JUHA KINNUNEN

Sobolev Spaces

Department of Mathematics, Aalto University
2025

Contents

1 Definitions and approximations

In this chapter we begin our study of Sobolev spaces. The Sobolev space is a vector space of functions that have weak derivatives. Motivation for studying these spaces is that solutions of partial differential equations, when they exist, belong naturally to Sobolev spaces.

1.1 Weak derivatives

Notation. Let $\Omega \subset \mathbb{R}^n$ be open, $f : \Omega \to \mathbb{R}$ and $k = 1, 2, \dots$. Then we use the following notations:

$$
\begin{aligned}
C(\Omega) &= \{f : f \text{ continuous in } \Omega\} \\
\operatorname{supp} f &= \overline{\{x \in \Omega : f(x) \neq 0\}} = \text{ the support of } f \\
C_0(\Omega) &= \{f \in C(\Omega) : \operatorname{supp} f \text{ is a compact subset of } \Omega\} \\
C^k(\Omega) &= \{f \in C(\Omega) : f \text{ is } k \text{ times continuously diferentiable}\} \\
C_0^k(\Omega) &= C^k(\Omega) \cap C_0(\Omega) \\
C^\infty &= \bigcap_{k=1}^{\infty} C^k(\Omega) = \text{ smooth functions} \\
C_0^\infty(\Omega) &= C^\infty(\Omega) \cap C_0(\Omega) \\
&= \text{ compactly supported smooth functions} \\
&= \text{ test functions}
\end{aligned}
$$

■

W A R N I N G : In general, $\operatorname{supp} f \not\subseteq \Omega$.

Examples 1.1:

(1) Let $u : B(0,1) \to \mathbb{R}$, $u(x) = 1 - |x|$. Then $\operatorname{supp} u = \overline{B(0,1)}$.

(2) Let $f:\mathbb{R}\to\mathbb{R}$ be

$$f(x)=\begin{cases} x^2, & x\geqslant 0,\\ -x^2, & x<0.\end{cases}$$

Now $f\in C^1(\mathbb{R})\setminus C^2(\mathbb{R})$ although the graph looks smooth.

(3) Let us define $\varphi:\mathbb{R}^n\to\mathbb{R}$,

$$\varphi(x)=\begin{cases} e^{\frac{1}{|x|^2-1}}, & x\in B(0,1),\\ 0, & x\in\mathbb{R}^n\setminus B(0,1).\end{cases}$$

Now $\varphi\in C_0^\infty(\mathbb{R}^n)$ and $\operatorname{supp}\varphi=\overline{B(0,1)}$ (exercise).

Let us start with a motivation for definition of weak derivatives. Let $\Omega\subset\mathbb{R}^n$ be open, $u\in C^1(\Omega)$ and $\varphi\in C_0^\infty(\Omega)$. Integration by parts gives

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

There is no boundary term, since φ has a compact support in Ω and thus vanishes near $\partial\Omega$.

Let then $u\in C^k(\Omega)$, $k=1,2,\dots$, and let $\alpha=(\alpha_1,\alpha_2,\dots,\alpha_n)\in\mathbb{N}_0^n$ be a multi-index such that the order of the multi-index $|\alpha|=\alpha_1+\ldots+\alpha_n$ is at most k. We denote

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

THE MORAL: A coordinate of a multi-index indicates how many times a function is differentiated with respect to the corresponding variable. The order of a multi-index tells the total number of differentiations.

Successive integration by parts gives

$$\int_\Omega uD^\alpha\varphi\,dx=(-1)^{|\alpha|}\int_\Omega D^\alpha u\varphi\,dx.$$

Notice that the left-hand side makes sense even under the assumption $u\in L^1_{\mathrm{loc}}(\Omega)$.

Definition 1.2. Assume that $u\in L^1_{\mathrm{loc}}(\Omega)$ and let $\alpha\in\mathbb{N}^n$ be a multi-index. Then $v\in L^1_{\mathrm{loc}}(\Omega)$ is the αth weak partial derivative of u, written $D^\alpha u=v$, if

$$\int_\Omega uD^\alpha\varphi\,dx=(-1)^{|\alpha|}\int_\Omega v\varphi\,dx$$

for every test function $\varphi\in C_0^\infty(\Omega)$. We denote $D^0u=D^{(0,\dots,0)}=u$. If $|\alpha|=1$, then

$$D_ju=\frac{\partial u}{\partial x_j}=D^{(0,\dots,1,\dots,0)}u,\qquad j=1,\dots,n,$$

(the jth component is 1) and

$$Du=(D_1u,D_2u\dots,D_nu)$$

is the weak gradient of u. Here

THE MORAL: Classical derivatives are defined as pointwise limits of difference quotients, but the weak derivatives are defined as functions satisfying the integration by parts formula. Observe, that changing the function on a set of measure zero does not affect its weak derivatives.

WARNING: We use the same notation for the weak and classical derivatives. It should be clear from the context which interpretation is used.

Remarks 1.3:

(1) If $u \in C^k(\Omega)$, then the classical partial derivatives up to order k are also the corresponding weak derivatives of u. In this sense, weak derivatives generalize classical derivatives.

(2) If $u = 0$ almost everywhere in an open set, then $D^\alpha u = 0$ almost everywhere in the same set.

(3) Let Ω' be an open subset of Ω and assume that u has a weak partial derivative $D^\alpha u$ in Ω. Then $D^\alpha u$ is the weak partial derivative of u in Ω'.

(4) Being a weak derivative is a local property in the following sense: if for every point $x \in \Omega$ there exists an open ball $B(x,r_x) \subset \Omega$, $r_x > 0$, so that u has a weak derivative $D^\alpha u$ in $B(x,r_x)$, then $D^\alpha u$ is the weak derivative u in Ω (exercise).

Lemma 1.4. A weak αth partial derivative of u, if it exists, is uniquely defined up to a set of measure zero.

Proof. Assume that $v, \tilde{v} \in L^1_{\text{loc}}(\Omega)$ are both weak αth partial derivatives of u, that is,

$$\int_\Omega u D^\alpha \varphi \, dx = (-1)^{|\alpha|} \int_\Omega v\varphi \, dx = (-1)^{|\alpha|} \int_\Omega \tilde{v}\varphi \, dx$$

for every $\varphi \in C_0^\infty(\Omega)$. This implies that

$$\int_\Omega (v - \tilde{v})\varphi \, dx = 0 \quad \text{for every} \quad \varphi \in C_0^\infty(\Omega). \tag{1.5}$$

Claim: $v = \tilde{v}$ almost everywhere in Ω.

Reason. Let $\Omega' \Subset \Omega$ (i.e. Ω' is open and $\overline{\Omega'}$ is a compact subset of Ω). The space $C_0^\infty(\Omega')$ is dense in $L^p(\Omega')$ (we shall return to this later). There exists a sequence of functions $\varphi_i \in C_0^\infty(\Omega')$ such that $|\varphi_i| \leqslant 2$ in Ω' and $\varphi_i \to \operatorname{sgn}(v - \tilde{v})$ almost everywhere in Ω' as $i \to \infty$. Here sgn is the signum function.

Identity (1.5) and the dominated convergence theorem, with the majorant $|(v - \tilde{v})\varphi_i| \leqslant 2(|v| + |\tilde{v}|) \in L^1(\Omega')$, give

$$\begin{aligned} 0 &= \lim_{i\to\infty} \int_{\Omega'} (v - \tilde{v})\varphi_i \, dx = \int_{\Omega'} \lim_{i\to\infty} (v - \tilde{v})\varphi_i \, dx \\ &= \int_{\Omega'} (v - \tilde{v}) \operatorname{sgn}(v - \tilde{v}) \, dx = \int_{\Omega'} |v - \tilde{v}| \, dx \end{aligned}$$

This implies that $v = \widetilde{v}$ almost everywhere in Ω' for every $\Omega' \Subset \Omega$. Thus $v = \widetilde{v}$ almost everywhere in Ω. ■

From the proof we obtain a very useful corollary.

Corollary 1.6 (Fundamental lemma of the calculus of variations). If $f \in L^1_{\text{loc}}(\Omega)$ satisfies

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

for every $\varphi \in C_0^\infty(\Omega)$, then $f = 0$ almost everywhere in Ω.

THE MORAL: This is an integral way to say that a function is zero almost everywhere.

Example 1.7. Let $n = 1$ and $\Omega = (0,2)$. Consider

$$u(x) = \begin{cases} x, & 0 < x < 1, \\ 1, & 1 \leqslant x < 2, \end{cases}$$

and

$$v(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & 1 \leqslant x < 2. \end{cases}$$

We claim that $u' = v$ in the weak sense. To see this, we show that

$$\int_0^2 u\varphi'\,dx = -\int_0^2 v\varphi\,dx$$

for every $\varphi \in C_0^\infty((0,2))$.

