Introduction to PDE

Spaces of functions

1 Spaces of smooth functions and distributions

Let us recall very rapidly facts about Banach spaces. A Banach space B is a real or complex vector space that is normed and complete. Normed means that it has a norm, i.e. a function

$$\|\cdot\| : B \to \mathbb{R}_+$$

that satisfies

$$\begin{array}{c} \|x+y\| \le \|x\| + \|y\|, \quad \forall\, x, y \in B \\ \|\lambda x\| = |\lambda|\|x\| \quad \forall \lambda \in \mathbb{C},\ x \in B \\ \|x\| = 0 \Rightarrow x = 0, \quad \forall x \in B \end{array}$$

Complete means that all Cauchy sequences converge. The norm makes B into a metric space with translation invariant distance, $d(x,y) = \|x-y\|$. We recall the Baire category theorem that says that all complete metric spaces are of the second category, i.e. cannot be written as a countable union of closed sets with empty interiors. A linear operator

$$T : X \to Y$$

between Banach spaces is continuous iff it is continuous at zero. Continuity is equivalent with the existence of a constant $C > 0$ so that

$$\|Tx\| \le C\|x\|, \quad \forall x \in X.$$

The norm of the operator is

$$\|T\| = \sup_{\|x\|\le 1} \|Tx\|$$

The space $\mathcal{L}(X,Y) := \{T : X \to Y \mid\ T \text{ linear, continuous}\}$ is itself a Banach space. We recall a few basic theorems in Banach space theory:

Theorem 1 *(Banach-Steinhaus, uniform boundedness principle). Let $T_a : X \to Y$ be a family (as $a \in A$, a set) of linear continuous operators between Banach spaces. If for every $x \in X$ there is C_x so that*

$$\|T_a x\| \le C_x$$

then there exists C so that

$$\sup_{a\in A} \|T_a\| \le C$$

Vice-versa: if $\sup_a \|T_a\| = \infty$ then there exists a dense G_δ set such that for every point x in it $\sup_{a\in A} \|T_a x\| = \infty$.

Theorem 2 *(open mapping) Let $T : X \to Y$ be a linear map between Banach spaces. If T is onto then T is open (maps open sets to open stes).*

Theorem 3 *(closed graph) If the graph $G_T \subset X \times Y$, $G_T = \{(x, y) \mid\ y = Tx\}$ of a linear map between Banach spaces is closed, then T is continuous.*

Examples of classical Banach spaces are ℓ_p, the spaces of sequences and the classical Lebesgue spaces $L^p(\Omega)$ with $1 \le p \le \infty$, $\Omega \subset \mathbb{R}^n$ open. The space of continuous functions on a compact is $C(K) = \{f : K \to \mathbb{C} \mid f \text{ continuous}\}$ where $K \subset \mathbb{R}^n$ is compact. The norm is $\|f\| = \sup_{x\in K} |f(x)|$. The Hölder class $C^\alpha(\Omega)$ is the space of bounded contuous functions with norm

$$\|f\|_{C^\alpha} = \sup_{x\in\Omega} |f(x)| + \sup_{x\ne y} \frac{|f(x) - f(y)|}{|x - y|^\alpha}$$

with $0 < \alpha < 1$. When $\alpha = 1$ we have the Lipschitz class. We will describe Sobolev classes shortly.

The dual of a Banach space B is $B' = \mathcal{L}(B, \mathbb{C})$. Classical examples are $L^p(\Omega)' = L^q(\Omega)$, $p^{-1} + q^{-1} = 1$, $1 \le p < \infty$.

Let X be a vector space over the real or complex numbers. We say that

$$p : X \mapsto \mathbb{R}_+$$

is a seminorm, if it obeys

$$p(x + y) \le p(x) + p(y), \quad \forall\ x, y \in X$$

and

$$p(\lambda x) = |\lambda| p(x) \quad \forall\ \lambda \in \mathbb{C}, x \in X.$$

For later use let us recall also the complex Hahn-Banach theorem.

Theorem 4 *(Hahn-Banach) Let $X \subset Y$ be a linear subspace of a complex vector space. We assume that a seminorm p is given on Y, and that a linear map*

$$F : X \to \mathbb{C}$$

is given on X such that

$$|F(x)| \leq p(x)$$

holds on X. Then there exists a linear extension $\widetilde{F}$ of F,

$$\widetilde{F} : Y \to \mathbb{C}$$

satisfying $\widetilde{F}(x) = F(x)$ for all $x \in X$ and

$$|\widetilde{F}(y)| \leq p(y)$$

for all $y \in Y$.

