reflexive, it follows that there exists a subsequence $\{u_{n_k}\}$ such that

$$\frac{\partial u_{n_k}}{\partial x_i} \ \rightharpoonup \ v_i$$

*weakly* in $L^p(\Omega)$. Thus if $\varphi \in \mathcal{D}(\Omega)$, then we have

$$\int_\Omega \frac{\partial u_{n_k}}{\partial x_i}\varphi \ dx \ = \ -\int_\Omega u_{n_k}\frac{\partial \varphi}{\partial x_i} \ dx$$

from which we deduce that

$$\int_\Omega v_i\varphi \ dx \ = \ -\int_\Omega u\frac{\partial \varphi}{\partial x_i} \ dx$$

so that it follows that $u \in W^{1,p}(\Omega)$ and that $\frac{\partial u}{\partial x_i} = v_i$. This is a very useful observation. Since $L^1(\Omega)$ is separable, the same idea will also work for the case $p = \infty$. ■

**Remark 2.2** Theorem 2.1 is true for all spaces $W^{m,p}(\Omega)$. ■

**Definition 2.2** *The closure of the subspace $\mathcal{D}(\Omega)$ in $W^{m,p}(\Omega)$ is the closed subspace $W_0^{m,p}(\Omega)$* ■.

We will see later that, in general, $W_0^{m,p}(\Omega)$ is a proper closed subspace of $W^{m,p}(\Omega)$. We will, however, see in the next section that the two spaces coincide when $\Omega = \mathbb{R}^N$.

# 3 The case $\Omega = \mathbb{R}^N$.

Let $p = 2$ and consider the case $H^m(\mathbb{R}^N)$. If $u$ belongs to this space, then it follows that $u, D^\alpha u \in L^2(\mathbb{R}^N)$, for all $|\alpha| \leq m$. Now, for a square integrable function on $\mathbb{R}^N$, we can define its Fourier transform, which will also be square integrable on $\mathbb{R}^N$, and by the Plancherel theorem, the $L^2$-norms of both the function and its Fourier transform will be the same. Further, we also know that

$$\widehat{D^\alpha u}(\xi) \ = \ (2\pi i)^{|\alpha|}\xi^\alpha \widehat{u}(\xi)$$

for $\xi \in \mathbb{R}^N$. Thus it follows that if $u \in H^m(\mathbb{R}^N)$, then $\widehat{u}(.)$ and $\xi \mapsto \xi^\alpha \widehat{u}(\xi),\ |\alpha| \leq m$ are all in $L^2(\mathbb{R}^N)$, and conversely.

We can see easily that the same powers of $\xi$ occur both in $(1+|\xi|^2)^m$ and in the sum $\sum_{|\alpha|\leq m}|\xi^\alpha|^2$ and so we have the existence of two constants $M_1>0$ and $M_2>0$, which depend only on $m$ and $N$, such that

$$M_1(1+|\xi|^2)^m \;\leq\; \sum_{|\alpha|\leq m}|\xi^\alpha|^2 \;\leq\; M_2(1+|\xi|^2)^m.$$

Consequently, we can write

$$H^m(\Omega) \;=\; \{u\in L^2(\mathbb{R}^N)\;|\;(1+|\xi|^2)^{\frac{m}{2}}\widehat{u}(\xi)\in L^2(\mathbb{R}^N)\}.$$

By the Plancherel theorem, it also follows that the norm defined by the following relation is equivalent to the norm in $H^m(\mathbb{R}^N)$ and we will denote it by the same symbol:

$$\|u\|^2_{m,\mathbb{R}^N} \;=\; \int_{\mathbb{R}^N}(1+|\xi|^2)^m|\widehat{u}(\xi)|^2\;d\xi.$$

**Lemma 3.1** *Let $\{\rho_\varepsilon\}_{\varepsilon>0}$ be the family of mollifiers.*
*(i) If $u:\mathbb{R}^N\to\mathbb{R}$ is continuous, then $\rho_\varepsilon * u\to u$ pointwise, as $\varepsilon\to 0$.*
*(ii) If $u:\mathbb{R}^N\to\mathbb{R}$ is continuous with compact support, then $\rho_\varepsilon * u\to u$ uniformly, as $\varepsilon\to 0$.*
*(iii) If $u\in L^p(\mathbb{R}^N)$, then $\rho_\varepsilon * u\to u$ in $L^p(\mathbb{R}^N)$, for $1\leq p<\infty$.*

**Proof:** (i) Let $x\in\mathbb{R}^N$. Then, given $\eta>0$, there exists $\delta>0$ such that for all $|y|<\delta$, we have $|u(x-y)-u(x)|<\eta$. Thus, if $\varepsilon<\delta$, we have

$$|\rho_\varepsilon * u(x)-u(x)|\leq\int_{|y|\leq\varepsilon}|u(x-y)-u(x)|\rho_\varepsilon(y)\;dy<\eta\int_{|y|\leq\varepsilon}\rho_\varepsilon(y)\;dy \;=\; \eta.$$

This proves the first statement.
(ii) If $\mathrm{supp}(u)=K$ which is compact, then

$$\mathrm{supp}(\rho_\varepsilon * u) \;\subset\; K+B(0;\varepsilon)$$

which is compact and is contained within a fixed compact set, say, $K+B(0;1)$ if we restrict $\varepsilon$ to be less than or equal to unity. On this compact set, $u$ is uniformly continuous and the $\delta$ corresponding to $\eta$ in the previous step is now independent of the point $x$ and so the pointwise convergence is now uniform.
(iii) From the step (ii) above it is immediate that if $u$ is continuous with compact support in $\mathbb{R}^N$, then $\rho_\varepsilon * u$ converges to $u$ in $L^p(\mathbb{R}^N)$ as well, since the entire family is supported in a single compact set and the convergence there is uniform. Now, we know that continuous functions with compact support are dense in $L^p(\mathbb{R}^N)$ for $1\leq p<\infty$ and so given $u\in L^p(\mathbb{R}^N)$, we can find, for every $\eta>0$, a continuous function $g$ with compact support such that

$$|u-g|_{0,p,\mathbb{R}^N} \;<\; \frac{\eta}{3}.$$

Then, for $\varepsilon$ sufficiently small, we have

$$|\rho_\varepsilon * g-g|_{0,p,\mathbb{R}^N} \;<\; \frac{\eta}{3}.$$

Then

$$|\rho_\varepsilon * u - u|_{0,p,\mathbb{R}^N} \leq |u - g|_{0,p,\mathbb{R}^N} + |g - \rho_\varepsilon * g|_{0,p,\mathbb{R}^N} + |\rho_\varepsilon * (g - u)|_{0,p,\mathbb{R}^N}.$$

But by Young's inequality

$$|\rho_\varepsilon * (g - u)|_{0,p,\mathbb{R}^N} \ \leq \ |\rho_\varepsilon|_{0,1,\mathbb{R}^N} |g - u|_{0,p,\mathbb{R}^N} \ < \ \frac{\eta}{3}$$

since the integral of $\rho_\varepsilon$ is unity and the result now follows immediately. ■

**Theorem 3.1** *For $1 \leq p < \infty$, we have*

$$W_0^{m,p}(\mathbb{R}^N) \ = \ W^{m,p}(\mathbb{R}^N).$$

**Proof:** We will prove it for $m = 1$.

Step 1. Let $\{\rho_\varepsilon\}_{\varepsilon>0}$ be the family of mollifiers. Then, if $u \in W^{1,p}(\mathbb{R}^N)$, it follows from the preceding lemma that $\rho_\varepsilon * u \to u$ in $L^p(\mathbb{R}^N)$. Also, since $D^\alpha(\rho_\varepsilon * u) = \rho_\varepsilon * D^\alpha u$, for any multi-index $\alpha$, it follows, again from the preceding lemma, that $\rho_\varepsilon * u \to u$ in $W^{1,p}(\mathbb{R}^N)$ as well. Further, notice that $\rho_\varepsilon * u \in \mathcal{C}^\infty(\mathbb{R}^N)$, by the properties of convolutions.

Step 2. Let $\zeta \in \mathcal{D}(\mathbb{R}^N)$, be such that $0 \leq \zeta \leq 1$, $\zeta \equiv 1$ on $B(0;1)$ and $\text{supp}(\zeta) \subset B(0;2)$. For every positive integer $k$, define

$$\zeta_k(x) \ = \ \zeta(x/k).$$

Then $\zeta_k \equiv 1$ on $B(0;k)$ and $\text{supp}(\zeta_k) \subset B(0;2k)$. Thus $\zeta_k \in \mathcal{D}(\mathbb{R}^N)$. Choose a sequence $\varepsilon_k$ decreasing to zero. Let $u_k = \rho_{\varepsilon_k} * u$. Define

$$\varphi_k \ = \ \zeta_k u_k.$$

Then $\varphi_k \in \mathcal{D}(\mathbb{R}^N)$. From the properties of $\zeta$ it now follows that $\varphi_k = u_k$ on $B(0;k)$ and also that $|\varphi_k| \leq |u_k|$. Thus,

$$|u_k - \varphi_k|^p_{0,p,\mathbb{R}^N} \ = \ \int_{|x|>k} |u_k - \varphi_k|^p \, dx \ \leq \ 2^p \int_{|x|>k} |u_k|^p \, dx.$$

Now, by the triangle inequality (Minkowski's inequalty), we have

$$\left( \int_{|x|>k} |u_k|^p \, dx \right)^{\frac{1}{p}} \leq \left( \int_{|x|>k} |u_k - u|^p \, dx \right)^{\frac{1}{p}} + \left( \int_{|x|>k} |u|^p \, dx \right)^{\frac{1}{p}}.$$

The first term on the right hand-side of the above inequlaity can be made as small as we please for large $k$ since $u_k \to u$ in $L^p(\mathbb{R}^N)$. The second term can also be made as small as we please for large $k$, since it represents the tail of a convergent integral over $\mathbb{R}^N$. Thus, $\varphi_k \to u$ in $L^p(\mathbb{R}^N)$.

Step 3. Notice that

$$\frac{\partial \varphi_k}{\partial x_i} \ = \ \zeta_k \frac{\partial u_k}{\partial x_i} + u_k \frac{\partial \zeta_k}{\partial x_i}, \ 1 \leq i \leq N.$$

Then, since the derivatives of $u_k$ converge to those of $u$ in $L^p(\mathbb{R}^N)$, we have that the first term in the above relation converges to $\frac{\partial u}{\partial x_i}$ in $L^p(\mathbb{R}^N)$, exactly as in Step 2 above. Further

$$\frac{\partial \zeta_k}{\partial x_i} \;=\; \frac{1}{k}\frac{\partial \zeta}{\partial x_i}\left(\frac{x}{k}\right).$$

Since the derivatives of $\zeta$ are uniformly bounded, we deduce that

$$u_k \frac{\partial \zeta_k}{\partial x_i} \;\to\; 0$$

in $L^p(\mathbb{R}^N)$. Thus, it follows that

$$\frac{\partial \varphi_k}{\partial x_i} \;\to\; \frac{\partial u}{\partial x_i}$$

in $L^p(\mathbb{R}^N)$ for all $1 \le i \le N$. Hence it follows that $\varphi_k \to u$ in $W^{1,p}(\mathbb{R}^N)$ and so $u \in W_0^{1,p}(\mathbb{R}^N)$. This completes the proof. ■

# 4 Friedrich's theorem and applications

In the preceding section we saw that any element of $W^{1,p}(\mathbb{R}^N)$ can be approximated by functions from $\mathcal{D}(\mathbb{R}^N)$. We now investigate to what extent we can approximate functions in $W^{m,p}(\Omega)$ by smooth functions when $\Omega \subset \mathbb{R}^N$ is a proper open subset. We begin with a useful technical lemma.