Reason. An integration by parts and the fundamental theorem of calculus give

$$\begin{aligned} \int_0^2 u(x)\varphi'(x)\,dx &= \int_0^1 x\varphi'(x)\,dx + \int_1^2 \varphi'(x)\,dx \\ &= \underbrace{x\varphi(x)\Big|_0^1}_{=\varphi(1)} - \int_0^1 \varphi(x)\,dx + \underbrace{\varphi(2)}_{=0} - \varphi(1) \\ &= -\int_0^1 \varphi(x)\,dx = -\int_0^2 v\varphi(x)\,dx \end{aligned}$$

for every $\varphi \in C_0^\infty((0,2))$. ■

1.2 Sobolev spaces

Definition 1.8. Assume that Ω is an open subset of $\mathbb{R}^n$. The Sobolev space $W^{k,p}(\Omega)$ consists of functions $u \in L^p(\Omega)$ such that for every multi-index α with $|\alpha| \leqslant k$, the weak derivative $D^\alpha u$ exists and $D^\alpha u \in L^p(\Omega)$. Thus

$$W^{k,p}(\Omega) = \{u \in L^p(\Omega) : D^\alpha u \in L^p(\Omega),\ |\alpha| \leqslant k\}.$$

If $u \in W^{k,p}(\Omega)$, we define its norm

$$\|u\|_{W^{k,p}(\Omega)} = \left(\sum_{|\alpha| \leqslant k} \int_\Omega |D^\alpha u|^p \, dx \right)^{\frac{1}{p}}, \quad 1 \leqslant p < \infty,$$

and

$$\|u\|_{W^{k,\infty}(\Omega)} = \sum_{|\alpha| \leqslant k} \|D^\alpha u\|_{L^\infty(\Omega)} = \sum_{|\alpha| \leqslant k} \operatorname*{ess\,sup}_\Omega |D^\alpha u|.$$

Notice that $D^0 u = D^{(0,\dots,0)} u = u$. Assume that Ω' is an open subset of Ω. We say that Ω' is compactly contained in Ω, denoted $\Omega' \Subset \Omega$, if $\overline{\Omega'}$ is a compact subset of Ω. A function $u \in W^{k,p}_{\text{loc}}(\Omega)$, if $u \in W^{k,p}(\Omega')$ for every $\Omega' \Subset \Omega$.

T H E M O R A L : Sobolev space $W^{k,p}(\Omega)$ consists of functions in $L^p(\Omega)$ that have weak partial derivatives up to order k and they belong to $L^p(\Omega)$.

Remarks 1.9:

(1) As in L^p spaces we identify $W^{k,p}$ functions which are equal almost everywhere.

(2) There are several ways to define a norm on $W^{k,p}(\Omega)$. The norm $\|\cdot\|_{W^{k,p}(\Omega)}$ is equivalent, for example, with the norm

$$\sum_{|\alpha| \leqslant k} \|D^\alpha u\|_{L^p(\Omega)}, \quad 1 \leqslant p \leqslant \infty.$$

and $\|\cdot\|_{W^{k,\infty}(\Omega)}$ is also equivalent with

$$\max_{|\alpha| \leqslant k} \|D^\alpha u\|_{L^\infty(\Omega)}.$$

(3) For $k = 1$ we have

$$\|u\|_{W^{1,p}(\Omega)} = \left(\|u\|^p_{L^p(\Omega)} + \sum_{j=1}^n \left\|D_j u\right\|^p_{L^p(\Omega)} \right)^{\frac{1}{p}},$$

and

$$\|u\|_{W^{1,\infty}(\Omega)} = \|u\|_{L^\infty(\Omega)} + \sum_{j=1}^n \|D_j u\|_{L^\infty(\Omega)}.$$

We may also consider equivalent norms

$$\|u\|_{W^{1,p}(\Omega)} = \left(\|u\|^p_{L^p(\Omega)} + \|Du\|^p_{L^p(\Omega)} \right)^{\frac{1}{p}},$$

$$\|u\|_{W^{1,p}(\Omega)} = \|u\|_{L^p(\Omega)} + \sum_{j=1}^n \left\|D_j u\right\|_{L^p(\Omega)},$$

and

$$\|u\|_{W^{1,p}(\Omega)} = \|u\|_{L^p(\Omega)} + \|Du\|_{L^p(\Omega)}$$

when $1 \leqslant p < \infty$. For $p = \infty$, we may consider

$$\|u\|_{W^{1,\infty}(\Omega)} = \|u\|_{L^\infty(\Omega)} + \|Du\|_{L^\infty(\Omega)}$$

and

$$\|u\|_{W^{1,\infty}(\Omega)} = \max\left\{\|u\|_{L^\infty(\Omega)}, \|D_1 u\|_{L^\infty(\Omega)}, \ldots, \|D_n u\|_{L^\infty(\Omega)}\right\}.$$

Here $Du = (D_1 u, \ldots, D_n u)$ is the weak gradient of u,

$$\|Du\|_{L^p(\Omega)} = \left(\int_\Omega |Du|^p \, dx\right)^{\frac{1}{p}}, \quad 1 < p < \infty,$$

and

$$\|Du\|_{L^\infty(\Omega)} = \operatorname*{ess\,sup}_\Omega |Du|$$

with $|Du| = \sqrt{|D_1 u|^2 + \cdots + |D_n u|^2}$.

(4) $u \in W^{k,p}_{\text{loc}}(\Omega)$ if and only if every point has a neighborhood $U_x \subset \Omega$ such that $u \in W^{k,p}(U_x)$, see Remark 1.3 (4).

(5) Let $u \in W^{k,p}_{\text{loc}}(\Omega)$ such that $u \in L^p(\Omega)$ and $D^\alpha u \in L^p(\Omega)$ for every multi-index α with $|\alpha| \leqslant k$. It follows from Remark 1.3 (4) that $u \in W^{k,p}(\Omega)$.

Example 1.10. Let $n \geqslant 2$ and $u : B(0,1) \to [0,\infty]$, $u(x) = |x|^{-\alpha}$, $\alpha > 0$. Clearly $u \in C^\infty(B(0,1) \setminus \{0\})$, but u is unbounded in any neighbourhood of the origin.

We start by showing that u has a weak derivative in the entire unit ball. When $x \neq 0$, we have

$$\frac{\partial u}{\partial x_j}(x) = -\alpha |x|^{-\alpha-1} \frac{x_j}{|x|} = -\alpha \frac{x_j}{|x|^{\alpha+2}}, \quad j = 1, \ldots, n.$$

Thus

$$Du(x) = -\alpha \frac{x}{|x|^{\alpha+2}}.$$

Gauss' theorem gives

$$\int_{B(0,1)\setminus\overline{B(0,\varepsilon)}} D_j(u\varphi)\, dx = \int_{\partial(B(0,1)\setminus\overline{B(0,\varepsilon)})} u\varphi\nu_j \, dS,$$

where $\nu = (\nu_1, \ldots, \nu_n)$ is the outward pointing unit ($|\nu| = 1$) normal of the boundary and $\varphi \in C_0^\infty(B(0,1))$. As $\varphi = 0$ on $\partial B(0,1)$, this can be written as

$$\int_{B(0,1)\setminus\overline{B(0,\varepsilon)}} D_j u\varphi \, dx + \int_{B(0,1)\setminus\overline{B(0,\varepsilon)}} u D_j\varphi \, dx = \int_{\partial B(0,\varepsilon)} u\varphi\nu_j \, dS.$$

By rearranging terms, we obtain

$$\int_{B(0,1)\setminus\overline{B(0,\varepsilon)}} u D_j\varphi \, dx = -\int_{B(0,1)\setminus\overline{B(0,\varepsilon)}} D_j u\varphi \, dx + \int_{\partial B(0,\varepsilon)} u\varphi\nu_j \, dS. \tag{1.11}$$

Let us estimate the last term on the right-hand side. Since $\nu(x) = -\frac{x}{|x|}$, we have $\nu_j(x) = -\frac{x_j}{|x|}$, when $x \in \partial B(0,\varepsilon)$. Thus

$$\begin{aligned}\left|\int_{\partial B(0,\varepsilon)} u\varphi\nu_j \, dS\right| &\leqslant \|\varphi\|_{L^\infty(B(0,1))} \int_{\partial B(0,\varepsilon)} \varepsilon^{-\alpha} \, dS \\ &= \|\varphi\|_{L^\infty(B(0,1))}\omega_{n-1}\varepsilon^{n-1-\alpha} \xrightarrow{\varepsilon\to 0} 0,\end{aligned}$$

if $n-1-\alpha>0$. Here $\omega_{n-1}=\mathscr{H}^{n-1}(\partial B(0,1))$ is the $(n-1)$-dimensional measure of the sphere $\partial B(0,1)$.