A topological vector space is a vector space that is also a Hausdorff topological space, with continuous operations. The system of neighborhoods at a point is the translate of the system of neighborhoods at zero. A topological vector space is locally convex if zero has a base of neighborhoods that are convex. (Convex sets are sets S such that, together with two points $x, y \in S$ they contain the whole segment $[x, y] \subset S$, where $[x, y] = \{z = (1-t)x + ty \mid \ 0 \leq t \leq 1\}$. A locally convex topological vector space is said to be metrizable if there is a translation-invariant metric that gives the same topology. A family of seminorms is said to be sufficient if for every x there is p in the family so that $p(x) \neq 0$. A sufficient famly of seminorms on a vector space X defines a locally convex topology: it is the coarsest topology such that all seminorms in the family are continuous. A base of neighborhoods is

$$V = \{x \mid p_i(x) < \epsilon_i, i = 1, \ldots m\}$$

with $m \in \mathbb{N}$, $\epsilon_i > 0$ and p_i belonging to the sufficient family. Classical examples of locally convex spaces are $C^m(\Omega)$ and $C^\infty(\Omega)$. Let $K \subset \Omega$ be a compact subset and let $j \leq m$ be an integer. We define

$$p_{K,j}(f) := \sup_{x \in K, \ |\alpha| \leq j} |\partial^\alpha f(x)|$$

where $\partial^\alpha = \frac{\partial^{|\alpha|}}{\partial_1^{\alpha_1}\dots\partial_n^{\alpha_n}}$. Then the topology of $C^m(\Omega)$ is defined by the seminorms $p_{K,j}$. The topology of $C^\infty(\Omega)$ is defined by the same seminorms, but with j arbitrarily large. The topology of $C^m(\Omega)$ is the topology of uniform convergence on compacts, together with derivatives of order up to m. These topologies are metrizable. A metric is obtained as follows. We take a sequence of compacts K_k such that $K_k \subset K_{k+1}$ and $\cup_k K_k = \Omega$. Then, for every k we define

$$d_k(f,g) = \sum_{j=0}^{m} 2^{-j} \frac{p_{K_k,j}(f-g)}{1 + p_{K_k,j}(f-g)}$$

and then set

$$d(f,g) = \sum_{k=0}^{\infty} 2^{-k} \frac{d_k(f,g)}{1 + d_k(f,g)}.$$

A locally convex space is said to be complete if all Cauchy sequences converge. The spaces $C^m(\Omega)$, $0 \le m \le \infty$ are complete. The space $\mathcal{D}(\Omega) = C_0^\infty(\Omega)$ of infinitely differentiable functions with compact support has a topology that is a strict inductive limit. We consider first compacts $K \subset \Omega$. For each such compact we consider $\mathcal{D}_K(\Omega)$, formed with those $C_0^\infty(\Omega)$ functions which have compact support included in K. This is a vector space and $p_{K,j}$ are sufficient seminorms for $j \ge 0$. If $K \subset L$, the spaces are included $\mathcal{D}_K(\Omega) \subset \mathcal{D}_L(\Omega)$. The inclusion is an isomorphism. Then we can take a sequence of compacts K_k as above and identify $\mathcal{D}(\Omega)$ as the set theoretical union of $\mathcal{D}_{K_k}(\Omega)$. The topology on $\mathcal{D}(\Omega)$ is the finest locally convex topology so that all the inclusions $\mathcal{D}_{K_k}(\Omega) \subset \mathcal{D}(\Omega)$ are continuous. A linear map $T : \mathcal{D}(\Omega) \to \mathbb{C}$ is continuous iff its restrictions $T : \mathcal{D}_K \to \mathbb{C}$ are continuous for all $K \subset \Omega$ compact.

Definition 1 *The dual of $\mathcal{D}(\Omega)$,*

$$\mathcal{D}'(\Omega) = \{T : \mathcal{D}(\Omega) \to \mathbb{C} \mid \quad T \text{ linear, continuous}\}$$

is the set of distributions on Ω.