**Lemma 4.1** *Let $u : \Omega \to \mathbb{R}$. Define $\widetilde{u} : \mathbb{R}^N \to \mathbb{R}$ by its extension by zero outside $\Omega$, i.e.*

$$\widetilde{u}(x) \;=\; \begin{cases} u(x), & \text{if} x \in \Omega, \\ 0, & \text{if} x \notin \Omega. \end{cases}$$

*Then, if $u \in W^{1,p}(\Omega)$ and if $\psi \in \mathcal{D}(\Omega)$, we have $\widetilde{\psi u} \in W^{1,p}(\mathbb{R}^N)$ and, for all $1 \le i \le N$,*

$$\frac{\partial}{\partial x_i}(\widetilde{\psi u}) \;=\; \left(\psi \frac{\partial u}{\partial x_i} + \frac{\partial \psi}{\partial x_i} u\right)^{\sim}.$$

**Proof:** Since the extension by zero maps functions in $L^p(\Omega)$ into functions in $L^p(\mathbb{R}^N)$, it is enough to prove the formula for the derivatives of $\widetilde{\psi u}$ given in the statement of the lemma. Let $\varphi \in \mathcal{D}(\mathbb{R}^N)$. Then

$$\begin{aligned} \int_{\mathbb{R}^N} \widetilde{\psi u}\frac{\partial \varphi}{\partial x_i}\, dx &= \int_\Omega \psi u \frac{\partial \varphi}{\partial x_i}\, dx = \int_\Omega u\left(\frac{\partial(\psi\varphi)}{\partial x_i} - \varphi\frac{\partial \psi}{\partial x_i}\right) dx \\ &= -\int_\Omega \frac{\partial u}{\partial x_i}\psi\varphi\, dx - \int_\Omega u\frac{\partial \psi}{\partial x_i}\varphi\, dx \\ &= -\int_{\mathbb{R}^N}\left(\psi\frac{\partial u}{\partial x_i} + u\frac{\partial \psi}{\partial x_i}\right)^{\sim}\varphi\, dx \end{aligned}$$

which completes the proof. ■

**Theorem 4.1** *(Friedrichs) Let $1 \le p < \infty$. Let $\Omega \subset \mathbb{R}^N$ be an open set and let $u \in W^{1,p}(\Omega)$. Then, there exists a sequence $\{u_n\}$ in $\mathcal{D}(\mathbb{R}^N)$ such that $u_n \to u$ in $L^p(\Omega)$ and $\frac{\partial u_n}{\partial x_i} \to \frac{\partial u}{\partial x_i}$ in $L^p(\Omega')$ for any relatively compact open set $\Omega'$ of $\Omega$.*

**Proof:** Let, as before, $\widetilde{u}$ denote the extension of $u$ by zero outside $\Omega$. Then if $\{\rho_\varepsilon\}_{\varepsilon>0}$ denotes the family of mollifiers, we know that (cf. Lemma 3.1) $\rho_\varepsilon * \widetilde{u} \to \widetilde{u}$ in $L^p(\mathbb{R}^N)$ and so $\rho_\varepsilon * \widetilde{u} \to u$ in $L^p(\Omega)$. Now, let $\Omega' \subset\subset \Omega$. Then we can find another relatively compact open subset $\Omega''$ such that

$$\Omega' \subset\subset \Omega'' \subset\subset \Omega.$$

Let $\psi \in \mathcal{D}(\Omega)$ be such that $\psi \equiv 1$ in $\Omega''$. Let $d = d(\partial\Omega', \partial\Omega'') > 0$. Now,

$$\begin{aligned} \operatorname{supp}(\rho_\varepsilon * \widetilde{\psi u} - \rho_\varepsilon * \widetilde{u}) &= \operatorname{supp}(\rho_\varepsilon * (1-\psi)\widetilde{u}) \\ &\subset B(0;\varepsilon) + \operatorname{supp}(1-\psi) \end{aligned}$$

which will be contained in $\mathbb{R}^N \backslash \Omega'$ if $\varepsilon < d$. Thus

$$\rho_\varepsilon * \widetilde{\psi u} \;=\; \rho_\varepsilon * \widetilde{u} \text{ in } \Omega'.$$

Now, $\rho_\varepsilon * \widetilde{\psi u} \in W^{1,p}(\mathbb{R}^N)$ and

$$\frac{\partial(\rho_\varepsilon * \widetilde{\psi u})}{\partial x_i} \;=\; \rho_\varepsilon * \frac{\partial(\widetilde{\psi u})}{\partial x_i} \;=\; \rho_\varepsilon * \left(\frac{\partial \psi}{\partial x_i} u + \frac{\partial u}{\partial x_i}\psi\right)^{\sim}$$

which converges, in $L^p(\mathbb{R}^N)$ to

$$\left(\frac{\partial \psi}{\partial x_i} u + \frac{\partial u}{\partial x_i}\psi\right)^{\sim}.$$

In particular, in $L^p(\Omega')$, we have

$$\frac{\partial(\rho_\varepsilon * \widetilde{u})}{\partial x_i} \;=\; \frac{\partial(\rho_\varepsilon * \widetilde{\psi u})}{\partial x_i} \to \frac{\partial u}{\partial x_i}.$$

Thus, for a sequence $\varepsilon_k \downarrow 0$, we have constructed $v_k = \rho_{\varepsilon_k} * \widetilde{u}$ such that $v_k \in \mathcal{C}^\infty(\mathbb{R}^N)$, $v_k \to u$ in $L^p(\Omega)$ and such that, for all $1 \le i \le N$, $\frac{\partial v_k}{\partial x_i} \to \frac{\partial u}{\partial x_i}$ in $L^p(\Omega')$, for any relatively compact open subset $\Omega'$ contained in $\Omega$. Now let $\zeta_k$ be as in the proof of Theorem 3.1 and set $u_k = \zeta_k v_k$ which will have the same convergence properties as the sequence $\{v_k\}$. ■

Thus, given $u \in W^{1,p}(\Omega)$, while we can approximate it in $L^p(\Omega)$ by smooth functions, its derivatives can be approximated by the corresponding sequences of derivatives only in $L^p$ of relatively compact subsets. We then ask ourselves whether it is at all possible to approximate elements in $W^{1,p}(\Omega)$ by smooth functions.

**Definition 4.1** *Let $\Omega \subset \mathbb{R}^N$ be an open set. A bounded linear operator $P : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^N)$ is said to be an* **extension operator** *if the restriction of $Pu$ to $\Omega$ is $u$ for every $u \in W^{1,p}(\Omega)$.* ■

**Theorem 4.2** *If there exists an extension operator on $W^{1,p}(\Omega)$, then given $u \in W^{1,p}(\Omega)$, there exists a sequence $\{u_n\}$ in $\mathcal{D}(\mathbb{R}^N)$ such that $u_n \to u$ in $W^{1,p}(\Omega)$.*

**Proof:** There exists a sequence $\{u_n\}$ in $\mathcal{D}(\mathbb{R}^N)$ converging to $Pu$ in $W^{1,p}(\Omega)$, as proved in Theorem 3.1. Since $Pu = u$ in $\Omega$, the result follows immediately. ■

**Corollary 4.1** *If there exists an extension operator on $W^{1,p}(\Omega)$, then $\mathcal{C}^\infty(\overline{\Omega})$ is dense in $W^{1,p}(\Omega)$.* ■

**Example 4.1** Let

$$\mathbb{R}^N_+ \;=\; \{x = (x_1, \cdots, x_N) \in \mathbb{R}^N \mid x_N > 0\}$$

be the upper half space. Then it admits an extension operator which is defined in the following way. Set $x = (x', x_N)$ where $x' = (x_1, \cdots, x_{N-1}) \in \mathbb{R}^{N-1}$. Define

$$Pu(x) \;=\; \begin{cases} u(x', x_N) & \text{if } x_N > 0, \\ u(x', -x_N) & \text{if } x_N < 0. \end{cases}$$

It can be shown that $P$ defines an exiension operator from $W^{1,p}(\mathbb{R}^N_+)$ to $W^{1,p}(\mathbb{R}^N)$. ■

**Definition 4.2** *Let $\Omega \subset \mathbb{R}^N$ be an open subset such that its boundary $\partial\Omega$ is bounded (and hence compact). Let $Q$ be a unit cube in $\mathbb{R}^N$ with centre at the origin and with edges parallel to the coordinate axes. Let $Q_+ = Q \cap \mathbb{R}^N_+$ and let $Q_0 = Q \cap \mathbb{R}^{N-1}$. The domain $\Omega$ is said to be of class $\mathcal{C}^k$, where $k$ is a non-negative integer, if for every $x \in \partial\Omega$, we can find a neighbourhood $U$ of $x$ in $\mathbb{R}^N$ and a bijective map $T : Q \to U$ such that both $T$ and $T^{-1}$ are $\mathcal{C}^k$-maps on $\overline{Q}$ and $\overline{U}$ respectively and the following relations hold:*

$$T(Q_+) \;=\; U \cap \Omega \text{ and } T(Q_0) \;=\; U \cap \partial\Omega.$$ ■

**Example 4.2** If $\Omega \subset \mathbb{R}^N$ is an open set such that its boundary is bounded and is of class $\mathcal{C}^1$, then there exists an extension operator on $W^{1,p}(\Omega)$. ■.

We now look at some very useful applications of Friedrich's theorem.

**Theorem 4.3** *(Chain Rule) Let $1 \leq p \leq \infty$. Let $\Omega \subset \mathbb{R}^N$ be an open subset. Let $G : \mathbb{R} \to \mathbb{R}$ be a continuously differentiable map such that $G(0) = 0$. Assume, further, that there exists a positive constant $M$ such that $|G'(s)| \leq M$ for all $s \in \mathbb{R}$. Then, if $u \in W^{1,p}(\Omega)$, we have $G \circ u \in W^{1,p}(\Omega)$ and,*

$$\frac{\partial(G \circ u)}{\partial x_i} \;=\; (G' \circ u)\frac{\partial u}{\partial x_i},\; 1 \leq i \leq N,$$

*where $(G \circ u)(x) = G(u(x))$.*

**Proof:** Since $G(0) = 0$, by the mean value theorem, it follows that $|G(s)| \leq Ms$, for all $s \in \mathbb{R}$. Thus, it follows that both $G \circ u$ and $(G' \circ u)\frac{\partial u}{\partial x_i}$ are in $L^p(\Omega)$. Thus it suffics to prove the formula for the derivative of $G \circ u$ given in the statement above to prove the theorem.

Case 1: Let $1 \leq p < \infty$. Let $\{u_n\}$ be a sequence in $\mathcal{D}(\mathbb{R}^N)$ as in Friedrich's theorem. Then, it is clear that $G \circ u_n \to G \circ u$ in $L^p(\Omega)$ and also that, for a subsequence that we will henceforth work with and therefore continue to denote as $\{u_n\}$, $G'(u_n(x)) \to G'(u(x))$ almost everywhere. Now, let $\varphi \in \mathcal{D}(\Omega)$. Choose a relatively compact subset $\Omega'$ contained in $\Omega$ such that $\text{supp}(\varphi) \subset \Omega'$. Since $u_n$ is smooth, we have by the classical Green's theorem (integration by parts)

$$\int_\Omega (G \circ u_n) \frac{\partial \varphi}{\partial x_i}\, dx \;=\; \int_{\Omega'} (G \circ u_n) \frac{\partial \varphi}{\partial x_i}\, dx \;=\; -\int_{\Omega'} (G' \circ u_n) \frac{\partial u_n}{\partial x_i} \varphi\, dx.$$

Then, with the convergences observed above and those guaranteed by Friedrich's theorem, we get, on passing to the limit,

$$\int_\Omega (G \circ u) \frac{\partial \varphi}{\partial x_i}\, dx \;=\; -\int_{\Omega'} (G' \circ u) \frac{\partial u}{\partial x_i} \varphi\, dx \;=\; -\int_\Omega (G' \circ u) \frac{\partial u}{\partial x_i} \varphi\, dx$$

which proves the result.