Next we study integrability of $D_j u$. We need this information in order to be able to use the dominated convergence theorem. A straightforward computation gives

$$\begin{aligned}
\int_{B(0,1)} |D_j u|\,dx &\leqslant \int_{B(0,1)} |Du|\,dx = \alpha\int_{B(0,1)} |x|^{-\alpha-1}\,dx \\
&= \alpha\int_0^1\int_{\partial B(0,r)} |x|^{-\alpha-1}\,dS\,dr = \alpha\omega_{n-1}\int_0^1 r^{-\alpha-1+n-1}\,dr \\
&= \alpha\omega_{n-1}\int_0^1 r^{n-\alpha-2}\,dr = \frac{\alpha\omega_{n-1}}{n-\alpha-1} r^{n-\alpha-1}\Big|_0^1 < \infty,
\end{aligned}$$

if $n-1-\alpha>0$.

The following argument shows that $D_j u$ is a weak derivative of u also in a neighbourhood of the origin. By the dominated convergence theorem

$$\begin{aligned}
\int_{B(0,1)} u D_j\varphi\,dx &= \int_{B(0,1)} \lim_{\varepsilon\to 0}\Big(u D_j\varphi\chi_{B(0,1)\setminus\overline{B(0,\varepsilon)}}\Big)\,dx \\
&= \lim_{\varepsilon\to 0}\int_{B(0,1)\setminus\overline{B(0,\varepsilon)}} u D_j\varphi\,dx \\
&= -\lim_{\varepsilon\to 0}\int_{B(0,1)\setminus\overline{B(0,\varepsilon)}} D_j u\varphi\,dx + \lim_{\varepsilon\to 0}\int_{\partial B(0,\varepsilon)} u\varphi\nu_j\,dS \\
&= -\int_{B(0,1)} \lim_{\varepsilon\to 0} D_j u\varphi\chi_{B(0,1)\setminus\overline{B(0,\varepsilon)}}\,dx \\
&= -\int_{B(0,1)} D_j u\varphi\,dx.
\end{aligned}$$

Here we used the dominated convergence theorem twice: First to the function

$$u D_j\varphi\chi_{B(0,1)\setminus\overline{B(0,\varepsilon)}},$$

which is dominated by $|u|\|D\varphi\|_\infty \in L^1(B(0,1))$, and then to the function

$$D_j u\varphi\chi_{B(0,1)\setminus\overline{B(0,\varepsilon)}},$$

which is dominated by $|Du|\|\varphi\|_\infty \in L^1(B(0,1))$. We also used (1.11) and the fact that the last term there converges to zero as $\varepsilon\to 0$.

We have proved that u has a weak derivative in the unit ball. Let $1\leqslant p<n$ and $\alpha>0$. We note that $u\in L^p(B(0,1))$ if and only if $-p\alpha+n>0 \iff \alpha<\frac{n}{p}$. On the other hand, $|Du|\in L^p(B(0,1))$, if and only if $-p(\alpha+1)+n>0 \iff \alpha<\frac{n-p}{p}$. Thus $u\in W^{1,p}(B(0,1))$ if and only if $\alpha<\frac{n-p}{p}$.

Let (q_i) be a countable and dense subset of $B(0,1)$ and let $u: B(0,1)\to[0,\infty]$,

$$u(x)=\sum_{i=1}^{\infty}\frac{1}{2^i}|x-q_i|^{-\alpha}.$$

Then $u\in W^{1,p}(B(0,1))$ if $\alpha<\frac{n-p}{p}$.

Reason. We observe that

$$\begin{aligned}
\||x-q_i|^{-\alpha}\|^p_{W^{1,p}(B(0,1))} &= \||x-q_i|^{-\alpha}\|^p_{L^p(B(0,1))} + \|D(|x-q_i|^{-\alpha})\|^p_{L^p(B(0,1))} \\
&= \||x-q_i|^{-\alpha}\|^p_{L^p(B(0,1))} + \|\alpha|x-q_i|^{-\alpha-1}\|^p_{L^p(B(0,1))} \\
&= \int_{B(0,1)} |x-q_i|^{-\alpha p}\,dx + \alpha^p \int_{B(0,1)} |x-q_i|^{-\alpha p-p}\,dx \\
&\leqslant \int_{B(0,2)} |x|^{-\alpha p}\,dx + \alpha^p \int_{B(0,2)} |x|^{-\alpha p-p}\,dx \\
&= \int_{B(0,1)} |2x|^{-\alpha p}2^n\,dx + \alpha^p \int_{B(0,1)} |2x|^{-\alpha p-p}2^n\,dx \\
&= 2^{-\alpha p+n}\int_{B(0,1)} |x|^{-\alpha p}\,dx + 2^{-\alpha p-p+n}\alpha^p \int_{B(0,1)} |x|^{-(\alpha+1)p}\,dx \\
&\leqslant 2^{-\alpha p+n}(1+2^{-p})\||x|^{-\alpha}\|^p_{W^{1,p}(B(0,1))} < \infty, \quad i=1,2,\dots.
\end{aligned}$$

Note that the right-hand side is independent of i. It follows that

$$\begin{aligned}
\|u\|_{W^{1,p}(B(0,1))} &\leqslant \sum_{i=1}^{\infty} \frac{1}{2^i}\||x-q_i|^{-\alpha}\|_{W^{1,p}(B(0,1))} \\
&\leqslant \sum_{i=1}^{\infty} \frac{1}{2^i}\big(2^{-\alpha p+n}(1+2^{-p})\big)^{\frac{1}{p}}\||x|^{-\alpha}\|_{W^{1,p}(B(0,1))} \\
&= \big(2^{-\alpha p+n}(1+2^{-p})\big)^{\frac{1}{p}}\||x|^{-\alpha}\|_{W^{1,p}(B(0,1))} < \infty,
\end{aligned}$$

if $\alpha < \frac{n-p}{p}$. ■

Since $\alpha > 0$, we note that u is unbounded and not differentiable in the classical sense in a dense subset of $B(0,1)$.

THE MORAL: Functions in $W^{1,p}$, $1 \leqslant p < n$, $n \geqslant 2$, may be unbounded in every open subset.

Example 1.12. Observe, that $u(x) = |x|^{-\alpha}$, $\alpha > 0$, does not belong to $W^{1,n}(B(0,1))$. However, there are unbounded functions in $W^{1,n}$, $n \geqslant 2$. Let $u : B(0,1) \to \mathbb{R}$,

$$u(x) = \begin{cases} \log\left(\log\left(1+\frac{1}{|x|}\right)\right), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Then $u \in W^{1,n}(B(0,1))$ when $n \geqslant 2$, but $u \notin L^\infty(B(0,1))$. This can be used to construct a function in $W^{1,n}(B(0,1)$ that is unbounded in every open subset of $B(0,1)$ (exercise).

THE MORAL: Functions in $W^{1,p}$, $1 \leqslant p \leqslant n$, $n \geqslant 2$, are not continuous. Later we shall see, that every $W^{1,p}$ function with $p > n$ coincides with a continuous function almost everywhere.

Example 1.13. The function $u : B(0,1) \to \mathbb{R}$,

$$u(x) = u(x_1, \dots, x_n) = \begin{cases} 1, & x_n \geqslant 0, \\ 0, & x_n < 0, \end{cases}$$

does not belong to $W^{1,p}(B(0,1)$ for any $1 \leqslant p \leqslant \infty$ (exercise).

1.3 Properties of weak derivatives

The following general properties of weak derivatives follow rather directly from the definition.

Lemma 1.14. Assume that $u, v \in W^{k,p}(\Omega)$ and $1 \leqslant |\alpha| \leqslant k$. Then

(1) $D^\alpha u \in W^{k-|\alpha|,p}(\Omega)$,
(2) $D^\beta(D^\alpha u) = D^\alpha(D^\beta u)$ for all multi-indices α, β with $|\alpha| + |\beta| \leqslant k$,
(3) for every $\lambda, \mu \in \mathbb{R}$, $\lambda u + \mu v \in W^{k,p}(\Omega)$ and

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

(4) if $\Omega' \subset \Omega$ is open, then $u \in W^{k,p}(\Omega')$,
(5) (Leibniz's formula) if $\eta \in C_0^\infty(\Omega)$, then $\eta u \in W^{k,p}(\Omega)$ and

$$D^\alpha(\eta u) = \sum_{\beta \leqslant \alpha} \binom{\alpha}{\beta} D^\beta \eta D^{\alpha - \beta} u,$$

where

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!(\alpha - \beta)!}, \quad \alpha! = \alpha_1! \dots \alpha_n!$$

and $\beta \leqslant \alpha$ means that $\beta_j \leqslant \alpha_j$ for every $j = 1, \dots, n$.