More concretely thus, a linear map u is a distribution, $u \in \mathcal{D}'(\Omega)$, if and only if, for every compact $K \subset \Omega$ there exist a nonnegative integer j and a constant C such that

$$|u(\phi)| \le C p_{K,j}(\phi)$$

holds for all $\phi \in \mathcal{D}_K(\Omega)$. The language of distributions is useful because it provide a generous framework in which problems can be placed and analyzed. It will turn out that locally distributions are derivatives of L^1 functions. Clearly, if $f \in L^1_{loc}(\Omega)$ (i.e. f is integrable on compacts in Ω) then f defines a distribution u by the rule

$$u(\phi) = \int_\Omega f(x)\phi(x)dx$$

The Dirac mass at a point in $x_0 \in \Omega$ is another example of a distribution

$$\delta_{x_0}(\phi) = \phi(x_0).$$

Another example is the principal value integral $P.V.(\frac{1}{x}) \in \mathcal{D}'(\mathbb{R})$ given by

$$P.V.\left(\frac{1}{x}\right)(\phi) = \lim_{\epsilon\to 0,\ \epsilon>0} \int_{|x|>\epsilon} \frac{\phi(x)}{x} dx$$

In order to go further, we recall convolutions, partition of unity and mollifiers. Let $f, g, h \in C_0^\infty(\mathbb{R}^n)$. The convolution

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

has the properties

$$f * g = g * f,$$
$$(f * g) * h = f * (g * h),$$
$$\partial^\alpha (f * g) = (\partial^\alpha f) * g = f * (\partial^\alpha g),$$
$$spt(f * g) \subset spt(f) + spt(g).$$

In the last line we denoted by $spt f$ the support of f, i.e. the closure of the set where f does not vanish, $spt f := \overline{\{x \mid \ f(x) \neq 0\}}$. We also use the notation $A + B$ for the sum of two sets, the set of all possible sums of elements $A + B = \{c \mid \ \exists a \in A,\ \exists b \in B,\ c = a + b\}$. Let ϕ be a smooth positive function $\phi > 0$ in $\mathbb{R}^n$, supported in the ball centered at zero of radius one,

$$spt\phi \subset B(0, 1)$$

with normalized integral $\int_{\mathbb{R}^n} \phi(x)dx = 1$. We define

$$\phi_\epsilon(x) = \epsilon^{-n}\phi\left(\frac{x}{\epsilon}\right)$$

and recall that

$$\lim_{\epsilon \to 0} f * \phi_\epsilon = f$$

holds in a variety of contexts. The convergence is true in L^p for $1 \le p < \infty$. This means that, if $f \in L^p(\mathbb{R}^n)$ then the convergence is in norm

$$\lim_{\epsilon \to 0} \|f - (f * \phi_\epsilon)\|_{L^p} = 0$$

If f is continuous, then the convergence holds uniformly on compacts. Note that even if $f \in L^1_{loc}(\mathbb{R}^n)$, $f * \phi_\epsilon$ is well defined at each $x \in \mathbb{R}^n$ and it is a C^∞ function. We call such a ϕ_ϵ a standard mollifier. We recall a partition of unity statement

Proposition 1 *Let Ω be an open set and assume it is the union of a family of open sets Ω_a, $\Omega = \cup_{a \in A} \Omega_a$. There exists a sequence ψ_j of $C_0^\infty(\mathbb{R}^n)$ functions so that for each j there exists an index a such that $spt\psi_j \subset \Omega_a$. For every compact $K \subset \Omega$ only finitely many occurrences $spt\psi_j \cap K \neq \emptyset$ can happen, and*

$$\sum_j \psi_j(x) = 1$$

holds for every $x \in \Omega$.

We say that ψ_j is a partition of unity subordinated to $(\Omega_a)_{a \in A}$.

Let now $\Omega_1 \subset \Omega$ be two open sets. If $u \in \mathcal{D}'(\Omega)$ it is clear how to define $u_{|\,\Omega_1}$: it is the restriction as a linear map to $\mathcal{D}(\Omega_1)$ which can be identified with a subset of $\mathcal{D}(\Omega)$. We say that two distributions $u_1, u_2 \in \mathcal{D}'(\Omega)$ agree in Ω_1 if $u_{1|\,\Omega_1} = u_{2|\,\Omega_1}$. Two distributions agree locally, if any point has an open neighborhood where the two distributions agree. A partition of unity can be used to show that if two distributions agree locally, they are identical. We consider the space of rapidly decreasing functions a subset $\mathcal{S}(\mathbb{R}^n)$ of $\mathcal{C}^\infty(\mathbb{R}^n)$ formed with functions that decay together with all derivatives faster than any polynomial. The seminorms

$$p_{m,j}(f) = \sup_{x \in \mathbb{R}^n,\, |\alpha| \le j} (1 + |x|)^m |\partial^\alpha f(x)|$$

define a metrizable, complete locally convex topology. It is customary to denote by $\mathcal{E}(\mathbb{R}^n) = C^\infty(\mathbb{R}^n)$ with the locally convex topology introduced earlier. Then the following are continuous inclusions:

$$\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset \mathcal{E}(\mathbb{R}^n)$$