Case 2: Let $p = \infty$. If $\varphi \in \mathcal{D}(\Omega)$, choose $\Omega'$ as before. Then $u \in W^{1,\infty}(\Omega)$ implies that $u \in W^{1,\infty}(\Omega')$ and hence in $W^{1,q}(\Omega')$ for any $1 \leq q < \infty$ since $\overline{\Omega'}$ is compact. Thus, the calculations in the preceding case are still valid and the result follows. ■

**Remark 4.1** The condition $G(0) = 0$ was used only to prove that if $u \in L^p(\Omega)$,then so does $G \circ u$. If $\Omega$ were *bounded*, then constant functions are in $L^p(\Omega)$ for all $1 \leq p \leq \infty$ and the mean value theorem yields $|G(u(x))| \leq |G(0)| + M|u(x)|$ which shows that $G \circ u \in L^p(\Omega)$ and so the condition $G(0) = 0$ is no longer necessary in that case. ■

**Theorem 4.4** *Let $\Omega \subset \mathbb{R}^N$ be a bounded open set. Let $1 < p < \infty$. If $u \in W^{1,p}(\Omega)$, then $|u| \in W^{1,p}(\Omega)$ and*

$$\frac{\partial |u|}{\partial x_i} \;=\; \text{sgn}(u) \frac{\partial u}{\partial x_i},\; 1 \leq i \leq N$$

*where*

$$\text{sgn}(u)(x) \;=\; \begin{cases} +1, & \text{if } u(x) > 0, \\ 0, & \text{if } u(x) = 0, \\ -1, & \text{if } u(x) < 0. \end{cases}$$

**Proof:** Let $\varepsilon > 0$. Let $f_\varepsilon(t) = \sqrt{t^2 + \varepsilon}$. Then $f_\varepsilon \in \mathcal{C}^1(\mathbb{R})$ and

$$f'_\varepsilon(t) \;=\; \frac{t}{\sqrt{t^2 + \varepsilon}}$$

which shows that $|f'_\varepsilon(t)| \leq 1$. Since $\Omega$ is bounded, the chain rule holds (even though $f_\varepsilon(0) \neq 0$). Thus, $u \in W^{1,p}(\Omega)$ implies that $f_\varepsilon \circ u \in W^{1,p}(\Omega)$ and

$$\frac{\partial (f_\varepsilon \circ u)}{\partial x_i} \;=\; \frac{u}{\sqrt{u^2 + \varepsilon}} \frac{\partial u}{\partial x_i},\; 1 \leq i \leq N.$$

Now, $|f_\varepsilon(t) - |t|\,| \le \sqrt{\varepsilon}$. Hence $f_\varepsilon \circ u \to |u|$ in $L^p(\Omega)$. Further,

$$\int_\Omega \left| \frac{\partial (f_\varepsilon \circ u)}{\partial x_i} \right|^p \, dx \;\le\; \int_\Omega \left| \frac{\partial u}{\partial x_i} \right|^p \, dx$$

which is bounded. Thus, since $1 < p < \infty$, we deduce that (cf. Remark 2.1) $|u| \in W^{1,p}(\Omega)$ and that

$$\frac{\partial |u|}{\partial x_i} \;=\; \lim_{\varepsilon \to 0} \frac{\partial (f_\varepsilon \circ u)}{\partial x_i}$$

in $L^p(\Omega)$ if that limit exists. But

$$\frac{u}{\sqrt{u^2 + \varepsilon}} \frac{\partial u}{\partial x_i} \;\to\; \operatorname{sgn}(u) \frac{\partial u}{\partial x_i}$$

pointwise and since the $p$-th powers of the absolute values of all these functions are bounded by $|\frac{\partial u}{\partial x_i}|^p$, it follows from the dominated convergence theorem that the $L^p$-norms of these functions converge to the $L^p$-norm of $\frac{\partial u}{\partial x_i}$. The pointwise convergence and the convergence of the norm implies convergence in $L^p$ and this completes the proof. ■

**Theorem 4.5** *Let $\Omega \subset \mathbb{R}^N$ be a bounded open set and let $1 < p < \infty$. Let $u \in W^{1,p}(\Omega)$. Then, for any $t \in \mathbb{R}$, and for any $1 \le i \le N$, we have that $\frac{\partial u}{\partial x_i} = 0$ almost everywhere on the set $\{x \in \Omega \mid u(x) = t\}$.*

**Proof:** Let $u \ge 0$. Then $u = |u|$. Thus,

$$\operatorname{sgn}(u) \frac{\partial u}{\partial x_i} \;=\; \frac{\partial |u|}{\partial x_i} \;=\; \frac{\partial u}{\partial x_i}.$$

It then follows that on the set $\{x \in \Omega \mid u(x) = 0\}$, we have that $\frac{\partial u}{\partial x_i} = 0$ almost everywhere. If $u \in W^{1,p}(\Omega)$, then $u = u^+ - u^-$, where

$$u^+ \;=\; \frac{1}{2}(|u| + u), \;\; u^- \;=\; \frac{1}{2}(|u| - u).$$

Then

$$\{x \in \Omega \mid u(x) = 0\} \;=\; \{x \in \Omega \mid u^+(x) = 0\} \cap \{x \in \Omega \mid u^-(x) = 0\}.$$

Then $\frac{\partial u}{\partial x_i} = \frac{\partial (u^+)}{\partial x_i} - \frac{\partial (u^-)}{\partial x_i} = 0$ almost everywhere on this set. For any $t \in \mathbb{R}$, consider the function $u - t$. ■

**Remark 4.2** Notice that the set $\{x \in \Omega \mid u(x) = t\}$ may itself be of measure zero, in which case the above result gives no new information. However, if the function takes a constant value on a set of positive measure, then its derivative vanishes almost everywhere in that set. ■

**Theorem 4.6** *(Stampacchia) Let $\Omega \subset \mathbb{R}^N$ be a bounded open set and let $1 < p < \infty$. Let $f : \mathbb{R} \to \mathbb{R}$ be a Lipschitz continuous function. Then, if $u \in W^{1,p}(\Omega)$, we have $f \circ u \in W^{1,p}(\Omega)$ and if $f'$ is continuous except at a finite number of points $\{t_1, \cdots, t_k\}$, then*

$$\frac{\partial(f \circ u)}{\partial x_i}(x) \ = v_i \ \stackrel{\text{def}}{=} \ \begin{cases} (f' \circ u)\frac{\partial u}{\partial x_i}(x), & \text{if } u(x) \notin \{t_1, \cdots, t_k\}, \\ 0, & \text{otherwise.} \end{cases}$$

**Proof:** First of all, it follows from the Lipschitz continuity of $f$ and the boundedness of $\Omega$ that $f \circ u \in L^p(\Omega)$. If $M$ is the Lipschitz constant of $f$, then we also have that $|f'(t)| \leq M$. Then it also follows that $v_i \in L^p(\Omega)$ for $1 \leq i \leq N$.

Let $\{\rho_\varepsilon\}_{\varepsilon>0}$ denote the family of mollifiers in $\mathbb{R}$. Choose a sequence $\varepsilon_n \downarrow 0$ and set $f_n = \rho_{\varepsilon_n} * f$. Then $f_n \in \mathcal{C}^\infty(\mathbb{R})$ and (cf. Lemma 3.1) $f_n(t) \to f(t)$ for all $t \in \mathbb{R}$. Now,

$$\begin{array}{rcl} |f_n(t) - f_n(t')| & = & \left|\int_{|s|<\varepsilon_n} (f(t-s) - f(t'-s))\rho_{\varepsilon_n}(s)\ ds\right| \\ & \leq & M|t-t'| \int_{|s|<\varepsilon_n} \rho_{\varepsilon_n}(s)\ ds \ = \ M|t-t'|. \end{array}$$

Thus, it follows that $|f_n'(t)| \leq M$ for all $t \in \mathbb{R}$. We also have that $f_n \circ u \to f \circ u$ in $L^p(\Omega)$ since for any $x \in \Omega$,

$$|f_n(u(x)) - f(u(x))| \ = \ \left|\int_{|s|<\varepsilon_n} (f(u(x)-s) - f(u(x)))\rho_{\varepsilon_n}(s)\ ds\right| \ \leq \ M\varepsilon_n$$

which implies that $f_n \circ u \to f \circ u$ uniformly on $\Omega$ and, as $\Omega$ is bounded, in $L^p(\Omega)$ as well. Now, $f_n \circ u \in W^{1,p}(\Omega)$ and

$$\frac{\partial(f_n \circ u)}{\partial x_i} = (f_n' \circ u)\frac{\partial u}{\partial x_i}$$

which is clearly bounded in $L^p(\Omega)$. Thus, by Remark 2.1, we do have that $f \circ u \in W^{1,p}(\Omega)$ and that

$$\frac{\partial(f \circ u)}{\partial x_i} \ = \ \lim_{n\to\infty} \frac{\partial(f_n \circ u)}{\partial x_i}$$

in $L^p(\Omega)$, if that limit exists.

Now, assume that the derivative of $f$ exists and is continuous except at a finite number of points $\{t_i\}_{i=1}^k$. Let $E_i = \{x \in \Omega \mid u(x) = t_i\}, 1 \leq i \leq k$. Set $E = \cup_{i=1}^k E_i$. Then

$$(f_n' \circ u)\frac{\partial u}{\partial x_i} \ \to \ (f' \circ u)\frac{\partial u}{\partial x_i}$$

on $\Omega \backslash E$. On $E$, we have that $\frac{\partial u}{\partial x_i} = 0$ almost everywhere by the preceding theorem. Thus $(f_n' \circ u)\frac{\partial u}{\partial x_i} = 0$ almost everywhere on $E$. It follows that $\frac{\partial f_n \circ u}{\partial x_i} \to v_i$ pointwise almost everywhere in $\Omega$. Also

$$\left|(f_n' \circ u)\frac{\partial u}{\partial x_i} - v_i\right| \ \leq \ 2M\left|\frac{\partial u}{\partial x_i}\right|$$

and so by the dominated convergence theorem, we deduce that the limit in $L^p(\Omega)$ of $\frac{\partial((f_n \circ u)}{\partial x_i}$ is indeed $v_i$ for all $1 \leq i \leq N$, which completes the proof. ■

**Proposition 4.1** *Let $u \in W^{1,p}(\Omega)$, $1 \leq p < \infty$. Let $K \subset \Omega$ be compact. If $u$ vanishes on $\Omega \backslash K$, then $u \in W_0^{1,p}(\Omega)$.*

**Proof:** Choose relatively compact sets $\Omega''$ and $\Omega'$ such that

$$K \subset \Omega'' \subset\subset \Omega' \subset\subset \Omega.$$

Let $\psi \in \mathcal{D}(\Omega)$ such that $\psi \equiv 1$ on $\Omega''$ and such that $\text{supp}(\psi) \subset \Omega'$. Then $\psi u = u$. Let $\{u_n\}$ be a sequence in $\mathcal{D}(\mathbb{R}^N)$ such that $u_n \to u$ in $L^p(\Omega)$ and $\frac{\partial u_n}{\partial x_i} \to \frac{\partial u}{\partial x_i}$ in $L^p$ of all relatively compact open subsets of $\Omega$.Then $\psi u_n \in \mathcal{D}(\Omega)$ and $\psi u_n \to \psi u = u$ in $W^{1,p}(\Omega)$ since all the supports are contained in $\Omega' \subset\subset \Omega$. This shows that $u \in W_0^{1,p}(\Omega)$. ■

**Proposition 4.2** *Let $\Omega$ be bounded and let $1 < p < \infty$. If $f : \mathbb{R} \to \mathbb{R}$ is a Lipschitz continuous function such that $f(0) = 0$ and whose derivative exists and is continuous except at a finite number of points, then, if $u \in W_0^{1,p}(\Omega)$,we also have $f \circ u \in W_0^{1,p}(\Omega)$.*

**Proof:** Let $\{u_n\}$ be a sequence in $\mathcal{D}(\Omega)$ converging in $W^{1,p}(\Omega)$ to $u$. Then, since $f$ is Lipschitz continuous, we have

$$|f(u_n(x)) - f(u(x))| \ \leq \ M|u_n(x) - u(x)|$$

and it is now immediate that $f \circ u_n \to f \circ u$ in $L^p(\Omega)$. Further, it is clear from the formula for the derivatives that $f' \circ u_n$ is bounded in $L^p(\Omega)$. Thus it follows that, for a subsequence,$f \circ u_{n_k} \rightharpoonup f \circ u$ *weakly* in $W^{1,p}(\Omega)$ (cf. Remark 2.1). But, since $f(0) = 0$, it follows that $f \circ u_n$ vanishes outside the support of $u_n$, which is compact. Thus, by the preceding proposition, $f \circ u_n \in W_0^{1,p}(\Omega)$ and so it follows that $f \circ u \in W_0^{1,p}(\Omega)$ as well (since a closed subspace is also weakly closed). ■

In particular, if $u \in W_0^{1,p}(\Omega)$, it follows that the functions $|u| \in W_0^{1,p}(\Omega)$ and so we also have that $u^+$ and $u^-$ are in $W_0^{1,p}(\Omega)$. This fact is very useful when deriving maximum principles for second order elliptic partial differential equations. Stampacchia's theorem is not valid in spaces $W^{m,p}(\Omega)$ when $m > 1$ and this will, in some sense, explain why we do not have good maximum principles for higher order elliptic problems.