T H E M O R A L : Weak derivatives have the same properties as classical derivatives of smooth functions.

Proof. (1) Follows directly from the definition of weak derivatives. See also (2).

(2) Let $\varphi \in C_0^\infty(\Omega)$. Then $D^\beta \varphi \in C_0^\infty(\Omega)$. Therefore

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

for all test functions $\varphi \in C_0^\infty(\Omega)$. Notice that

$$\begin{aligned}|\alpha|+|\alpha+\beta| &= \alpha_1+\ldots+\alpha_n+(\alpha_1+\beta_1)+\ldots+(\alpha_n+\beta_n)\\ &= 2(\alpha_1+\ldots+\alpha_n)+\beta_1+\ldots+\beta_n\\ &= 2|\alpha|+|\beta|.\end{aligned}$$

As $2|\alpha|$ is an even number, the estimate above, together with the uniqueness results Lemma 1.4 and Corollary 1.6, implies that $D^\beta(D^\alpha u) = D^{\alpha+\beta}u$.

(3) and (4) Clear.

(5) The proof is by induction on $|\alpha|$. Let $|\alpha| = 1$ and $\varphi \in C_0^\infty(\Omega)$. Since $D^\alpha(\eta\varphi) = \eta D^\alpha\varphi + \varphi D^\alpha \eta$, we have

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

This shows that $D^\alpha(\eta u) = \eta D^\alpha u + uD^\alpha\eta$.

Then we make the induction assumption. Let $l < k$ and assume that

$$D^\alpha(\eta u) = \sum_{\beta\leqslant\alpha}\binom{\alpha}{\beta} D^\beta\eta D^{\alpha-\beta}u,$$

for every multi-index α with $|\alpha| \leqslant l$ and for every $\eta \in C_0^\infty(\Omega)$.

Let α be a multi-index with $|\alpha| = l+1$. Then $\alpha = \beta+\gamma$ with $|\beta| = l$ and $|\gamma| = 1$. As above, we have

$$\begin{aligned}\int_\Omega \eta u D^\alpha\varphi\,dx = \int_\Omega \eta u D^{\beta+\gamma}\varphi\,dx &= \int_\Omega \eta u D^\beta(D^\gamma u)\varphi\,dx\\ &= (-1)^{|\beta|}\int_\Omega \sum_{\sigma\leqslant\beta}\binom{\beta}{\sigma} D^\sigma\eta D^{\beta-\sigma}u D^\gamma\varphi\,dx.\end{aligned}$$

By the induction assumption on $D^{\beta-\sigma}u \in W^{k-|\beta|+|\sigma|,p}(\Omega)$ and $D^\gamma\varphi \in C_0^\infty(\Omega)$, $|\gamma| = 1$, we have

$$\begin{aligned}&(-1)^{|\beta|}\int_\Omega \sum_{\sigma\leqslant\beta}\binom{\beta}{\sigma} D^\sigma\eta D^{\beta-\sigma}u D^\gamma\varphi\,dx\\ &\qquad = (-1)^{|\beta|+|\gamma|}\int_\Omega \sum_{\sigma\leqslant\beta}\binom{\beta}{\sigma} D^\gamma(D^\sigma\eta D^{\beta-\sigma}u)\varphi\,dx.\end{aligned}$$

By the induction assumption on $D^{\beta-\sigma}u \in W^{k-|\beta|+|\sigma|,p}(\Omega)$ and $D^\sigma\varphi \in C_0^\infty(\Omega)$, we have

$$\begin{aligned}&(-1)^{|\beta|+|\gamma|}\int_\Omega \sum_{\sigma\leqslant\beta}\binom{\beta}{\sigma} D^\gamma(D^\sigma\eta D^{\beta-\sigma}u)\varphi\,dx\\ &= (-1)^{|\alpha|}\int_\Omega \sum_{\sigma\leqslant\beta}\binom{\beta}{\sigma}(D^\rho\eta D^{\alpha-\rho}u + D^\sigma\eta D^{\alpha-\sigma}u)\varphi\,dx,\end{aligned}$$

where $\rho = \sigma + \gamma$ so that $\alpha - \sigma = \beta - \rho$. It follows that

$$
\begin{aligned}
&(-1)^{|\alpha|} \int_\Omega \sum_{\sigma \leqslant \beta} \binom{\beta}{\sigma} (D^\rho \eta D^{\alpha-\rho} u + D^\sigma \eta D^{\alpha-\sigma} u)\varphi \, dx \\
&= (-1)^{|\alpha|} \int_\Omega \left(\sum_{\sigma \leqslant \beta} \binom{\alpha-\gamma}{\sigma} D^\sigma \eta D^{\alpha-\sigma} u + \sum_{\gamma \leqslant \rho \leqslant \alpha} \binom{\alpha-\gamma}{\rho-\gamma} D^\rho \eta D^{\alpha-\rho} u \right) \varphi \, dx \\
&= (-1)^{|\alpha|} \int_\Omega \left(\sum_{\sigma \leqslant \alpha} \binom{\alpha}{\sigma} D^\sigma \eta D^{\alpha-\sigma} u \right) \varphi \, dx,
\end{aligned}
$$

since

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

□

1.4 Completeness of Sobolev spaces

One of the most useful properties of Sobolev spaces is that they are complete. Thus Sobolev spaces are closed under limits of Cauchy sequences.

A sequence (u_i) of functions $u_i \in W^{k,p}(\Omega)$, $i = 1,2,\dots$, converges in $W^{k,p}(\Omega)$ to a function $u \in W^{k,p}(\Omega)$, if for every $\varepsilon > 0$ there exists i_ε such that

$$
\|u_i - u\|_{W^{k,p}(\Omega)} < \varepsilon \quad \text{when} \quad i \geqslant i_\varepsilon.
$$

Equivalently,

$$
\lim_{i\to\infty} \|u_i - u\|_{W^{k,p}(\Omega)} = 0.
$$

A sequence (u_i) is a Cauchy sequence in $W^{k,p}(\Omega)$, if for every $\varepsilon > 0$ there exists i_ε such that

$$
\|u_i - u_j\|_{W^{k,p}(\Omega)} < \varepsilon \quad \text{when} \quad i,j \geqslant i_\varepsilon.
$$

WARNING: This is not the same condition as

$$
\|u_{i+1} - u_i\|_{W^{k,p}(\Omega)} < \varepsilon \quad \text{when} \quad i \geqslant i_\varepsilon.
$$

Indeed, the Cauchy sequence condition implies this, but the converse is not true (exercise).

Theorem 1.15 (Completeness). The Sobolev space $W^{k,p}(\Omega)$, $1 \leqslant p \leqslant \infty$, $k = 1,2,\dots$, is a Banach space.

THE MORAL: The spaces $C^k(\Omega)$, $k = 1,2,\dots$, are not complete with respect to the Sobolev norm, but Sobolev spaces are. This is important in existence arguments for PDEs.

Proof. **Step 1:** $\|\cdot\|_{W^{k,p}(\Omega)}$ is a norm.

Reason. $\boxed{(1)}$ $\|u\|_{W^{k,p}(\Omega)} = 0 \Longleftrightarrow u = 0$ almost everywhere in Ω.
$\boxed{\Longrightarrow}$ $\|u\|_{W^{k,p}(\Omega)} = 0$ implies $\|u\|_{L^p(\Omega)} = 0$, which implies that $u = 0$ almost everywhere in Ω.
$\boxed{\Longleftarrow}$ $u = 0$ almost everywhere in Ω implies

$$\int_\Omega D^\alpha u \varphi \, dx = (-1)^{|\alpha|} \int_\Omega u D^\alpha \varphi \, dx = 0$$

for all $\varphi \in C_0^\infty(\Omega)$. This together with Corollary 1.6 implies that $D^\alpha u = 0$ almost everywhere in Ω for all α, $|\alpha| \leqslant k$.

$\boxed{(2)}$ $\|\lambda u\|_{W^{k,p}(\Omega)} = |\lambda| \|u\|_{W^{k,p}(\Omega)}$, $\lambda \in \mathbb{R}$. Clear.