Naturally, the duals (sets of linear continuous maps to the complex or real field) are included in reverse order

$$\mathcal{E}'(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n)$$

We can put weak topologies on these duals, (pointwise convergence) and then the inclusions are continuous. The elements of $\mathcal{E}'(\mathbb{R}^n)$ are distributions with compact support. Indeed, if $u \in \mathcal{E}'(\mathbb{R}^n)$, then there exists a constant C and a seminorm $p_{K,m}$ so that

$$|u(\phi)| \le Cp_{K,m}(\phi)$$

holds for all $\phi \in \mathcal{E}(\mathbb{R}^n)$. This implies that, as a element of $\mathcal{D}'(\mathbb{R}^n)$, $u_{|\mathbb{R}^n\setminus K} = 0$. The complement of the support of a distribution is naturally defined as the largest open set where the distribution vanishes, and so it follows that $spt(u) \subset K$. Conversely, if $u \in \mathcal{D}'(\mathbb{R}^n)$ has compact support K, then it is easy to see that $u(\phi) = u(\phi\chi)$ where χ is a C_0^∞ function identically equal to 1 on a neighborhood of K. Then the map $\phi \mapsto u(\phi\chi)$ extends uniquely to $\phi \in \mathcal{E}(\mathbb{R}^n)$ and defines a continuous linear map there. The distributions in $\mathcal{S}'(\mathbb{R}^n)$ are called temperate distributions.

Distributions can be differentiated any number of times

Definition 2 *Let $u \in \mathcal{D}'(\Omega)$. Let α be a multi-index. Then we define $\partial^\alpha u \in \mathcal{D}'(\Omega)$ by*

$$(\partial^\alpha u)(\phi) = (-1)^{|\alpha|}u(\partial^\alpha \phi)$$

For instance if $H(x)$ is the Heaviside function, $H(x) = 0$ for $x < 0$ and $H(x) = 1$ for $x \ge 0$ then $H' = \delta$. In general, we can generate thus locally distributions by taking derivatives of functions. The next proposition says that this is actually the most general case.

Proposition 2 *Let $\Omega_1 \subset \overline{\Omega_1} \subset \Omega$, with Ω open, Ω_1 open and bounded. Let $u \in \mathcal{D}'(\Omega)$. Then there exist a function $f \in L^\infty(\Omega_1)$ and m so that*

$$u_{|\Omega_1} = \partial_1^m \dots \partial_n^m f$$

holds in the sense of distributions in $\mathcal{D}'(\Omega_1)$.

The idea of the proof is the following. First we prove an inequality, namely that there exists m and a constant C such that

$$|u(\phi)| \le C \int_{\Omega_1} |\partial_1^m \dots \partial_n^m \phi(x)|\, dx$$

holds for all $\phi \in \mathcal{D}(\Omega_1)$. Suppose for a moment we have this inequality. Then we consider the linear subspace X of $L^1(\Omega_1)$ defined by

$$X = \{\psi \mid \exists \phi \in \mathcal{D}(\Omega_1),\ \psi = \partial_1^m \dots \partial_n^m \phi\}$$

On this linear subspace we define the linear functional $F : X \to \mathbb{C}$ by

$$F(\psi) = (-1)^{mn} u(\phi)$$

Note that the inequality which we have not yet proved guarantees two things, first that this is well defined, (meaning that it does not depend on the choice of ϕ) and that

$$|F(\psi)| \le C \|\psi\|_{L^1(\Omega_1)}$$

Then by Hahn-Banach, there exists an extension $\widetilde{F}$ of F to $L^1(\Omega_1)$ that is a continuous linear functional on $L^1(\Omega_1)$. But then, by duality, this implies that there exists a function $f \in L^\infty(\Omega_1)$ so that

$$\widetilde{F}(\psi) = \int_{\Omega_1} f(x)\psi(x)dx$$

holds for all $\psi \in L^1(\Omega_1)$. This shows that, for any $\phi \in \mathcal{D}(\Omega_1)$

$$u(\phi) = (-1)^{mn} \int_{\Omega_1} f(x)(\partial_1^m \dots \partial_n^m \phi)(x)dx$$

which finishes the proof, modulo the inequality. Now, the proof of the inequality: Because $K = \overline{\Omega_1}$ is a compact in Ω and $u \in \mathcal{D}(\Omega)$ there exists an integer j and a constant C_1 such that

$$|u(\phi)| \le C_1 \sup_{x \in \mathbb{R}^n, |\alpha| \le j} |\partial^\alpha \phi(x)|$$

Let $m = j + 1$ and write each derivative as a multiple integral

$$\partial^\alpha \phi(x) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} \partial_1^m \dots \partial_n^m \phi(y)dy$$

where the integral in the first direction is the integral of the integral, $m - \alpha_1$ times and the last integral is the integral of the integral $m - \alpha_n$ times. The inequality follows because the variables belong to the projections of Ω_1 on the coordinate axes and those are included in finite intervals.