# 5 Poincaré's inequality

In the previous section, we saw that the existence of an extension operator for $W^{1,p}(\Omega)$ depended on the nature of the domain. In general, the extension by zero will not map $W^{1,p}(\Omega)$ into $W^{1,p}(\mathbb{R}^N)$. For example, let $\Omega = (0,1) \subset \mathbb{R}$ and consider the function $u \equiv 1$ in $\Omega$. Then $\widetilde{u} \in L^p(\mathbb{R})$ for any $1 \leq p \leq \infty$. But

$$\frac{d\widetilde{u}}{dx} \ = \ \delta_0 - \delta_1$$

where $\delta_0(\varphi) = \varphi(0)$ and $\delta_1(\varphi) = \varphi(1)$ for $\varphi \in \mathcal{D}(\mathbb{R})$, and this cannot come from any locally integrable function, as already observed earlier.

We will now show that, irrespective of the nature of the domain, the extension by zero provides an extension operator from $W_0^{1,p}(\Omega)$ to $W^{1,p}(\mathbb{R}^N)$ for all $1 \leq p < \infty$.

**Theorem 5.1** *Let $1 \leq p < \infty$. Let $\widetilde{u}$ denote the extension by zero outside $\Omega$ for any function $u$ defined on $\Omega \subset \mathbb{R}^N$. Then, if $u \in W_0^{1,p}(\Omega)$, we have that $\widetilde{u} \in W^{1,p}(\mathbb{R}^N)$ and that*

$$\frac{\partial \widetilde{u}}{\partial x_i} \;=\; \widetilde{\frac{\partial u}{\partial x_i}},\; 1 \leq i \leq N.$$

**Proof:** The extension by zero maps functions in $L^p(\Omega)$ into the space $L^p(\mathbb{R}^N)$. Thus it suffices to prove the above formula for the derivatives of $\widetilde{u}$.

Since $u \in W_0^{1,p}(\Omega)$, let $\{u_n\}$ be a sequence in $\mathcal{D}(\Omega)$ converging to $u$ in $W^{1,p}(\Omega)$. Let $\varphi \in \mathcal{D}(\mathbb{R}^N)$. Then, by classical integration by parts for smooth functions, we have

$$\int_\Omega u_n \frac{\partial \varphi}{\partial x_i}\, dx \;=\; -\int_\Omega \frac{\partial u_n}{\partial x_i}\varphi\, dx.$$

There is no boundary term since $u_n \in \mathcal{D}(\Omega)$. Passing to the limit, we see that

$$\int_\Omega u \frac{\partial \varphi}{\partial x_i}\, dx \;=\; -\int_\Omega \frac{\partial u}{\partial x_i}\varphi\, dx.$$

In other words,

$$\int_{\mathbb{R}^N} \widetilde{u} \frac{\partial \varphi}{\partial x_i}\, dx \;=\; -\int_{\mathbb{R}^N} \widetilde{\frac{\partial u}{\partial x_i}}\varphi\, dx$$

which completes the proof. ■

**Theorem 5.2** *(Poincaré's inequality) Let $\Omega \subset \mathbb{R}^N$ be a bounded open subset. Let $1 \leq p < \infty$. Then there exists a constant $C = C(p, \Omega) > 0$ such that*

$$|u|_{0,p,\Omega} \;\leq\; C|u|_{1,p,\Omega}$$

*for all $u \in W_0^{1,p}(\Omega)$. In particular, the mapping $u \mapsto |u|_{1,p,\Omega}$ is a norm on $W_0^{1,p}(\Omega)$ which is equivalent to the norm $\|u\|_{1,p,\Omega}$. On $H_0^1(\Omega)$, the bilinear form*

$$< u, v >_{1,\Omega} = \int_\Omega \sum_{i=1}^N \frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i}\, dx \;=\; \int_\Omega \nabla u.\nabla v\, dx$$

*defines an inner-product yielding the norm $|.|_{1,\Omega}$ which is equivalent to the norm $\|.\|_{1,\Omega}$.*

**Proof:** Let $\Omega = (-a, a)^N,\ a > 0$. Let $u \in \mathcal{D}(\Omega)$. Assume that $1 < p < \infty$. Then,

$$u(x) \;=\; \int_{-a}^{x_N} \frac{\partial u}{\partial x_N}(x', t)\, dt$$

where $x = (x', x_N)$ and $x' = (x_1, \cdots, x_{N-1})$. Then, by Hölder's inequality,

$$|u(x)| \ \leq \ \left( \left| \frac{\partial u}{\partial x_N}(x', t) \right|^p \ dt \right)^{\frac{1}{p}} |x_N + a|^{\frac{1}{p'}}$$

where $p' = p/(p-1)$ is the conjugate exponent. Thus,

$$|u(x)|^p \ \leq \ (2a)^{\frac{p}{p'}} \int_{-a}^{a} \left| \frac{\partial u}{\partial x_N}(x', t) \right|^p \ dt$$

which yields, on integrating with respect to $x'$,

$$\int |u(x', x_N)|^p \ dx' \ \leq \ (2a)^{\frac{p}{p'}} \int_{\Omega} \left| \frac{\partial u}{\partial x_N} \right|^p \ dx$$

which in turn yields, on integrating with respect to $x_N$,

$$\int_{\Omega} |u|^p \ dx \ \leq \ (2a)^{\frac{p}{p'}+1} \int_{\Omega} \left| \frac{\partial u}{\partial x_N} \right|^p \ dx.$$

It is easy to deduce this inequality when $p = 1$ as well. Thus for $1 \leq p < \infty$, we deduce from this that

$$|u|_{0,p,\Omega} \ \leq \ 2a|u|_{1,p,\Omega}$$

for all $u \in \mathcal{D}(\Omega)$. But since $\mathcal{D}(\Omega)$ is dense in $W_0^{1,p}(\Omega)$, the above inequality also holds for all $u \in W_0^{1,p}(\Omega)$.

Now if $\Omega$ is an arbitrary bounded domain, then we can find $a > 0$ such that $\Omega \subset (-a, a)^N = \widetilde{\Omega}$. Let $\widetilde{u}$ denote the extension of $u$ by zero outside $\Omega$. Then, if $u \in W_0^{1,p}(\Omega)$, it follows that $\widetilde{u} \in W_0^{1,p}(\widetilde{\Omega})$. Thus

$$|u|_{0,p,\Omega} \ = \ |\widetilde{u}|_{0,p,\widetilde{\Omega}} \ \leq \ 2a|\widetilde{u}|_{1,p,\widetilde{\Omega}} \ = \ 2a|u|_{1,p,\Omega}$$

since $\frac{\partial \widetilde{u}}{\partial x_i} = \widetilde{\frac{\partial u}{\partial x_i}}$. This completes the proof. ■

We now make several important remarks.

- This inequality will be very crucial in the study of Dirichlet boundary value problems for elliptic partial differential equations.
- It is easy to see that this proof works when $\Omega$ is unbounded, but is bounded in some particular direction, *i.e.* if $\Omega$ is contained in some infinite strip of finite width. However, it is not true for truly unbounded regions. In particular, if $\Omega = \mathbb{R}^N$, consider $\zeta_k$ as in Step 2 of the proof of Theorem 3.1. Then

  $$\frac{\partial \zeta_k}{\partial x_i} \ = \ \frac{1}{k} \frac{\partial \zeta}{\partial x_i}(\left(\frac{x}{k}\right).$$

  If $p > N$, it is then easy to see that $|\zeta_k|_{1,p,\mathbb{R}^N} \to 0$ as $k \to \infty$ while

  $$|\zeta_k|_{0,p,\mathbb{R}^N} \ \geq \ |B(0;k)|^{\frac{1}{p}} \ \to \ \infty$$

  (where $|B(0;k)|$ denotes the Lebesgue measure of the ball $B(0;k)$) since $\zeta_k \equiv 1$ on $B(0;k)$. Thus we cannot have Poincaré's inequality in such domains.

- This inequality also shows that if $\Omega$ is bounded, then $W^{1,p}(\Omega)$ and $W_0^{1,p}(\Omega)$ cannot be equal (which was the case for $\Omega = \mathbb{R}^N$). For the constant function $u \equiv 1$ on $\Omega$ is in $W^{1,p}(\Omega)$ and $|u|_{1,p,\Omega} = 0$ but $|u|_{0,p,\Omega} > 0$. Hence the constant function cannot belong to $W_0^{1,p}(\Omega)$.

- If $u \in W_0^{2,p}(\Omega)$ where $\Omega$ is bounded, then $u, \frac{\partial u}{\partial x_i} \in W_0^{1,p}(\Omega)$, for all $1 \leq i \leq N$. Thus, it follows that

$$\left|\frac{\partial u}{\partial x_i}\right|_{0,p,\Omega} \leq C \left|\frac{\partial u}{\partial x_i}\right|_{1,p,\Omega}$$

  which is the same as saying

$$|u|_{1,p,\Omega} \leq C|u|_{2,p,\Omega}.$$

  Thus, for some constant $C' = C'(p, \Omega) > 0$ we get

$$|u|_{0,p,\Omega} \leq C'|u|_{2,p,\Omega}.$$

  In general, there exists a constant $C > 0$, delpending only on $m, p$ and $\Omega$, such that

$$|u|_{0,p,\Omega} \leq C|u|_{m,p,\Omega}$$

  for all $u \in W_0^{m,p}(\Omega)$ when $m \geq 1$ is an integer and $1 \leq p < \infty$.

- We saw that the constant $C$ in Poincaré's inequality depended in some way on the diameter of $\Omega$. This is not the best possible constant. If $p = 2$, it turns out that the best constant is connected to 'the principal eigenvalue of the Laplace operator with homogeneous Dirichlet boundary conditions on $\partial\Omega$'. For instance, if $\Omega = (0, 1)$, we get that the constant is unity from the proof of Poincaré's inequality. But the best constant can be shown to be $\frac{1}{\pi}$. Thus, for all $u \in H_0^1(0, 1)$, we have

$$\left(\int_0^1 |u|^2 \, dx\right)^{\frac{1}{2}} \leq \frac{1}{\pi}\left(\int_0^1 |u'|^2 \, dx\right)^{\frac{1}{2}}.$$

  Equality is achieved for the function $u = \sin \pi x$ and constant multiples thereof.