$\boxed{(3)}$ The triangle inequality for $1 \leqslant p < \infty$ follows from Minkowski's inequality applied first for the Lebesgue measure and then for the counting measure, since

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

■

Step 2: Let (u_i) be a Cauchy sequence in $W^{k,p}(\Omega)$. As

$$\|D^\alpha u_i - D^\alpha u_j\|_{L^p(\Omega)} \leqslant \|u_i - u_j\|_{W^{k,p}(\Omega)}, \quad |\alpha| \leqslant k,$$

it follows that $(D^\alpha u_i)$ is a Cauchy sequence in $L^p(\Omega)$, $|\alpha| \leqslant k$. The completeness of $L^p(\Omega)$ implies that there exists $u_\alpha \in L^p(\Omega)$ such that $D^\alpha u_i \to u_\alpha$ in $L^p(\Omega)$ as $i \to \infty$. In particular, $u_i \to u_{(0,\dots,0)} = u$ in $L^p(\Omega)$ as $i \to \infty$.

Step 3: We show that $D^\alpha u = u_\alpha$, $|\alpha| \leqslant k$. We would like to argue

$$\begin{aligned}
\int_\Omega u D^\alpha \varphi \, dx &= \lim_{i\to\infty} \int_\Omega u_i D^\alpha \varphi \, dx \\
&= \lim_{i\to\infty} (-1)^{|\alpha|} \int_\Omega D^\alpha u_i \varphi \, dx \\
&= (-1)^{|\alpha|} \int_\Omega u_\alpha \varphi \, dx
\end{aligned}$$

for every $\varphi \in C_0^\infty(\Omega)$. On the second line we used the definition of the weak derivative. Next we show how to conclude the first and last equalities above.

$\boxed{1 < p < \infty}$ Let $\varphi \in C_0^\infty(\Omega)$. By Hölder's inequality we have

$$\begin{aligned}
\left| \int_\Omega u_i D^\alpha \varphi \, dx - \int_\Omega u D^\alpha \varphi \, dx \right| &= \left| \int_\Omega (u_i - u) D^\alpha \varphi \, dx \right| \\
&\leqslant \|u_i - u\|_{L^p(\Omega)} \|D^\alpha \varphi\|_{L^{p'}(\Omega)} \xrightarrow{i\to\infty} 0
\end{aligned}$$

and consequently we obtain the first inequality above. The last inequality follows in the same way, since

$$\left|\int_\Omega D^\alpha u_i \varphi\,dx - \int_\Omega u_\alpha \varphi\,dx\right| \leqslant \|D^\alpha u_i - u_\alpha\|_{L^p(\Omega)} \|\varphi\|_{L^{p'}(\Omega)} \xrightarrow{i\to\infty} 0.$$

$\boxed{p=1, p=\infty}$ A similar argument as above (exercise).

This means that the weak derivatives $D^\alpha u$ exist and $D^\alpha u = u_\alpha$, $|\alpha| \leqslant k$. As we also know that $D^\alpha u_i \to u_\alpha = D^\alpha u$, $|\alpha| \leqslant k$, we conclude that $\|u_i - u\|_{W^{k,p}(\Omega)} \to 0$ as $i \to \infty$. Thus $u_i \to u$ in $W^{k,p}(\Omega)$ as $i \to \infty$. □

Remark 1.16. $W^{k,p}(\Omega)$, $1 \leqslant p < \infty$ is separable. In the case $k = 1$ consider the mapping $u \mapsto (u, Du)$ from $W^{1,p}(\Omega)$ to $L^p(\Omega) \times L^p(\Omega)^n$ and recall that a subset of a separable space is separable. However, $W^{1,\infty}(\Omega)$ is not separable (exercise).

1.5 Hilbert space structure

The space $W^{k,2}(\Omega)$ is a Hilbert space with the inner product

$$\langle u, v\rangle_{W^{k,2}(\Omega)} = \sum_{|\alpha|\leqslant k} \langle D^\alpha u, D^\alpha v\rangle_{L^2(\Omega)},$$

where

$$\langle D^\alpha u, D^\alpha v\rangle_{L^2(\Omega)} = \int_\Omega D^\alpha u D^\alpha v\,dx.$$

Observe that

$$\|u\|_{W^{k,2}(\Omega)} = \langle u, u\rangle_{W^{k,2}(\Omega)}^{\frac{1}{2}}.$$

1.6 Approximation by smooth functions

This section deals with the question whether every function in a Sobolev space can be approximated by a smooth function.

Define $\phi \in C_0^\infty(\mathbb{R}^n)$ by

$$\phi(x) = \begin{cases} c\, e^{\frac{1}{|x|^2-1}}, & |x| < 1, \\ 0, & |x| \geqslant 1, \end{cases}$$

where $c > 0$ is chosen so that

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

For $\varepsilon > 0$, set

$$\phi_\varepsilon(x) = \frac{1}{\varepsilon^n}\phi\left(\frac{x}{\varepsilon}\right).$$

The function ϕ is called the standard mollifier or Friedrich's mollifier. Observe that $\phi_\varepsilon \geqslant 0$, $\operatorname{supp}\phi_\varepsilon = \overline{B(0,\varepsilon)}$ and

$$\int_{\mathbb{R}^n} \phi_\varepsilon(x)\,dx = \frac{1}{\varepsilon^n}\int_{\mathbb{R}^n} \phi\left(\frac{x}{\varepsilon}\right)dx = \frac{1}{\varepsilon^n}\int_{\mathbb{R}^n} \phi(y)\varepsilon^n\,dy = \int_{\mathbb{R}^n} \phi(x)\,dx = 1$$

for all $\varepsilon > 0$. Here we used the change of variable $y = \frac{x}{\varepsilon}$, $dx = \varepsilon^n\,dy$.

Notation. If $\Omega \subset \mathbb{R}^n$ is open with $\partial\Omega \neq \emptyset$, we write

$$\Omega_\varepsilon = \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > \varepsilon\}, \quad \varepsilon > 0.$$

If $f \in L^1_{\text{loc}}(\Omega)$, we obtain its standard convolution mollification $f_\varepsilon : \Omega_\varepsilon \to [-\infty, \infty]$,

$$f_\varepsilon(x) = (f * \phi_\varepsilon)(x) = \int_\Omega f(y)\phi_\varepsilon(x-y)\,dy.$$

■

T H E M O R A L : Since the convolution is a weighted integral average of f over the ball $B(x,\varepsilon)$ for every x, instead of Ω it is well defined only in Ω_ε. If $\Omega = \mathbb{R}^n$, we do not have this problem.

Remarks 1.17:

(1) For every $x \in \Omega_\varepsilon$, we have

$$f_\varepsilon(x) = \int_\Omega f(y)\phi_\varepsilon(x-y)\,dy = \int_{B(x,\varepsilon)} f(y)\phi_\varepsilon(x-y)\,dy.$$

(2) By a change of variables $z = x - y$ we have

$$\int_\Omega f(y)\phi_\varepsilon(x-y)\,dy = \int_\Omega f(x-z)\phi_\varepsilon(z)\,dz$$

(3) For every $x \in \Omega_\varepsilon$, we have

$$|f_\varepsilon(x)| \leqslant \left|\int_{B(x,\varepsilon)} f(y)\phi_\varepsilon(x-y)\,dy\right| \leqslant \|\phi_\varepsilon\|_{L^\infty(\mathbb{R}^n)} \int_{B(x,\varepsilon)} |f(y)|\,dy < \infty.$$

(4) If $f \in C_0(\Omega)$, then $f_\varepsilon \in C_0(\Omega_\varepsilon)$, whenever

$$0 < \varepsilon < \varepsilon_0 = \frac{1}{2}\operatorname{dist}(\operatorname{supp} f, \partial\Omega).$$

Reason. If $x \in \Omega_\varepsilon$ such that $\operatorname{dist}(x, \operatorname{supp} f) > \varepsilon_0$ (in particular, for every $x \in \Omega_\varepsilon \setminus \Omega_{\varepsilon_0}$) then $B(x,\varepsilon) \cap \operatorname{supp} f = \emptyset$, which implies that $f_\varepsilon(x) = 0$. ■

Lemma 1.18 (Properties of mollifiers).

(1) $f_\varepsilon \in C^\infty(\Omega_\varepsilon)$.

(2) $f_\varepsilon \to f$ almost everywhere as $\varepsilon \to 0$.

(3) If $f \in C(\Omega)$, then $f_\varepsilon \to f$ uniformly in every $\Omega' \Subset \Omega$.

(4) If $f \in L^p_{\text{loc}}(\Omega)$, $1 \leqslant p < \infty$, then $f_\varepsilon \to f$ in $L^p(\Omega')$ for every $\Omega' \Subset \Omega$.