2 Sobolev spaces

We define the norm

$$\|f\|_{m,p} = \sum_{|\alpha|\le m} \|\partial^\alpha f\|_{L^p(\Omega)}$$

and the space $W^{m,p}(\Omega)$ to be the set

$$W^{m,p}(\Omega) = \{f \in L^p(\Omega) \mid \partial^\alpha f \in L^p(\Omega),\ |\alpha| \le m\}$$

The derivatives are taken in the sense of distributions. The space endowed with the norm $\|\cdot\|_{m,p}$ is a Banach space. Here $m \ge 0$ and $1 \le p \le \infty$. Let us prove the fact that $W^{m,p}(\Omega)$ is complete. Let f_j be a Cauchy sequence. Because $L^p(\Omega)$ is complete, it follows that for every α, $|\alpha| \le m$, there exist functions $f^{(\alpha)} \in L^p(\Omega)$ so that $\partial^\alpha f_j$ converge in the norm of $L^p(\Omega)$ to $f^{(\alpha)}$. It remains to show that $f^{(\alpha)} = \partial^\alpha f^{(0)}$ in $\mathcal{D}'(\Omega)$. This is an easy consequence of the definitions:

$$\begin{gathered}\int f^{(\alpha)}\phi dx = \lim_{j\to\infty}\int(\partial^\alpha f_j)\phi dx \\ = (-1)^{|\alpha|}\lim_{j\to\infty}\int f_j(\partial^\alpha\phi)dx = (-1)^{|\alpha|}\int f^{(0)}(\partial^\alpha\phi)dx \\ = \int(\partial^\alpha f^{(0)})\phi dx\end{gathered}$$

holds for any $\phi \in C_0^\infty(\Omega)$.

Proposition 3 *Let $1 \le p < \infty$, $m \ge 0$. Let $\Omega \subset \mathbb{R}^n$ be open. Then $C^\infty(\Omega) \cap W^{m,p}(\Omega)$ is dense in $W^{m,p}(\Omega)$.*

For the proof, let us take a sequence of open, relatively compact sets Ω_j so that $\overline{\Omega_j} \subset \Omega_{j+1}$ and $\cup\Omega_j = \Omega$. Let ψ_j be a partition of unity subordinated to the family Ω_j. Let $u \in W^{m,p}(\Omega)$. Note that $\psi_j u \in W^{m,p}(\Omega)$ and $u = \sum_j \psi_j u$. Note also that $spt(\psi_j u) \subset \Omega_j$ is compact, so by choosing δ_j small enough we have that

$$g_j := (\psi_j u) * \phi_{\delta_j} \in C_0^\infty(\Omega_j)$$

(Here ϕ is a standard mollifier.) By choosing δ_j small enough we can make sure that

$$\|g_j - \psi_j u\|_{m,p} \le \epsilon 2^{-j-1}.$$

Now $g = \sum_j g_j$ belongs to $C^\infty(\Omega)$. Indeed, the sum is locally finite. (Taking $\delta_j \le 2^{-j}$, if $x \in \Omega$ and if $spt(\psi_j) \cap B(x, 2r) = \emptyset$ for $j \ge N(r)$ then, taking M so large that $2^{-M} \le r$ and taking $j \ge N, M$ we deduce that $spt(g_j) \cap B(x, r) = \emptyset$.

The main point is that we do not need to change the choice of δ_j according to where x is, just the number of intersecting sets depends on that.) Now

$$\|g-u\|_{m,p} \le \sum_j \|g_j - \psi_j u\|_{m,p} \le \epsilon$$

The fact that $g \in W^{m,p}(\Omega)$ follows from this inequality. This concludes the proof. The closure of $\mathcal{D}(\Omega)$ in $W^{m,p}(\Omega)$ is denoted by $W_0^{m,p}(\Omega)$. It is customary to denote by $H^m(\Omega) = W^{m,2}(\Omega)$. Note that $W_0^{m,p}(\mathbb{R}^n) = W^{m,p}(\mathbb{R}^n)$. In general however the closure of smooth functions with compact support is smaller than $W^{m,p}(\Omega)$.