# 6 Imbedding theorems

We know that if $u \in W^{1,p}(\Omega)$, then, $u \in L^p(\Omega)$. We now ask the question if the information that the first order partial derivatives of $u$ are also in $L^p(\Omega)$ will give us more information on the function $u$, *vis-à-vis* its smoothness or integrability with respect to other exponents. In particular we would like to know if $W^{1,p}(\mathbb{R}^N)$ is continuously imbedded in $L^{p^*}(\mathbb{R}^N)$, where $p \neq p^*$. In other words, we are looking for an inequality of the form

$$|u|_{0,p^*,\mathbb{R}^N} \leq C|u|_{1,p,\mathbb{R}^N}$$

for all $u \in W^{1,p}(\mathbb{R}^N)$, where $C > 0$ is a constant which depends only on $p$ and $N$. A simple analysis will show us when this will be possible, if at all, and what is the value of $p^*$ that we should expect. Let $\lambda > 0$ be a fixed real number. Then consider the scaling $x \mapsto \lambda x$ which maps $\mathbb{R}^N$ onto itself. Let $u \in W^{1,p}(\mathbb{R}^N)$. Then if we set $u_\lambda(x) = u(\lambda x)$, it follows that $u_\lambda \in W^{1,p}(\mathbb{R}^N)$ as well. We have

$$\frac{\partial u_\lambda}{\partial x_i}(x) \;=\; \lambda \frac{\partial u}{\partial x_i}(\lambda x).$$

Consequently, by the change of variable formula, we get

$$\int_{\mathbb{R}^N} \left| \frac{\partial u_\lambda}{\partial x_i}(x) \right|^p \, dx \;=\; \lambda^p \int_{\mathbb{R}^N} \left| \frac{\partial u}{\partial x_i}(\lambda x) \right|^p \, dx \;=\; \lambda^{p-N} \int_{\mathbb{R}^N} \left| \frac{\partial u}{\partial x_i}(x) \right|^p \, dx.$$

Thus,

$$|u_\lambda|_{1,p,\mathbb{R}^N} \;=\; \lambda^{1-\frac{N}{p}} |u|_{1,p,\mathbb{R}^N}.$$

Similarly

$$|u_\lambda|_{0,p^*,\mathbb{R}^N} \;=\; \lambda^{-\frac{N}{p^*}} |u|_{0,p^*,\mathbb{R}^N}.$$

hence, for all $\lambda > 0$, we must have

$$0 \;<\; C \;\leq\; \frac{|u_\lambda|_{1,p,\mathbb{R}^N}}{|u_\lambda|_{0,p^*,\mathbb{R}^N}} \;=\; \lambda^{1-\frac{N}{p}+\frac{N}{p^*}} \frac{|u|_{1,p,\mathbb{R}^N}}{|u|_{0,p^*,\mathbb{R}^N}}.$$

If $1 - \frac{N}{p} + \frac{N}{p^*}$ is strictly positive, then we let $\lambda \to 0$ to get a contradiction and if it is strictly negative, we can let $\lambda \to \infty$ to get a contradiction. Thus, if at all we can hope for such an inequality, it follows that we must have that this number is zero, *i.e.*

$$\frac{1}{p^*} \;=\; \frac{1}{p} - \frac{1}{N}.$$

For this to be possible, it is necessary that $p < N$ as well. In that case notice that we also have $p^* > p$.

We now split our investigation into three cases, *viz.*, $p < N, p = N$ and $p > N$.

**Theorem 6.1** *(Sobolev's Inequality) Let $1 \leq p < N$ and define $p^*$ as above. Then, there exists $C = C(p, N) > 0$ such that for all $u \in W^{1,p}(\mathbb{R}^N)$,*

$$|u|_{0,p^*,\mathbb{R}^N} \;\leq\; C|u|_{1,p,\mathbb{R}^N}.$$

*In particular, we have the continuous inclusion*

$$W^{1,p}(\mathbb{R}^N) \;\hookrightarrow L^{p^*}(\mathbb{R}^N) \;\blacksquare$$

.

The proof of this result is rather technical and we refer the reader to Kesavan [1].

**Corollary 6.1** *Let $1 \leq p < N$. Then we have the continuous inclusions*

$$W^{1,p}(\mathbb{R}^N) \;\hookrightarrow\; L^q(\mathbb{R}^N)$$

*for all $q \in [p, p^*]$.*

**Proof:** Let $p < q < p^*$. Then, there exists $\alpha \in (0,1)$ such that

$$\frac{1}{q} \;=\; \frac{\alpha}{p} + \frac{1-\alpha}{p^*}.$$

Then $|u|^{\alpha q} \in L^{\frac{p}{\alpha q}}(\mathbb{R}^N)$ and $|u|^{(1-\alpha)q} \in L^{\frac{p^*}{(1-\alpha)q}}(\mathbb{R}^N)$ if $u \in W^{1,p}(\mathbb{R}^N)$.Thus, since $q = \alpha q + (1-\alpha)q$, we get, by Hölder's inequality,

$$\begin{array}{rcl} |u|_{0,q,\mathbb{R}^N} & \leq & |u|^{\alpha}_{0,p,\mathbb{R}^N} |u|^{1-\alpha}_{0,p^*,\mathbb{R}^N} \\ & \leq & \alpha |u|_{0,p,\mathbb{R}^N} + (1-\alpha)|u|_{0,p^*,\mathbb{R}^N} \\ & \leq & \alpha |u|_{0,p,\mathbb{R}^N} + C(1-\alpha)|u|_{1,p,\mathbb{R}^N} \\ & \leq & C'\|u\|_{1,p,\mathbb{R}^N}. \end{array}$$

(The inequality in the second line above comes from the generalised AM-GM inequality.) This completes the proof. ■

**Corollary 6.2** *Let $\Omega \subset \mathbb{R}^N$, and let $1 \leq p < N$. Then, there exists $C > 0$ such that, for all $u \in W^{1,p}_0(\Omega)$,*

$$\begin{array}{rcl} |u|_{0,p^*,\Omega} & \leq & C|u|_{1,p,\Omega} \\ |u|_{0,q,\Omega} & \leq & C\|u\|_{1,p,\Omega}, \text{ for all } p < q < p^*. \end{array}$$

*In particular, for all $p \leq q \leq p^*$, we have the continuous inclusions*

$$W^{1,p}_0(\Omega) \;\hookrightarrow L^q(\Omega).$$

*If $\Omega = \mathbb{R}^N_+$ or if $\Omega$ has bounded boundary and is of class $\mathcal{C}^1$, then we also have the continuous inclusions*

$$W^{1,p}(\Omega) \;\hookrightarrow L^q(\Omega)$$

*for all $p \leq q \leq p^*$.*

**Proof:** The extension by zero imbeds $W^{1,p}_0(\Omega)$ into $W^{1,p}(\mathbb{R}^N)$ and now it is easy to see the conclusions for the space $W^{1,p}_0(\Omega)$. If $\Omega = \mathbb{R}^N_+$ or if $\Omega$ has bounded boundary and is of class $\mathcal{C}^1$, then there exists an extension operator

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

and again the results follow immediately. ■

Let us rewrite the Sobolev inequality for $W^{1,p}_0(\Omega)$ , where $\Omega$ is a bounded open subset of $\mathbb{R}^N$ in the following form: there exists a constant $S > 0$ such that, for all $u \in W^{1,p}_0(\Omega)$,

$$|u|_{1,p,\Omega} \;\geq\; S|u|_{0,p^*,\Omega}.$$

The best possible constant $S(p,N,\Omega)$ is then given by

$$S(p,N,\Omega) \;=\; \inf_{\substack{u \in W^{1,p}_0(\Omega) \\ u \neq 0}} \frac{|u|_{1,p,\Omega}}{|u|_{0,p^*,\Omega}}.$$

Now if $\Omega_1 \subset \Omega_2$ are two bounded open subsets of $\mathbb{R}^N$, then the extension by zero outside $\Omega_1$ maps an element of $W_0^{1,p}(\Omega_1)$ into $W_0^{1,p}(\Omega_2)$, keeping the norms unchanged in their value. It follows from this immediately that $S(p, N, \Omega_1) \geq S(p, N, \Omega_2)$. Now let $B_1 = B(0; r_1)$ and $B_2 = B(0; r_2)$ be two concentric balls centered at the origin. Then the scalings

$$x \mapsto \frac{r_2}{r_1} x \text{ and } x \mapsto \frac{r_1}{r_2} x$$

map functions in $W_0^{1,p}(B_1)$ into functions in $W_0^{1,p}(B_2)$ and *vice-versa.* It then follows that

$$S(p, N, B_1) \;=\; S(p, N, B_2).$$

Since the Lebesgue measure is translation invariant, this is also true if the balls are concentric but are centered elsewhere in $\mathbb{R}^N$. Now if $\Omega$ is any bounded open set, we can always find two concentric balls one within $\Omega$ and the other containing $\Omega$. Thus it follows that $S(p, N, \Omega)$ is independent of the domain. In fact for all bounded domains $\Omega$, we have that $S(p, N, \Omega) = S(p, N)$, where

$$S(p, N) \;=\; \inf_{\substack{u \in W^{1,p}(\mathbb{R}^N) \\ u \neq 0}} \frac{|u|_{1,p,\mathbb{R}^N}}{|u|_{0,p^*,\mathbb{R}^N}}.$$

The value of this best constant has been worked out (independently) by Aubin and Talenti. When $p = 2$, the only minimizers of this optimization problem have been shown to be the functions $U(x), U_\varepsilon(x - x_0)$, where $\varepsilon > 0, x_0 \in \mathbb{R}^N$ and

$$\begin{aligned} U(x) &= C(1 + |x|^2)^{-\frac{N-2}{2}} \\ U_\varepsilon(x) &= C_\varepsilon(\varepsilon + |x|^2)^{-\frac{N-2}{2}} \end{aligned}$$

where $C$ and $C_\varepsilon$ are positive normalization constants.

In particular, this shows that for any $\Omega$ a bounded open subset of $\mathbb{R}^N$,the minimization problem stated above can never have a solution. For, if there were one, then the extension by zero outside $\Omega$ would be a minimizer for the problem in $\mathbb{R}^N$ as well, but it has been shown that the only minimizers in $\mathbb{R}^N$ are the functions $U$ and $U_\varepsilon$ above and they never vanish anywhere. This is true for all $1 \leq p < \infty$.

**Theorem 6.2** *Let $p = N$. Then we have the continuous inclusions*

$$W^{1,N}(\mathbb{R}^N) \;\hookrightarrow L^q(\mathbb{R}^N)$$

*for all $q \in [N, \infty)$. If $\Omega = \mathbb{R}^N_+$ or if $\Omega$ is an open subset with bounded boundary and is of class $\mathcal{C}^1$, then the above result is true with $\Omega$ replacing $\mathbb{R}^N$. If $\Omega$ is any open set, we have the continuous inclusions*

$$W_0^{1,N}(\Omega) \;\hookrightarrow L^q(\Omega)$$

*for all $q \in [N, \infty)$.* ■

Once again, we refer the reader to Kesavan [1] for the proof.

We do not have the inclusion of any of the spaces mentioned in the preceding theorem in $L^\infty$ of the corresponding domain.

**Example 6.1** Let $\Omega = B(0; \frac{1}{2}) \subset \mathbb{R}^2$. Let

$$u(x) \;=\; \log\log\frac{2}{|x|},\; x \neq 0.$$

Then $u \in H^1(\Omega)$, *i.e.* $p = 2 = N$, but clearly $u \notin L^\infty(\Omega)$. ■

We now turn to the case $p > N$. To motivate the theorem, let us consider the case $N = 1$ and $p > 1$.