W A R N I N G : Assertion (4) does not hold for $p = \infty$, since the limit of a uniformly converging sequence of continuous functions is continuous, whereas exist functions in L^∞ that are not continuous.

Proof. (1) Let $x \in \Omega_\varepsilon$, $j = 1,\dots,n$, $e_j = (0,\dots,1,\dots,0)$ (the jth component is 1). Let $h_0 > 0$ such that $B(x,h_0) \subset \Omega_\varepsilon$ and let $h \in \mathbb{R}$, $|h| < h_0$. Then

$$\frac{f_\varepsilon(x+he_j)-f_\varepsilon(x)}{h} = \frac{1}{\varepsilon^n}\int_{B(x+he_j,\varepsilon)\cup B(x,\varepsilon)} \frac{1}{h}\left[\phi\Big(\frac{x+he_j-y}{\varepsilon}\Big)-\phi\Big(\frac{x-y}{\varepsilon}\Big)\right] f(y)\,dy.$$

Let $\Omega' = B(x,h_0+\varepsilon)$. Then $\Omega' \Subset \Omega$ and $B(x+he_j,\varepsilon)\cup B(x,\varepsilon) \subset \Omega'$.

Claim:

$$\frac{1}{h}\left[\phi\Big(\frac{x+he_j-y}{\varepsilon}\Big)-\phi\Big(\frac{x-y}{\varepsilon}\Big)\right] \xrightarrow{h\to 0} \frac{1}{\varepsilon}\frac{\partial\phi}{\partial x_j}\Big(\frac{x-y}{\varepsilon}\Big) \quad \text{for every } y \in \Omega'.$$

Reason. Let $\psi(x) = \phi\big(\frac{x-y}{\varepsilon}\big)$. Then

$$\frac{\partial\psi}{\partial x_j}(x) = \frac{1}{\varepsilon}\frac{\partial\phi}{\partial x_j}\Big(\frac{x-y}{\varepsilon}\Big), \quad j = 1,\dots,n$$

and by the fundamental theorem of calculus, we have

$$\psi(x+he_j)-\psi(x) = \int_0^h \frac{\partial}{\partial t}(\psi(x+te_j))\,dt = \int_0^h D\psi(x+te_j)\cdot e_j\,dt.$$

■

Thus

$$\begin{aligned}|\psi(x+he_j)-\psi(x)| &\leqslant \int_0^{|h|} |D\psi(x+te_j)\cdot e_j|\,dt\\ &\leqslant \frac{1}{\varepsilon}\int_0^{|h|}\left|D\phi\Big(\frac{x+te_j-y}{\varepsilon}\Big)\right|dt\\ &\leqslant \frac{|h|}{\varepsilon}\|D\phi\|_{L^\infty(\mathbb{R}^n)}.\end{aligned}$$

This estimate shows that we can use the Lebesgue dominated convergence theorem (on the third row) to obtain

$$\begin{aligned}\frac{\partial f_\varepsilon}{\partial x_j}(x) &= \lim_{h\to 0}\frac{f_\varepsilon(x+he_j)-f_\varepsilon(x)}{h}\\ &= \lim_{h\to 0}\frac{1}{\varepsilon^n}\int_{\Omega'}\frac{1}{h}\left[\phi\Big(\frac{x+he_j-y}{\varepsilon}\Big)-\phi\Big(\frac{x-y}{\varepsilon}\Big)\right]f(y)\,dy\\ &= \frac{1}{\varepsilon^n}\int_{\Omega'}\frac{1}{\varepsilon}\frac{\partial\phi}{\partial x_j}\Big(\frac{x-y}{\varepsilon}\Big)f(y)\,dy\\ &= \int_{\Omega'}\frac{\partial\phi_\varepsilon}{\partial x_j}(x-y)f(y)\,dy = \Big(\frac{\partial\phi_\varepsilon}{\partial x_j}*f\Big)(x).\end{aligned}$$

A similar argument shows that $D^\alpha f_\varepsilon$ exists and

$$D^\alpha f_\varepsilon = D^\alpha \phi_\varepsilon * f \quad \text{in } \Omega_\varepsilon$$

for every multi-index α.

(2) Recall that $\int_{B(x,\varepsilon)} \phi_\varepsilon(x-y)\,dy = 1$. Therefore we have

$$\begin{aligned}
|f_\varepsilon(x) - f(x)| &= \left| \int_{B(x,\varepsilon)} \phi_\varepsilon(x-y) f(y)\,dy - f(x) \int_{B(x,\varepsilon)} \phi_\varepsilon(x-y)\,dy \right| \\
&= \left| \int_{B(x,\varepsilon)} \phi_\varepsilon(x-y)(f(y) - f(x))\,dy \right| \\
&\leqslant \frac{1}{\varepsilon^n} \int_{B(x,\varepsilon)} \phi\Big(\frac{x-y}{\varepsilon}\Big) |f(y) - f(x)|\,dy \\
&\leqslant \Omega_n \|\phi\|_{L^\infty(\mathbb{R}^n)} \frac{1}{|B(x,\varepsilon)|} \int_{B(x,\varepsilon)} |f(y) - f(x)|\,dy \xrightarrow{\varepsilon \to 0} 0
\end{aligned}$$

for almost every $x \in \Omega$. Here $\Omega_n = |B(0,1)|$ and the last convergence follows from the Lebesgue differentiation theorem.

(3) Let $\Omega' \Subset \Omega'' \Subset \Omega$, $0 < \varepsilon < \operatorname{dist}(\Omega', \partial\Omega'')$, and $x \in \Omega'$. Because $\overline{\Omega''}$ is compact and $f \in C(\Omega)$, f is uniformly continuous in Ω'', that is, for every $\varepsilon' > 0$ there exists $\delta > 0$ such that

$$|f(x) - f(y)| < \varepsilon' \text{ for every } x, y \in \Omega'' \text{ with } |x - y| < \delta.$$

By combining this with an estimate from the proof of (ii), we conclude that

$$\begin{aligned}
|f_\varepsilon(x) - f(x)| &\leqslant \Omega_n \|\phi\|_{L^\infty(\mathbb{R}^n)} \frac{1}{|B(x,\varepsilon)|} \int_{B(x,\varepsilon)} |f(y) - f(x)|\,dy \\
&< \Omega_n \|\phi\|_{L^\infty(\mathbb{R}^n)} \varepsilon'
\end{aligned}$$

for every $x \in \Omega'$ if $\varepsilon < \delta$.

(4) Let $\Omega' \Subset \Omega'' \Subset \Omega$.

Claim:

$$\int_{\Omega'} |f_\varepsilon|^p\,dx \leqslant \int_{\Omega''} |f|^p\,dx$$

whenever $0 < \varepsilon < \operatorname{dist}(\Omega', \partial\Omega'')$ and $0 < \varepsilon < \operatorname{dist}(\Omega'', \partial\Omega)$.

Reason. Take $x \in \Omega'$. Hölder's inequality implies

$$\begin{aligned}
|f_\varepsilon(x)| &= \left| \int_{B(x,\varepsilon)} \phi_\varepsilon(x-y) f(y)\,dy \right| \\
&\leqslant \int_{B(x,\varepsilon)} \phi_\varepsilon(x-y)^{1-\frac{1}{p}} \phi_\varepsilon(x-y)^{\frac{1}{p}} |f(y)|\,dy \\
&\leqslant \left(\int_{B(x,\varepsilon)} \phi_\varepsilon(x-y)\,dy \right)^{\frac{1}{p'}} \left(\int_{B(x,\varepsilon)} \phi_\varepsilon(x-y) |f(y)|^p\,dy \right)^{\frac{1}{p}}
\end{aligned}$$

By raising the previous estimate to power p and by integrating over Ω', we obtain

$$\begin{aligned}\int_{\Omega'} |f_\varepsilon(x)|^p\,dx &\leqslant \int_{\Omega'}\int_{B(x,\varepsilon)} \phi_\varepsilon(x-y)|f(y)|^p\,dy\,dx\\ &= \int_{\Omega''}\int_{\Omega'} \phi_\varepsilon(x-y)|f(y)|^p\,dx\,dy\\ &= \int_{\Omega''} |f(y)|^p \int_{\Omega'} \phi_\varepsilon(x-y)\,dx\,dy\\ &= \int_{\Omega''} |f(y)|^p\,dy.\end{aligned}$$

Here we used Fubini's theorem and once more the fact that the integral of ϕ_ε is one. ■

Since $C(\Omega'')$ is dense in $L^p(\Omega'')$. Therefore for every $\varepsilon' > 0$ there exists $g \in C(\Omega'')$ such that

$$\left(\int_{\Omega''} |f-g|^p\,dx\right)^{\frac{1}{p}} \leqslant \frac{\varepsilon'}{3}.$$