The Sobolev embedding theorems are our next topic. We start with the Gagliardo-Nirenberg-Sobolev inequality

Theorem 5 *Let $1 \le p < n$. Let*

$$p* = \frac{np}{n-p}$$

There exists a constant C such that

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

holds for all $f \in W^{1,p}(\mathbb{R}^n)$.

The proof goes like this. First we consider the case $p = 1$. WLOG we can take $f \in C_0^\infty(\mathbb{R}^n)$. We write

$$f(x) = \int_{-\infty}^{x_i} \partial_i f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots x_n)dt$$

So,

$$|f(x)| \le \int_{-\infty}^{x_i} |\nabla f|(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots x_n)dt$$

We do this for all $i = 1, \dots n$:

$$|f(x)|^{\frac{n}{n-1}} \le \left[\Pi_{i=1}^n \int_{-\infty}^{\infty} |\nabla f|(x_1, \dots, t_i, \dots, x_n)dt_i\right]^{\frac{1}{n-1}}$$

We integrate this with respect to x_1. One factor in the product does not depend on x_1:

$$\int_{-\infty}^{\infty} |f(x)|^{\frac{n}{n-1}} dx_1 \leq$$
$$\left[\int_{-\infty}^{\infty} |\nabla f|(t_1, x_2, \dots, x_n) dt_1\right]^{\frac{1}{n-1}} \int_{-\infty}^{\infty} dx_1 \left(\Pi_{i=2}^{n} \int_{-\infty}^{\infty} |\nabla f| dt_i\right)^{\frac{1}{n-1}}$$
$$\leq \left[\int_{-\infty}^{\infty} |\nabla f|(t_1, x_2, \dots, x_n) dt_1\right]^{\frac{1}{n-1}} \left(\Pi_{i=2}^{n} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\nabla f| dt_1 dt_i\right)^{\frac{1}{n-1}}$$

In the last piece we used Hölder's inequality with $n-1$ factors. Now we integrate dx_2. We obtain the result by induction:

$$\int |f|^{\frac{n}{n-1}} dx_1 \dots dx_k \leq$$
$$\left[\int |\nabla f|(t_1, \dots, t_k, x_{k+1}, \dots x_n) dt_1 \dots dt_k\right]^{\frac{k}{n-1}} \times \tag{1}$$
$$\left[\Pi_{i=k+1}^{n} \int |\nabla f|(t_1, \dots t_k, x_{k+1}, \dots, t_i, \dots x_n) dt_1 \dots dt_k dt_i\right]^{\frac{1}{n-1}}$$

For general p, we apply the result for $p=1$ to an appropriately chosen $|f|^q$.

$$\left[\int |f|^{\frac{qn}{n-1}} dx\right]^{\frac{n-1}{n}} \leq C \int |f|^{q-1} |\nabla f|$$

We use Hölder and note that if $\frac{qn}{n-1} = p*$ then $(q-1)p' = p*$ where $p' = \frac{p}{p-1}$ is the conjugate exponent of p.

The general statement is

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

for $mp < n$, $p \leq q \leq \frac{np}{n-mp}$. These inclusions hold for Ω bounded with smooth boundary. (C^1). If $mp > n$ then $W^{m,p}(\Omega) \subset C^{[m-\frac{n}{p}],\alpha}(\overline{\Omega})$ where $\alpha = m - \frac{n}{p} - [m - \frac{n}{p}] > 0$. If Ω is bounded with C^r boundary, there exists an extension operator

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

such that $(Eu)(x) = u(x)$ a.e. in Ω,

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

We will give some proofs. For instance in the case $\Omega = \mathbb{R}^n$ and $p=2$ it is easy to see that $u \in H^m$ if and only if

$$(1 + |\xi|^2)^{\frac{m}{2}} \widehat{u}(\xi) \in L^2(\mathbb{R}^n)$$

where $\widehat{u}$ is the Fourier transform. Then, if $k < m - \frac{n}{2}$, it is also easy to see that

$$\|(1+|\xi|^2)^{\frac{k}{2}}\widehat{u}(\xi)\|_{L^1(\mathbb{R}^n)} \le C\|u\|_{W^{m,2}(\mathbb{R}^n)}$$