**Example 6.2** Let $1 < p < \infty$. Let $I = (0, 1)$ and consider $u \in W^{1,p}(I)$. Then $u' \in L^p(I)$ and so is integrable. Thus the function

$$\overline{u}(x) \;=\; \int_0^x u'(t)\; dt$$

is absolutely continuous and its derivative (both in the classical and distributional sense) is $u'$. Thus, we have that $(u - \overline{u})' = 0$ and it follows that

$$u \;=\; \overline{u} + c,\; \text{a.e.}$$

where $c$ is a constant, since $u - \overline{u}$ is a constant distribution as was shown in Section 1. Hence it follows that we can consider $u$ as an absolutely continuous function on $I$ and so extends continuously to $\overline{I} = [0, 1]$. (Recall that elements of $L^p$ are only equivalence classes of functions; by saying that an element of $W^{1,p}(I)$ is continuous, we mean that the equivalence class of any $u \in W^{1,p}(I)$ has a representative that is absolutely continuous.) Thus, we can write

$$u(x) \;=\; u(0) + \int_0^x u'(t)\; dt,\; x \in I.$$

We then see that

$$|u(0)| \;\leq\; |u(x)| + |u'|_{0,p,I}|x|^{\frac{1}{p'}} \;\leq\; |u(x)| + |u'|_{0,p,I}$$

where $p'$ is the conjugate exponent of $p$, using Hölder's inequality. Now applying the triangle inequality in $L^p$ to the functions $|u(x)|$ and the constant function $|u'|_{0,p,I}$, we get that

$$|u(0)| \;\leq\; |u|_{0,p,I} + |u'|_{0,p,I} \;\leq\; C\|u\|_{1,p,I}$$

where $C$ only depends on $p$. Again,

$$|u(x)| \;\leq\; |u(0)| + |u'|_{0,p,I} \;\leq\; C\|u\|_{1,p,I}$$

for any $x \in \overline{I}$. Thus we have established the continuous inclusion

$$W^{1,p}(I) \;\hookrightarrow \mathcal{C}(\overline{I}).$$

In addition, we have, for $x, y \in \overline{I}$,

$$u(x) - u(y) \;=\; \int_x^y u'(t)\; dt$$

from which we deduce that

$$|u(x) - u(y)| \;\leq\; C|u|_{1,p,I}|x-y|^{\frac{1}{p'}}$$

and notice that $\frac{1}{p'} = 1 - \frac{1}{p}$. Thus all the functions in $W^{1,p}(I)$ are not only absolutely continuous, but are also Hölder continuous with exponent $1-\frac{1}{p}$. ■

**Theorem 6.3** *Let $p > N$. Then we have the continuous inclusion*

$$W^{1,p}(\mathbb{R}^N) \;\hookrightarrow\; L^\infty(\mathbb{R}^N)$$

*and there exists a constant $C = C(N,p) > 0$ such that, for all $u \in W^{1,p}(\mathbb{R}^N)$, and for almost all $x, y \in \mathbb{R}^N$, we have*

$$|u(x) - u(y)| \;\leq\; C|u|_{1,p,\mathbb{R}^N}|x-y|^{1-\frac{N}{p}}.$$

*The same conclusions hold when $\mathbb{R}^N$ is replaced by $\mathbb{R}^N_+$ or by $\Omega$ with a bounded boundary and of class $\mathcal{C}^1$. The analogous result is also true for $W^{1,p}_0(\Omega)$ when $\Omega \subset \mathbb{R}^N$ is any open set. In particular, if $\Omega$ is a bounded open set, we can consider its elements as being Hölder continuous with exponent $1 - \frac{N}{p}$.* ■

We refer the reader to Kesavan [1] for the proof.

Now let $u \in W^{2,p}(\mathbb{R}^N)$. Then $u$ and $\frac{\partial u}{\partial x_i}, 1 \leq i \leq N$ are all in $W^{1,p}(\mathbb{R}^N)$. If $p < N$, then this implies that $u \in W^{1,p^*}(\mathbb{R}^N)$. Now assume that $p^* < N$ as well. This will happen if

$$\frac{1}{N} \;<\; \frac{1}{p^*} \;=\; \frac{1}{p} - \frac{1}{N}$$

or, in other words, when $p < \frac{N}{2}$. In this case $u \in L^{p^{**}}(\mathbb{R}^N)$, where

$$\frac{1}{p^{**}} \;=\; \frac{1}{p^*} - \frac{1}{N} \;=\; \frac{1}{p} - \frac{2}{N}.$$

More genrally, we have the following result.

**Theorem 6.4** *Let $m \geq 1$ be an integer. Let $1 \leq p < \infty$.*
*(i) If $\frac{1}{p} - \frac{m}{N} > 0$, then*

$$W^{m,p}(\mathbb{R}^N) \;\hookrightarrow L^q(\mathbb{R}^N),\; \frac{1}{q} \;=\; \frac{1}{p} - \frac{m}{N}.$$

*(ii) If $\frac{1}{p} - \frac{m}{N} = 0$, then*

$$W^{m,p}(\mathbb{R}^N) \;\hookrightarrow L^q(\mathbb{R}^N),\; q \in [p, \infty).$$

*(iii) If* $\frac{1}{p} - \frac{m}{N} < 0$, *then*

$$W^{m,p}(\mathbb{R}^N) \hookrightarrow L^\infty(\mathbb{R}^N).$$

*In the last case, set* $k$ *to be the integral part and* $\theta$ *to be the fractional part of* $m - \frac{N}{p}$. *Then there exists* $C > 0$ *such that for all* $u \in W^{m,p}(\mathbb{R}^N)$, *we have*

$$|D^\alpha u|_{0,\infty,\mathbb{R}^N} \leq C\|u\|_{m,p,\mathbb{R}^N}, \text{ for all } |\alpha| \leq k$$

*and, for almost all* $x, y \in \mathbb{R}^N$, *and for all* $|\alpha| = k$, *we have*

$$|D^\alpha u(x) - D^\alpha u(y)| \leq C\|u\|_{m,p,\mathbb{R}^N}|x - y|^\theta.$$

*In particular we have the continuous inclusion*

$$W^{m,p}(\mathbb{R}^N) \hookrightarrow \mathcal{C}^k(\mathbb{R}^N)$$

*for* $m > \frac{N}{p}$. *The same results are true when* $\mathbb{R}^N$ *is replaced by* $\mathbb{R}^N_+$ *or by* $\Omega$ *of class* $\mathcal{C}^m$ *with bounded boundary and for the spaces* $W_0^{m,p}(\Omega)$ *for any open subset* $\Omega$ *of* $\mathbb{R}^N$. *Thus, if* $\Omega$ *is bounded and sufficiently smooth, we have, for* $m > \frac{N}{p}$,

$$W^{m,p}(\Omega) \hookrightarrow \mathcal{C}^k(\overline{\Omega}). \blacksquare$$

**Remark 6.1** If $m > \frac{N}{p}$, and if $|\alpha| < k$, where $k$ is the integral part of $m - \frac{N}{p}$, it follows that $D^\alpha u$ is Lipschitz continuous, by virtue of the mean value theorem and the fact that the highest order derivatives are bounded. $\blacksquare$

# 7 Compactness Theorems

In the previous section we saw various continuous inclusions of the Sobolev spaces in the Lebesgue spaces and spaces of smooth functions. We now investigate the compactness of these inclusions.

For example consider the case $p > N$ and $W^{1,p}(\Omega)$ where $\Omega \subset \mathbb{R}^N$ is bounded. Then we saw (cf. Theorem 6.3) that the functions in $W^{1,p}(\Omega)$ are in $\mathcal{C}(\overline{\Omega})$ and that they are also Hölder continuous with exponent $1 - \frac{1}{p}$. Thus, it follows that if $B$ is the unit ball in $W^{1,p}(\Omega)$, then the elements of $B$ are uniformly bounded and that they are equicontinuous as well. Since $\overline{\Omega}$ is compact, it then follows that $B$ is relatively compact in $\mathcal{C}(\overline{\Omega})$, by the Ascoli-Arzela theorem. Consequently the inclusion of $W^{1,p}(\Omega)$ in $\mathcal{C}(\overline{\Omega})$ is compact. This argument works for all the cases covered in Theorem 6.4 as well, when $p > N$.

When $p \leq N$, we have continuous inclusions into the Lebesgue spaces. To examine the compactness of these inclusions, we need an analogue of the Ascoli-Arzela theorem which describes criteria for the relative compactness of subsets in the Lebesgue spaces. This comes from the theorem of Fréchet and Kolmogorov. We omit the proofs and refer the reader to Kesavan [1].

**Theorem 7.1** *(Rellich-Kondrasov) Let $\Omega \subset \mathbb{R}^N$ be a bounded open set. Let $1 \leq p < \infty$. Then, the following inclusions are compact:*
*(i) if $p < N, W^{1,p}(\Omega) \hookrightarrow L^q(\Omega)$, for all $1 \leq q < p^*$;*
*(ii) if $p = N$, $W^{1,N}(\Omega) \hookrightarrow L^q(\Omega)$, for all $1 \leq q < \infty$;*
*(iii) if $p > N$, $W^{1,p}(\Omega) \hookrightarrow \mathcal{C}(\overline{\Omega})$.* ■

**Example 7.1** We do not have compactness if $\Omega$ is unbounded. For example, consider $\Omega = \mathbb{R}$. Let $I = (0,1)$ and let $I_j = (j, j+1)$ for $j \in \mathbb{Z}$. Let $f$ be a $\mathcal{C}^1$ function supported in $I$. Define

$$f_j(x) \;=\; f(x-j),\; j \in \mathbb{Z}.$$

Then all the $f_j$'s are in $W^{1,p}(\mathbb{R})$ for any $1 \leq p < \infty$ and the sequence $\{f_j\}$ is clearly bounded in that space. However, since they all have disjoint supports, we have that if $i \neq j$, then

$$|f_i - f_j|_{0,q,\mathbb{R}} \;=\; 2^{\frac{1}{q}} |f|_{0,q,\mathbb{R}}.$$

Thus the sequence $\{f_j\}$ cannot have a convergent subsequence in any $L^q(\mathbb{R})$ for $1 \leq q < \infty$ and so none of the inclusions

$$W^{1,p}(\mathbb{R}) \;\hookrightarrow\; L^q(\mathbb{R})$$

can be compact. ■

**Example 7.2** When $\Omega \subset \mathbb{R}^N$ is bounded and $p < N$ the inclusion

$$W^{1,p}(\Omega) \hookrightarrow L^{p^*}(\Omega)$$

is not compact. For example, when $p = 2 < N$, assume that this inclusion is compact. Consider the minimization problem: find $u \in H_0^1(\Omega)$ such that

$$J(u) \;=\; \min_{\substack{v \in H_0^1(\Omega) \\ |v|_{0,2^*,\Omega}=1}} J(v)$$

where

$$J(v) \;=\; |v|_{1,\Omega}.$$

Denote the infimum by $m \geq 0$. Let $\{v_n\}$ be a minimizing sequence. Then, it is clearly bounded in $H_0^1(\Omega)$ and so it has a weakly convergent subsequence $\{v_{n_k}\}$ converging (weakly) to, say, $v \in H_0^1(\Omega)$. Since the inclusion

$$H_0^1(\Omega) \;\hookrightarrow\; L^{2^*}(\Omega)$$

is assumed to be compact, it follows that the subsequence converges in norm in $L^{2^*}(\Omega)$ and so we have that $|v|_{0,2^*,\Omega} = 1$. Now $|.|_{1,\Omega}$ is a norm on $H_0^1(\Omega)$ equivalent to the usual norm (Poincaré's inequality) and since the norm is weakly lower semi-continuous, we have that

$$m \;\leq\; |v|_{1,\Omega} \;\leq\; \liminf_{k\to\infty} |v_{n_k}|_{1,\Omega} \;=\; m.$$

Thus $v \in H_0^1(\Omega)$ is a minimizer of $J$ and extending it by zero outside $\Omega$ we see that it achieves the best Sobolev constant (when $p = 2$) in $\mathbb{R}^N$. But we saw in the last section that such minimizers never vanish and so we get a contradiction. Thus, the inclusion of $H_0^1(\Omega)$ in $L^{2^*}(\Omega)$ cannot be compact and so, *a fortiori*, the inclusion of $H^1(\Omega)$ in $L^{2^*}(\Omega)$ cannot be compact either. This is true for all $p < N$ as well. ■

# 8 The spaces $W^{s,p}(\Omega)$

In this section, we will try to define the spaces $W^{s,p}(\Omega)$ when $1 < p < \infty$ and when $s \in \mathbb{R}$ is an arbitrary real number. We start with the case of negative integers.