By (3), we have $g_\varepsilon \to g$ uniformly in Ω' as $\varepsilon \to 0$. Thus

$$\left(\int_{\Omega'} |g_\varepsilon - g|^p\,dx\right)^{\frac{1}{p}} \leqslant \sup_{\Omega'} |g_\varepsilon - g|\,|\Omega'|^{\frac{1}{p}} < \frac{\varepsilon'}{3},$$

when $\varepsilon > 0$ is small enough. Now we use Minkowski's inequality and the previous claim to conclude that

$$\begin{aligned}\left(\int_{\Omega'} |f_\varepsilon - f|^p\,dx\right)^{\frac{1}{p}} &\leqslant \left(\int_{\Omega'} |f_\varepsilon - g_\varepsilon|^p\,dx\right)^{\frac{1}{p}}\\ &\qquad + \left(\int_{\Omega'} |g_\varepsilon - g|^p\,dx\right)^{\frac{1}{p}} + \left(\int_{\Omega'} |g - f|^p\,dx\right)^{\frac{1}{p}}\\ &\leqslant 2\left(\int_{\Omega''} |g - f|^p\,dx\right)^{\frac{1}{p}} + \left(\int_{\Omega'} |g_\varepsilon - g|^p\,dx\right)^{\frac{1}{p}}\\ &\leqslant 2\frac{\varepsilon'}{3} + \frac{\varepsilon'}{3} = \varepsilon'.\end{aligned}$$

Thus $f_\varepsilon \to f$ in $L^p(\Omega')$ as $\varepsilon \to 0$. □

1.7 Local approximation in Sobolev spaces

Next we show that the convolution approximation converges locally in Sobolev spaces.

Theorem 1.19. Let $u \in W^{k,p}(\Omega)$, $1 \leqslant p < \infty$. then

(1) $D^\alpha u_\varepsilon = D^\alpha u * \phi_\varepsilon$ in Ω_ε and
(2) $u_\varepsilon \to u$ as $\varepsilon \to 0$ in $W^{k,p}(\Omega')$ for every $\Omega' \Subset \Omega$.

THE MORAL: Smooth functions are dense in local Sobolev spaces. Thus every Sobolev function can be locally approximated with a smooth function in the Sobolev norm.

Proof. (1) By Young's theorem

$$\|u*\phi_\varepsilon\|_{L^p(\Omega_\varepsilon)} \leqslant \|u\|_{L^p(\Omega_\varepsilon)}\|\phi_\varepsilon\|_{L^1(\Omega_\varepsilon)} = \|u\|_{L^p(\Omega_\varepsilon)} < \infty$$

for every $\varepsilon > 0$, since $\|\phi_\varepsilon\|_{L^1(\Omega_\varepsilon)} = 1$ for every $r > 0$. This shows that $u*\phi_\varepsilon \in L^p(\Omega_\varepsilon)$ and, by Hölder's inequality, that $u*\phi_\varepsilon \in L^1_{\text{loc}}(\Omega_\varepsilon)$ for every $\varepsilon > 0$. A similar argument shows that

$$\|D^\alpha u*\phi_\varepsilon\|_{L^p(\Omega_\varepsilon)} \leqslant \|D^\alpha u\|_{L^p(\Omega_\varepsilon)}\|\phi_\varepsilon\|_{L^1(\Omega_\varepsilon)} = \|D^\alpha u\|_{L^p(\Omega_\varepsilon)} < \infty$$

for every $\varepsilon > 0$. Thus $D^\alpha u*\phi_\varepsilon \in L^p(\Omega_\varepsilon)$ and, by Hölder's inequality $D^\alpha u*\phi_\varepsilon \in L^1_{\text{loc}}(\Omega_\varepsilon)$ for every $\varepsilon > 0$. An alternative way to show that $u*\phi_\varepsilon \in L^1_{\text{loc}}(\Omega_\varepsilon)$ and $D^\alpha u*\phi_\varepsilon \in L^1_{\text{loc}}(\Omega_\varepsilon)$ is to apply Lemma 1.18 (1) to conclude that $u*\phi_\varepsilon \in C^\infty(\Omega_\varepsilon)$ and $D^\alpha u*\phi_\varepsilon \in C^\infty(\Omega_\varepsilon)$ for every $\varepsilon > 0$.

Let $\varphi \in C_0^\infty(\Omega_\varepsilon)$. By a repeated application of Fubini's theorem, we have

$$\begin{aligned}
\int_{\Omega_\varepsilon}(u*\phi_\varepsilon)(x)D^\alpha\varphi(x)\,dx &= \int_{\Omega_\varepsilon}\left(\int_{\Omega_\varepsilon}u(x-y)\phi_\varepsilon(y)\,dy\right)D^\alpha\varphi(x)\,dx \\
&= \int_{\Omega_\varepsilon}\left(\int_{\Omega_\varepsilon}u(x-y)D^\alpha\varphi(x)\,dx\right)\phi_\varepsilon(y)\,dy \\
&= (-1)^{|\alpha|}\int_{\Omega_\varepsilon}\left(\int_{\Omega_\varepsilon}\varphi(x)D^\alpha u(x-y)\,dx\right)\phi_\varepsilon(y)\,dy \\
&= (-1)^{|\alpha|}\int_{\Omega_\varepsilon}\left(\int_{\Omega_\varepsilon}\phi_\varepsilon(y)D^\alpha u(x-y)\,dy\right)\varphi(x)\,dx \\
&= (-1)^{|\alpha|}\int_{\Omega_\varepsilon}(D^\alpha u*\phi_\varepsilon)(x)\varphi(x)\,dx.
\end{aligned}$$

This shows that $D^\alpha(u*\phi_\varepsilon) = D^\alpha u*\phi_\varepsilon$ in Ω_ε. Here D^α denotes the weak partial derivative. By Lemma 1.18 (1) we have $u*\phi_\varepsilon \in C^\infty(\Omega_\varepsilon)$ and thus the classical derivative $D^\alpha(u*\phi_\varepsilon)$ equals with the weak derivative $D^\alpha(u*\phi_\varepsilon)$ in Ω_ε.

(2) Let $\Omega' \Subset \Omega$, and choose $\varepsilon > 0$ s.t. $\Omega' \subset \Omega_\varepsilon$. By (i) we know that $D^\alpha u_\varepsilon = D^\alpha u*\phi_\varepsilon$ in Ω', $|\alpha| \leqslant k$. By Lemma 1.18, we have $D^\alpha u_\varepsilon \to D^\alpha u$ in $L^p(\Omega')$ as $\varepsilon \to 0$, $|\alpha| \leqslant k$. Consequently

$$\|u_\varepsilon - u\|_{W^{k,p}(\Omega')} = \left(\sum_{|\alpha|\leqslant k}\|D^\alpha u_\varepsilon - D^\alpha u\|^p_{L^p(\Omega')}\right)^{\frac{1}{p}} \xrightarrow{\varepsilon\to 0} 0.$$

□

Remark 1.20. Theorem 1.19 (1) can also be proved by applying the fact that

$u * \phi_\varepsilon \in C^\infty(\Omega_\varepsilon)$, see Lemma 1.18 (1). Fix $x \in \Omega_\varepsilon$. Then

$$\begin{aligned} D^\alpha u_\varepsilon(x) &= D^\alpha(u * \phi_\varepsilon)(x) = (u * D^\alpha \phi_\varepsilon)(x) \\ &= \int_\Omega D_x^\alpha \phi_\varepsilon(x-y) u(y)\,dy \\ &= (-1)^{|\alpha|} \int_\Omega D_y^\alpha(\phi_\varepsilon(x-y)) u(y)\,dy. \end{aligned}$$

Here we first used the proof of Lemma 1.18 (1) and then the fact that

$$\frac{\partial}{\partial x_j}\left(\phi\left(\frac{x-y}{\varepsilon}\right)\right) = -\frac{\partial}{\partial x_j}\left(\phi\left(\frac{y-x}{\varepsilon}\right)\right) = -\frac{\partial}{\partial y_j}\left(\phi\left(\frac{x-y}{\varepsilon}\right)\right).$$

For every $x \in \Omega_\varepsilon$, the function $\varphi(y) = \phi_\varepsilon(x-y)$ belongs to $C_0^\infty(\Omega)$. Therefore

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

By combining the above facts, we see that

$$D^\alpha u_\varepsilon(x) = (-1)^{|\alpha|+|\alpha|} \int_\Omega D^\alpha u(y) \phi_\varepsilon(x-y)\,dy = (D^\alpha u * \phi_\varepsilon)(x).$$

Notice that $(-1)^{|\alpha|+|\alpha|} = 1$.