Therefore it follows that $\partial^\alpha u$ is bounded and uniformly continuous if $|\alpha| < m - \frac{n}{2}$. The extension theorem is done by looking locally near the boundary. if the boundary is flat $\Sigma = \{x \mid \ x_n = 0\}$ and if u has compact support in $x_n \ge 0$, then the formula

$$(Eu)(x) = \left\{ \begin{array}{c} u(x) \quad \text{for } x_n > 0 \\ \sum_{j=1}^{m+1} \lambda_j u(x', -jx_n), \quad \text{for } x_n < 0 \end{array} \right.$$

defines the extension if λ_j are chosen to satisfy

$$\sum_{j=1}^{m+1} j^k \lambda_j = (-1)^k, \quad k = 0, 1, \dots m.$$

Because of the Vandermonde determinant

$$Det(j^k)_{j=1,\dots,m+1,\, k=0,\dots m} = \Pi_{1\le i<j\le m+1}(j-i)$$

we can find λ_j. Note that if $u \in C^m(\overline{\mathbb{R}^n_+})$ then $Eu \in C^m(\mathbb{R}^n)$. This gives the extension, by checking the norms and using the density. This also shows that if the boundary of the domain is smooth then we can approximate functions in $W^{m,p}(\Omega)$ by functions that are in $C^\infty(\mathbb{R}^n)$ restricted to Ω. Indeed, by a partition of unity, it is enough to consider functions compactly supported near the boundary of Ω. By a change of variables and a translation we reduce the problem to a half-space. If u is supported there, we can first approximate it by a smooth function f, compactly supported in $x_n \ge 0$. Then we consider $\widetilde{f_\epsilon}(x) = f(x', x_n + \epsilon)$. This dips f below the boundary. We have that $\widetilde{f_\epsilon}$ is close to f in $W^{m,p}(\Omega)$ and $\widetilde{f_\epsilon} \in C^\infty(\overline{\Omega})$. Convolution with a mollifier finishes the job.

Lemma 1 *Let $1 \le p < \infty$. There exists $C > 0$ so that for every $f \in C^\infty(\mathbb{R}^n)$, $x \in \mathbb{R}^n$, $r > 0$, $z \in B(x,r)$ we have*

$$\int_{B(x,r)} |f(y) - f(z)|^p dy \le Cr^{n+p-1} \int_{B(x,r)} |y-z|^{1-n} |\nabla f(y)|^p dy$$

The proof goes like this. First we note that

$$\begin{gathered}\int_{B(x,r)} |f(y)-f(z)|^p dy = \int_{B(x,r)\cap B(z,2r)} |f(y)-f(z)|^p dy \\ = \int_{B(z,2r)} |f(y)-f(z)|^p \chi_{B(x,r)}(y) dy \\ = \int_0^{2r} d\rho \int_{|y-z|=\rho} |f(y)-f(z)| \chi_{B(x,r)}(y) dS_\rho(y) \\ = \int_0^{2r} d\rho \int_{|y-z|=\rho,\, y\in B(x,r)} |f(y)-f(z)|^p dS_\rho(y)\end{gathered}$$

Now, because

$$f(y)-f(z) = \int_0^1 (y-z)\cdot\nabla f(z+t(y-z))dt$$

it follows that

$$|f(y)-f(z)|^p \leq |y-z|^p \int_0^1 |\nabla f(z+t(y-z)|^p dt$$

and thus

$$\begin{gathered}\int_{|y-z|=\rho,\, y\in B(x,r)} |f(y)-f(z)|^p dS_\rho(y) \\ \leq \rho^p \int_0^1 dt \int_{|y-z|=\rho,\, y\in B(x,r)} |\nabla f(z+t(y-z)|^p dS_\rho(y)\end{gathered}$$

Changing variables to $w = z+t(y-z) \in B(x,r)$ we have

$$\begin{gathered}\int_{|y-z|=\rho,\, y\in B(x,r)} |f(y)-f(z)|^p dS_\rho(y) \\ \leq \rho^p \int_0^1 t^{1-n} dt \int_{|w-z|=t\rho,\, w\in B(x,r)} |\nabla f(w)|^p dS_{t\rho}(w)\end{gathered}$$

and then setting $t\rho = \rho'$

$$\begin{gathered}\int_{|y-z|=\rho,\, y\in B(x,r)} |f(y)-f(z)|^p dS_\rho(y) \\ \leq \rho^{n+p-2} \int_0^\rho d\rho' \int_{|w-z|=\rho',\, w\in B(x,r)} |w-z|^{1-n} |\nabla f(w)|^p dS_{\rho'}(w) \\ = \rho^{n+p-2} \int_{B(x,r)\cap B(z,\rho)} |w-z|^{1-n} |\nabla f(w)|^p dw\end{gathered}$$