**Proposition 8.1** *Let $\Omega \subset \mathbb{R}^N$ be an open set and let $1 < p < \infty$. Let $F$ belong to the dual space of $W^{1,p}(\Omega)$ (respectively, $W_0^{1,p}(\Omega)$). Then, there exist $f_0, f_1, \cdots, f_N \in L^{p'}(\Omega)$, where $p'$ is the conjugate exponent of $p$, such that, for all $v \in W^{1,p}(\Omega)$ (respectively, $v \in W_0^{1,p}(\Omega)$), we have*

$$F(v) \;=\; \int_\Omega f_0 v \; dx + \sum_{i=1}^N \int_\Omega f_i \frac{\partial v}{\partial x_i} \; dx.$$

*If we define*

$$\|v\|_{1,p,\Omega} \;=\; |v|_{0,p,\Omega} + \sum_{i=1}^N \left| \frac{\partial v}{\partial x_i} \right|_{0,p,\Omega}$$

*(which is equivalent to the norm defined in Section 2), then*

$$\|F\| \;=\; \max_{0 \le i \le N} |f_i|_{0,p',\Omega}.$$

*If $\Omega$ is bounded and if $F$ is in the dual of $W_0^{1,p}(\Omega)$, then we may take $f_0 = 0$.*

**Proof:** Let $E = (L^p(\Omega))^{N+1}$ and let $T$ be the natural isometry from $W^{1,p}(\Omega)$ (respectively, $W_0^{1,p}(\Omega)$) into $E$ described in Section 2, *i.e.*

$$T(v) \;=\; \left( v, \frac{\partial v}{\partial x_1}, \cdots, \frac{\partial v}{\partial x_N} \right).$$

Let $G \subset E$ denote the image of $T$ and let $S : G \to W^{1,p}(\Omega)$ (respectively, $W_0^{1,p}(\Omega)$) be the inverse mapping of $T$. Consider the functional

$$h \in G \;\mapsto\; F(S(h)).$$

By the Hahn-Banach theorem, we can extend this to a continuous linear functional $\Phi$ on all of $E$, preserving the norm. Then, by the Riesz representation theorem, we can find $f_i \in L^{p'}(\Omega), 0 \le i \le N$ such that, if $v = (v_0, v_1, \cdots, v_N) \in E$, then

$$\Phi(v) \;=\; \sum_{i=0}^N \int_\Omega f_i v_i \; dx.$$

Further, if we define $\|v\|_E = \sum_{i=0}^N |v_i|_{0,p,\Omega}$, then

$$\|\Phi\| \;=\; \max_{0 \le i \le N} |f_i|_{0,p',\Omega}$$

and of course, $\|\Phi\| = \|F\|$ by definition. Now, for any $v \in W^{1,p}(\Omega)$ (respectively, $W_0^{1,p}(\Omega)$), we have $v = S(T(v))$ and so

$$F(v) \;=\; \Phi(T(v)) \;=\; \int_\Omega f_0 v \; dx + \sum_{i=1}^N \int_\Omega f_i \frac{\partial v}{\partial x_i} \; dx.$$

If $\Omega$ is bounded then we can work (thanks to Poincaré's inequality) with the isometry $T : W_0^{1,p}(\Omega) \to (L^p(\Omega))^N$ given by

$$T(u) \;=\; \left(\frac{\partial v}{\partial x_1}, \cdots, \frac{\partial v}{\partial x_N}\right)$$

and so we may take $f_0 = 0$. This completes the proof. ■

Now, if $\varphi \in \mathcal{D}(\Omega)$, then, by definition of the distribution derivative,

$$F(\varphi) \;=\; \int_\Omega f_0 \varphi \; dx - \sum_{i=1}^N < \frac{\partial f}{\partial x_i}, \varphi >$$

where the bracket $< ., . >$ denotes the action of a distribution on an element of $\mathcal{D}(\Omega)$. Since $\mathcal{D}(\Omega)$ is dense in $W_0^{1,p}(\Omega)$, a continuous linear functional on $W_0^{1,p}(\Omega)$ is completely defined by its action on $\mathcal{D}(\Omega)$ and so, in this case, we can identify $F$ with the distribution

$$f_0 - \sum_{i=1}^N \frac{\partial f_i}{\partial x_i}.$$

This identification is not possible for functionals on $W^{1,p}(\Omega)$ since $\mathcal{D}(\Omega)$ is not dense there. Now, if $m > 1$ is any positive integer, any first order derivative of an element in $W^{m,p}(\Omega)$ falls in $W^{m-1,p}(\Omega)$. For consistency, we would like this to be true for $m = 0$ as well. Since derivatives of $L^{p'}(\Omega)$ functions are in the dual of $W_0^{1,p}(\Omega)$, we therefore define this dual space as $W^{-1,p'}(\Omega)$.

**Definition 8.1** *Let $1 < p < \infty$ and let $m \geq 1$ be an integer. Then, the space $W^{-m,p'}(\Omega)$ is the dual of the space $W_0^{m,p}(\Omega)$, where $p'$ is the conjugate exponent of $p$.* ■

Let $1 < p < \infty$ and let $s > 0$. Then the spaces $W^{s,p}(\Omega)$ can be defined in a variety of ways when $s$ is not an integer. Usually interpolation theory in $L^p$ spaces is used. We will not go into these aspects here. The space $W_0^{s,p}(\Omega)$ will be the closure of $\mathcal{D}(\Omega)$ in $W^{s,p}(\Omega)$ and its dual will be the space $W^{-s,p'}(\Omega)$ where $p'$ is the conjugate exponent of $p$.

We will henceforth concentrate on the case $p = 2$. We saw that $H^m(\mathbb{R}^N)$ can be defined via the Fourier transform. We can immediately generalize this.

**Definition 8.2** *Let $s > 0$ be a real number. Then*

$$H^s(\mathbb{R}^N) \;=\; \{u \in L^2(\mathbb{R}^N) \mid (1 + |\xi|^2)^{\frac{s}{2}} \widehat{u}(\xi) \in L^2(\mathbb{R}^N)\}.$$

*The norm in this space is given by*

$$\|u\|_{s,\mathbb{R}^N} \;=\; \left(\int_{\mathbb{R}^N} (1 + |\xi|^2)^s |\widehat{u}(\xi)|^2 \; d\xi\right)^{\frac{1}{2}}.$$

*If $\Omega \subset \mathbb{R}^N$ is a sufficiently smooth open subset, then $H^s(\Omega)$ is the space of restriction of functions in $H^s(\mathbb{R}^N)$ to $\Omega$. If $\Omega$ is a proper subset of $\mathbb{R}^N$, then $H_0^s(\Omega)$ is the closure of $\mathcal{D}(\Omega)$ in $H^s(\Omega)$. The dual of $H_0^s(\Omega)$ (respectively, $H^s(\mathbb{R}^N)$) is the space $H^{-s}(\Omega)$ (respectively, $H^{-s}(\mathbb{R}^N)$).* ■

Let $\Omega \subset \mathbb{R}^N$ be a bounded open set with 'sufficiently smooth' boundary $\partial\Omega$ and is such that $\Omega$ always lies to the same side of each connected component of the boundary. Then, at each point of the boundary, there is a neighbourhood $U$ and a bijective mapping $T : Q \to U$, where $Q$ is the unit cube in $\mathbb{R}^N$ centered at the origin, with the properties given in Definition 4.2. Since $\partial\Omega$ is compact, it can be covered by a finite number $\{U_i\}_{i=1}^m$ of such neighbourhoods. Let $\{T_i\}_{i=1}^m$ be the corresponding maps. Let $\{\psi_i\}_{i=1}^m$ be a $\mathcal{C}^\infty$ partition of unity subordinate to the collection $\{U_i\}_{i=1}^m$.

Being a smooth $(N-1)$-dimensional manifold, $\partial\Omega$ can be provided with the $(N-1)$- dimensional surface measure induced on it from $\mathbb{R}^N$. Thus, we can easily define the spaces $L^p(\partial\Omega)$, for $1 \leq p \leq \infty$. Now, given $u \in L^2(\Omega)$, we can write

$$u \;=\; \sum_{i=1}^{m} \psi_i u.$$

The function $\psi_i u$ is supported inside the open set $U_i$. Consider the functions

$$v_i(y', 0) \;=\; (\psi_i u)(T(y', 0)),\ (y', 0) \in Q_0,\ 1 \leq i \leq m$$

where, as usual, $y \in \mathbb{R}^N$ is written as $(y', y_N)$ with $y' \in \mathbb{R}^{N-1}$. Since $v_i$ is supported in $Q_0$, we can extend it to all of $\mathbb{R}^{N-1}$ by setting it to be zero outside its support. It is easy to see that the maps $u \mapsto v_i$ for $1 \leq i \leq m$ all map $L^2(\partial\Omega)$ into $L^2(\mathbb{R}^{N-1})$ and also maps smooth functions to smooth functions. We now define, for $s > 0$,

$$H^s(\partial\Omega) \;=\; \{u \in L^2(\partial\Omega) \mid v_i \in H^s(\mathbb{R}^{N-1}),\ \text{for all } 1 \leq i \leq m\}.$$

It can be checked that this definition is independent of the choice of the atlas $\{U_i\}_{i=1}^m$ on $\partial\Omega$. We also define

$$H^{-s}(\partial\Omega) \;=\; (H^s(\partial\Omega))^*$$

(where the star denotes the dual space) for $s > 0$.

# 9 Trace theory

Sobolev spaces are the ideal functional analytic setting to study boundary value problems. In that case given $u \in W^{1,p}(\Omega)$, where $\Omega \subset \mathbb{R}^N$ is a bounded open set, we would like to assign meanings expressions like '$u$ restricted to $\partial\Omega$' or 'the outer normal derivative $\frac{\partial u}{\partial \nu}$ of $u$ on $\partial\Omega$' and so on. But when $u \in W^{1,p}(\Omega)$, it is *a priori* an element of $L^p(\Omega)$ and so is only defined almost everywhere. Since $\partial\Omega$ has measure zero in $\mathbb{R}^N$, it is therefore not meaningful to talk of the values $u$ takes on the boundary.

However, since we have additional information on the derivatives of $u$, when $u$ is in some Sobolev space, we can indeed give such notions a meaning consistent with our intuitive understanding of terms such as boundary value and exterior normal derivative. We will now make this more precise.

While the theory outlined below can be done for all $1 < p < \infty$, for simplicity, we will restrict ourselves to the case $p = 2$.

**Theorem 9.1** *There exists a continuous linear map $\gamma_0 : H^1(\mathbb{R}^N_+) \to L^2(\mathbb{R}^{N-1})$ such that, if $v \in H^1(\mathbb{R}^N_+)$ is continuous on $\overline{R^N_+}$, then $\gamma_0(v)$ is the restriction of $v$ to $\partial\mathbb{R}^N_+ = \mathbb{R}^{N-1}$.*

**Proof:** Let $v \in \mathcal{D}(\mathbb{R}^N)$. Then

$$\begin{aligned} |v(x',0)|^2 &= -\int_0^\infty \tfrac{\partial}{\partial x_N}(|v(x',x_N)|^2)\ dx_N \\ &= -2\int_0^\infty \tfrac{\partial v}{\partial x_N}(x',x_N).v(x',x_N)\ dx_N \\ &\leq \int_0^\infty \left( \left| \tfrac{\partial v}{\partial x_N}(x',x_N) \right|^2 + |v(x',x_N)|^2 \right)\ dx_N. \end{aligned}$$

Integrating both sides with respect to $x'$ over $\mathbb{R}^{N-1}$, we get

$$\int_{\mathbb{R}^{N-1}} |v(x',0)|^2\ dx' \ \leq\ \int_{\mathbb{R}^N_+} \left( \left| \frac{\partial v}{\partial x_N} \right|^2 + |v|^2 \right)\ dx.$$

In other words, we have

$$\left| v|_{\mathbb{R}^{N-1}} \right|_{0,\mathbb{R}^{N-1}} \ \leq\ \|v\|_{1,\mathbb{R}^N_+}.$$

Since we know that (cf. Example 4.1 and Corollary 4.1) the restrictions of functions in $\mathcal{D}(\mathbb{R}^N)$ are dense in $H^1(\mathbb{R}^N_+)$, the above inequality implies that the map

$$v \ \mapsto v|_{\mathbb{R}^{N-1}}$$

extends uniquely to a continuous linear map $\gamma_0$ from $H^1(\mathbb{R}^N_+)$ into $L^2(\mathbb{R}^{N-1})$.