Returning to the begining of the proof,

$$\begin{gathered}\int_{B(x,r)} |f(y)-f(z)|^p dy \leq \int_0^{2r} \rho^{n+p-2} d\rho \int_{B(x,r)\cap B(z,\rho)} |w-z|^{1-n} |\nabla f(w)|^p dw \\ \leq \tfrac{(2r)^{n+p-1}}{n+p-1} \int_{B(x,r)} |w-z|^{1-n} |\nabla f(w)|^p dw\end{gathered}$$

Now we can prove the Morrey inequality:

Theorem 6 *Let $n < p < \infty$. there exists a constant C so that for all $r > 0$,$x \in \mathbb{R}^n$, $y, z \in B(x,r)$ we have*

$$|f(y) - f(z)| \leq Cr \left(\frac{1}{\omega_n r^n} \int_{B(x,r)} |\nabla f|^p \right)^{\frac{1}{p}}$$

For the proof, we apply Lemma (1) with $p = 1$. We compare $f(y)$ and $f(z)$ to the average

$$\overline{f}(x,r) = \frac{1}{\omega_n r^n} \int_{B(x,r)} f(z)dz$$

So,

$$\begin{array}{c} \left|f(y) - \overline{f}(x,r)\right| \leq \frac{1}{\omega_n r^n} \int_{B(x,r)} |f(y) - f(z)|dz \\ \leq C \int_{B(x,r)} |\nabla f(w)||w-y|^{1-n} dw \\ \leq \left(\int_{B(x,r)} |\nabla f|^p \right)^{\frac{1}{p}} \left(\int_{B(x,r)} |w-y|^{\frac{p(1-n)}{p-1}} dw \right)^{\frac{p-1}{p}} \end{array}$$

It is clear that

$$\left(\int_{B(x,r)} |w-y|^{\frac{p(1-n)}{p-1}} dw \right)^{\frac{p-1}{p}} \leq \left(\int_{B(y,2r)} |w-y|^{\frac{p(1-n)}{p-1}} dw \right)^{\frac{p-1}{p}} = Cr^{1-\frac{n}{p}}$$

and that finishes the proof.

Lemma 2 *(Local Poincaré). Let $1 \leq p < n$. There exists a constant so that*

$$\left(\frac{1}{\omega_n r^n} \int_{B(x,r)} |f(y) - \overline{f}(x,r)|^{p*} dy \right)^{\frac{1}{p*}} \leq Cr \left(\frac{1}{\omega_n r^n} \int_{B(x,r)} |\nabla f(y)|^p dy \right)^{\frac{1}{p}}$$

The proof uses first the Lemma 1.

$$\begin{array}{c} \frac{1}{\omega_n r^n} \int_{B(x,r)} |f - \overline{f}|^p dy \leq \left(\frac{1}{\omega_n r^n} \right)^2 \int_{B(x,r)} dy \int_{B(x,r)} |f(y) - f(w)|^p dw \\ \leq \frac{C}{\omega_n r^n} r^{p-1} \int_{B(x,r)} dy \int_{B(x,r)} |\nabla f(w)|^p |w-y|^{1-n} dw \\ \leq Cr^p \frac{1}{\omega_n r^n} \int_{B(x,r)} |\nabla f|^p dw \end{array}$$

So, we have that

$$\left(\frac{1}{\omega_n r^n} \int_{B(x,r)} |f(y) - \overline{f}(x,r)|^p dy \right)^{\frac{1}{p}} \leq Cr \left(\frac{1}{\omega_n r^n} \int_{B(x,r)} |\nabla f(y)|^p dy \right)^{\frac{1}{p}}$$

On the other hand, the inequality

$$\left(\frac{1}{\omega_n r^n}\int_{B(x,r)}|g|^{p*}\right)^{\frac{1}{p*}} \leq C\left[r^p\frac{1}{\omega_n r^n}\int_{B(x,r)}|\nabla g|^p + \frac{1}{\omega_n r^n}\int_{B(x,r)}|g|^p\right]^{\frac{1}{p}}$$

valid for all g, follows by rescaling from the same inequality for $r = 1$. This follows by the extension theorem and the Gagliardo-Nirenberg-Sobolev inequality. Indeed, let $\overline{g} \in W^{1,p}(\mathbb{R}^n)$ extend $g \in W^{1,p}(B(0,1))$, with $\|\overline{g}\|_{1,p} \leq C\|g\|_{1,p}$. Then

$$\|g\|_{L^{p*}(B(0,1))} \leq \|\overline{g}\|_{L^{p*}(\mathbb{R}^n)} \leq C\|\nabla\overline{g}\|_{L^p(\mathbb{R}^n)} \leq C\|g\|_{W^{1,p}(B(0,1))}$$