Now let $v \in H^1(\mathbb{R}^N_+)$ be continuous on $\overline{\mathbb{R}^N_+}$. Extend $v$ to all of $\mathbb{R}^N$ by reflection on $\mathbb{R}^{N-1}$ (cf. Example 4.1). Choose a sequence $\varepsilon_m \downarrow 0$ and let $\rho_{\varepsilon_m}$ be the corresponding mollifiers. Let $\zeta \in \mathcal{D}(\mathbb{R}^N)$ be such that $0 \leq \zeta \leq 1,\ \zeta \equiv 1$ on $B(0;1)$ and such that $\text{supp}(\zeta) \subset B(0;2)$. Define $\zeta_k(x) = \zeta(x/k)$. Then (cf. Lemma 3.1), $\rho_{\varepsilon_m} * v(x) \to v(x)$ for all $x \in \mathbb{R}^N$ and also, since $\zeta_m \equiv 1$ on $B(0;m)$, we have $v_m(x) \ \to v(x)$ for all $x \in \mathbb{R}^N$, where

$$v_m \ =\ \zeta_m.(\rho_{\varepsilon_m} * v).$$

We also know that (cf. Theorem 3.1) $v_m \ \to v$ in $H^1(\mathbb{R}^N)$ and hence in $H^1(\mathbb{R}^N_+)$ as well. Thus $\gamma_0(v_m) \to \gamma_0(v)$ in $L^2(\mathbb{R}^{N-1})$. But, since $v_m \in \mathcal{D}(\mathbb{R}^N)$, it follows that $\gamma_0(v_m)$ is the restriction of $v_m$ to $\mathbb{R}^{N-1}$ and we saw that this converges pointwise to the restriction of $v$ to $\mathbb{R}^{N-1}$. Thus, it follows that $\gamma_0(v)$ is the restriction of $v$ to $\mathbb{R}^{N-1}$. This completes the proof. ■

It can be shown (cf. Kesavan [1]) that the range of $\gamma_0$ is the space $H^{\frac{1}{2}}(\mathbb{R}^{N-1})$ and that its kernel is $H^1_0(\mathbb{R}^N_+)$. Thus we can interpret elements of $H^1_0(\mathbb{R}^N_+)$ as those of $H^1(\mathbb{R}^N_+)$ which 'vanish on the boundary'.

In the space $H^2(\mathbb{R}^N_+)$, apart from $\gamma_0$, we can imitate the proof of the preceding theorem to show the existence of a map $\gamma_1 : H^2(\mathbb{R}^N_+) \to L^2(\mathbb{R}^{N-1})$ such that if $v \in H^2(\mathbb{R}^N_+) \cap \mathcal{C}^1(\overline{\mathbb{R}^N_+})$, then

$$\gamma_1(v) \ =\ -\frac{\partial v}{\partial x_N}\bigg|_{\mathbb{R}^{N-1}}.$$

Its range would be $H^{\frac{1}{2}}(\mathbb{R}^{N-1})$ while that of $\gamma_0$ will be $H^{\frac{3}{2}}(\partial\Omega)$. The kernel of the map

$$(\gamma_0, \gamma_1) : H^2(\mathbb{R}^N_+) \ \rightarrow \ H^{\frac{3}{2}}(\mathbb{R}^{N-1}) \times H^{\frac{1}{2}}(\mathbb{R}^{N-1})$$

is the space $H^2_0(\mathbb{R}^N_+)$. More, generally, we can define maps $\gamma_j : H^m(\mathbb{R}^N_+) \rightarrow L^2(\mathbb{R}^{N-1})$ which, for a smooth function $v$ is the restriction of $(-\frac{\partial v}{\partial x_N})^j$ to $\mathbb{R}^{N-1}$ and whose range is $H^{m-j-\frac{1}{2}}(\mathbb{R}^{N-1})$ for $1 \leq j \leq m-1$. The kernel of the map

$$(\gamma_0, \gamma_1, \cdots, \gamma_{m-1}) : H^m(\mathbb{R}^N_+) \ \rightarrow \ \Pi_{j=0}^{m-1} H^{m-j-\frac{1}{2}}(\mathbb{R}^{N-1})$$

is $H^m_0(\mathbb{R}^N_+)$.

Let us now turn to the case of a sufficently smooth bounded open set $\Omega$ of $\mathbb{R}^N$. Let $\{U_i\}_{i=1}^m$ together with the associated maps $\{T_i\}_{i=1}^m$ be an atlas for the boundary $\partial\Omega$. Let $\{\psi_i\}_{i=1}^m$ be an associated $\mathcal{C}^\infty$ partition of unity subordinate to the collection $\{U_i\}_{i=1}^m$. If $u \in H^1(\Omega)$, then (after extension by zero) we have that for each $1 \leq i \leq m$, the functions $(\psi_i u|_{U_i \cap \Omega}) \circ T_i \in H^1(\mathbb{R}^N_+)$ and so we can define its trace $\gamma_0$ as an element of $H^{\frac{1}{2}}(\mathbb{R}^{N-1})$. Coming back by $T_i^{-1}$, we can define the trace on $U_i \cap \partial\Omega$. Piecing these together, we get the trace $\gamma_0 u \in L^2(\partial\Omega)$ and, by our definition of the spaces on $\partial\Omega$, the range will be precisely $H^{\frac{1}{2}}(\partial\Omega)$. Similarly, we can define higher order traces, which generalize the notion of exterior normal derivatives of various orders.

**Theorem 9.2** *(Trace Theorem) Let $\Omega \subset \mathbb{R}^N$ be a bounded open set of class $\mathcal{C}^{m+1}$. Then there exist maps $\gamma_0, \gamma_1, \cdots, \gamma_{m-1}$ from $H^m(\Omega)$ into $L^2(\partial\Omega)$ such that*
*(i) if $v \in H^m(\Omega)$ is sufficently smooth, then,*

$$\gamma_0(v) \ = \ v|_{\partial\Omega}, \ \gamma_1(v) \ = \ \left.\frac{\partial v}{\partial \nu}\right|_{\partial\Omega}, \cdots, \gamma_{m-1}(v) \ = \ \left.\frac{\partial^{m-1} v}{\partial \nu^{m-1}}\right|_{\partial\Omega}$$

*where $\nu$ is the unit outward normal on $\partial\Omega$;*
*(ii) The range of the map $(\gamma_0, \gamma_1, \cdots, \gamma_{m-1})$ is*

$$\Pi_{j=0}^{m-1} H^{m-j-\frac{1}{2}}(\partial\Omega);$$

*(iii) The kernel of the map $(\gamma_0, \gamma_1, \cdots, \gamma_{m-1})$ is the space $H^m_0(\Omega)$.* ■

**Theorem 9.3** *(Green's theorem) If $\Omega \subset \mathbb{R}^N$ is a bounded open set of class $\mathcal{C}^1$ and if $u, v \in H^1(\Omega)$, then, for $1 \leq i \leq N$,*

$$\int_\Omega u \frac{\partial v}{\partial x_i} \, dx \ = \ -\int_\Omega \frac{\partial u}{\partial x_i} v \, dx + \int_{\partial\Omega} \gamma_0(u)\gamma_0(v)\nu_i \, d\sigma$$

*where $\nu = (\nu_1, \cdots, \nu_N)$ is the unit outward normal on $\partial\Omega$. In particular, if one of the two functions is in $H^1_0(\Omega)$, we have*

$$\int_\Omega u \frac{\partial v}{\partial x_i} \, dx \ = \ -\int_\Omega \frac{\partial u}{\partial x_i} v \, dx.$$

**Proof:** We know that $\mathcal{C}^\infty(\overline{\Omega})$ is dense in $H^1(\Omega)$. If $\{u_n\}$ and $\{v_n\}$ are sequences of smooth functions converging in $H^1(\Omega)$ to $u$ and $v$ respectively, we have, by the classical Green's theorem,

$$\int_\Omega u_m \frac{\partial v_m}{\partial x_i}\, dx \;=\; -\int_\Omega \frac{\partial u_m}{\partial x_i} v_m\, dx + \int_{\partial\Omega} u_m v_m \nu_i\, d\sigma$$

for any $1 \le i \le N$. The result now follows on passing to the limit as $n \to \infty$. If one of them is in $H_0^1(\Omega)$, then the integrand in the boundary integral vanishes. ■

We conclude this section by showing that the kernel of the trace map $\gamma_0$ is indeed $H_0^1(\Omega)$ in the one-dimensional case.

**Theorem 9.4** *Let $1 < p < \infty$. Let $I = (0,1) \subset \mathbb{R}$. Then*

$$W_0^{1,p}(I) \;=\; \{u \in W^{1,p}(I) \mid u(0) = u(1) = 0\}.$$

**Proof:** Let $u \in W_0^{1,p}(I)$. Then there exists a sequence $\{u_n\}$ in $\mathcal{D}(I)$ such that $u_n \to u$ in $W^{1,p}(I)$. But this implies that (cf. Example 6.2) that $u_n \to u$ uniformly on $\overline{I}$ and so it follows that $u(0) = u(1) = 0$.

Conversely, assume that $u \in W^{1,p}(I)$ (which is therefore continuous on $\overline{I}$) is such that $u(0) = u(1) = 0$. Then

$$u(x) \;=\; \int_0^x u'(t)\, dt \text{ and } \int_0^1 u'(t)\, dt = 0.$$

Since $\mathcal{D}(I)$ is dense in $L^p(I)$, let $\{v_n\}$ be a sequence in $\mathcal{D}(I)$ coverging to $u'$ in $L^p(I)$. Set $a_n = \int_0^1 v_n(t)\, dt$. Since $I$ has finite measure, we also have that $v_n \to u'$ in $L^1(I)$ and so it follows that $a_n \to 0$. Now let $\psi_0 \in \mathcal{D}(I)$ such that $\int_0^1 \psi(t)\, dt = 1$. Define

$$\psi_n(t) \;=\; v_n(t) - a_n \psi_0(t).$$

Then $\psi_n \in \mathcal{D}(I)$ and its integral over $I$ vanishes, by construction. Hence, there exists $u_n \in \mathcal{D}(I)$ such that $u_n' = \psi_n$ (cf. Lemma 1.1).
Now,

$$(u_n - u)(x) \;=\; \int_0^x (u_n' - u')(t)\, dt \;=\; \int_0^x (\psi_n - u')(t)\, dt.$$

Thus

$$|(u_n - u)(x)| \;\le\; |\psi_n - u'|_{0,p,I}$$

which yields

$$\begin{aligned} |u_n - u|_{0,p,I} &\le |\psi_n - u'|_{0,p,I} \\ &\le |\psi_n - v_n|_{0,p,I} + |v_n - u'|_{0,p,I} \\ &= a_n |\psi_0|_{0,p,I} + |v_n - u'|_{0,p,I} \\ &\to 0. \end{aligned}$$

Also

$$|u_n' - u'|_{o,p,I} \;=\; |\psi_n - u'|_{0,p,I} \;\to\; 0$$

as already shown. Thus $\{u_n\}$ is a sequence in $\mathcal{D}(I)$ which converges to $u$ in $W^{1,p}(I)$ and so $u \in W_0^{1,p}(I)$ whih completes the proof. ■

# References

[1] Kesavan, S. **Topics in Functional Analysis and Applications**, New Age International (formerly Wiley-Eastern), 1989.