1.8 Global approximation in Sobolev spaces

The next result shows that the convolution approximation converges also globally in Sobolev spaces.

Theorem 1.21 (Meyers-Serrin). If $u \in W^{k,p}(\Omega)$, $1 \leqslant p < \infty$, then there exist functions $u_i \in C^\infty(\Omega) \cap W^{k,p}(\Omega)$ such that $u_i \to u$ in $W^{k,p}(\Omega)$.

THE MORAL: Smooth functions are dense in Sobolev spaces. Thus every Sobolev function can be approximated with a smooth function in the Sobolev norm. In particular, this holds true for the function with a dense infinity set in Example 1.10.

Proof. Let $\Omega_0 = \emptyset$ and

$$\Omega_i = \left\{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > \tfrac{1}{i}\right\} \cap B(0,i), \quad i = 1,2,\dots.$$

Then

$$\Omega = \bigcup_{i=1}^\infty \Omega_i \quad \text{and} \quad \Omega_1 \Subset \Omega_2 \Subset \ldots \Subset \Omega.$$

Claim: There exist $\eta_i \in C_0^\infty(\Omega_{i+2}\setminus\overline{\Omega}_{i-1})$, $i = 1,2,\dots$, such that $0 \leqslant \eta_i \leqslant 1$ and

$$\sum_{i=1}^{\infty} \eta_i(x) = 1 \quad \text{for every} \quad x \in \Omega.$$

This is a partition of unity subordinate to the covering $\{\Omega_i\}$.

Reason. By using the distance function and convolution approximation we can construct $\widetilde{\eta}_i \in C_0^\infty(\Omega_{i+2}\setminus\overline{\Omega}_{i-1})$ such that $0 \leqslant \widetilde{\eta}_i \leqslant 1$ and $\widetilde{\eta}_i = 1$ in $\overline{\Omega}_{i+1}\setminus\Omega_i$ (exercise). Then we define

$$\eta_i(x) = \frac{\widetilde{\eta}_i(x)}{\sum_{j=1}^{\infty}\widetilde{\eta}_j(x)}, \quad i = 1,2,\dots.$$

Observe that the sum is only over four indices in a neighbourhood of a given point. ■

By Lemma 1.14 (5), $\eta_i u \in W^{k,p}(\Omega)$ and

$$\operatorname{supp}(\eta_i u) \subset \Omega_{i+2}\setminus\overline{\Omega}_{i-1}.$$

Let $\varepsilon > 0$. Choose $\varepsilon_i > 0$ so small that

$$\operatorname{supp}(\phi_{\varepsilon_i} * (\eta_i u)) \subset \Omega_{i+2}\setminus\overline{\Omega}_{i-1}$$

(see Remark 1.17 (4)) and

$$\|\phi_{\varepsilon_i} * (\eta_i u) - \eta_i u\|_{W^{k,p}(\Omega)} < \frac{\varepsilon}{2^i}, \quad i = 1,2,\dots.$$

By Theorem 1.19 (2), this is possible. Let

$$v = \sum_{i=1}^{\infty} \phi_{\varepsilon_i} * (\eta_i u).$$

This function belongs to $C^\infty(\Omega)$, since in a neighbourhood of any point $x \in \Omega$, there are at most finitely many nonzero terms in the sum. Since

$$u = u\sum_{i=1}^{\infty}\eta_i = \sum_{i=1}^{\infty}\eta_i u,$$

we have

$$\begin{aligned}
\|v-u\|_{W^{k,p}(\Omega)} &= \left\| \sum_{i=1}^{\infty} \phi_{\varepsilon_i} * (\eta_i u) - \sum_{i=1}^{\infty} \eta_i u \right\|_{W^{k,p}(\Omega)} \\
&\leqslant \sum_{i=1}^{\infty} \left\| \phi_{\varepsilon_i} * (\eta_i u) - \eta_i u \right\|_{W^{k,p}(\Omega)} \leqslant \sum_{i=1}^{\infty} \frac{\varepsilon}{2^i} = \varepsilon.
\end{aligned}$$ □

Remarks 1.22:

(1) The Meyers-Serrin theorem 1.21 gives the following characterization for the Sobolev spaces $W^{k,p}(\Omega)$, $1 \leqslant p < \infty$: $u \in W^{k,p}(\Omega)$ if and only if there exist functions $u_i \in C^\infty(\Omega)\cap W^{k,p}(\Omega)$, $i = 1,2,\dots$, such that $u_i \to u$ in $W^{k,p}(\Omega)$ as $i \to \infty$. More precisely, $W^{k,p}(\Omega)$ is the completion of $C^\infty(\Omega)$ in the Sobolev norm.

Reason. $\boxed{\Longrightarrow}$ Theorem 1.21.
$\boxed{\Longleftarrow}$ Theorem 1.15. ■

(2) The Meyers-Serrin theorem 1.21 is false for $p=\infty$. Indeed, if $u_i \in C^\infty(\Omega)\cap W^{1,\infty}(\Omega)$ such that $u_i \to u$ in $W^{1,\infty}(\Omega)$, then $u \in C^1(\Omega)$ (exercise). Thus special care is required when we consider approximations in $W^{1,\infty}(\Omega)$.

(3) Let $\Omega' \Subset \Omega$. The proof of Theorem 1.19 and Theorem 1.21 shows that for every $\varepsilon > 0$ there exists $v \in C_0^\infty(\Omega)$ such that $\|v-u\|_{W^{1,p}(\Omega')} < \varepsilon$.

(4) The proof of Theorem 1.21 shows that not only $C^\infty(\Omega)$ but also $C_0^\infty(\Omega)$ is dense in $L^p(\Omega)$, $1 \leqslant p < \infty$.

1.9 Sobolev spaces with zero boundary values

In this section we study definitions and properties of first order Sobolev spaces with zero boundary values in an open subset of $\mathbb{R}^n$. A similar theory can be developed for higher order Sobolev spaces as well. Recall that, by Theorem 1.21, the Sobolev space $W^{1,p}(\Omega)$ can be characterized as the completion of $C^\infty(\Omega)$ with respect to the Sobolev norm when $1 \leqslant p < \infty$.

Definition 1.23. Let $1 \leqslant p < \infty$. The Sobolev space with zero boundary values $W_0^{1,p}(\Omega)$ is the completion of $C_0^\infty(\Omega)$ with respect to the Sobolev norm. Thus $u \in W_0^{1,p}(\Omega)$ if and only if there exist functions $u_i \in C_0^\infty(\Omega)$, $i = 1,2,\dots$, such that $u_i \to u$ in $W^{1,p}(\Omega)$ as $i \to \infty$. The space $W_0^{1,p}(\Omega)$ is endowed with the norm of $W^{1,p}(\Omega)$.

THE MORAL: The only difference compared to $W^{1,p}(\Omega)$ is that functions in $W_0^{1,p}(\Omega)$ can be approximated by $C_0^\infty(\Omega)$ functions instead of $C^\infty(\Omega)$ functions, that is,

$$W^{1,p}(\Omega) = \overline{C^\infty(\Omega)} \quad \text{and} \quad W_0^{1,p}(\Omega) = \overline{C_0^\infty(\Omega)},$$

where the completions are taken with respect to the Sobolev norm. Observe that $C_0^\infty(\Omega)$ functions have finite Sobolev norm, but for $C^\infty(\Omega)$ we consider functions with finite Sobolev norm. A function in $W_0^{1,p}(\Omega)$ has zero boundary values in Sobolev's sense. We may say that $u, v \in W^{1,p}(\Omega)$ have the same boundary values in Sobolev's sense, if $u - v \in W_0^{1,p}(\Omega)$. This is useful, for example, in Dirichlet problems for PDEs.

WARNING: Roughly speaking a function in $W^{1,p}(\Omega)$ belongs to $W_0^{1,p}(\Omega)$, if it vanishes on the boundary. This is a delicate issue, since the function does not have to be zero pointwise on the boundary. We shall return to this question later.

Remarks 1.24:

(1) Clearly $C_0^\infty(\Omega) \subset W_0^{1,p}(\Omega) \subset W^{1,p}(\Omega) \subset L^p(\Omega)$.

(2) $W_0^{1,p}(\Omega)$ is a closed subspace of $W^{1,p}(\Omega)$ and thus complete (exercise).

(3) By the Meyers-Serrin theorem, see Theorem 1.19 and Theorem 1.21, we conclude that $u \in W_0^{1,p}(\Omega)$ if and only if there exist functions $u_i \in C_0(\Omega) \cap W^{1,p}(\Omega)$, $i = 1,2,\dots$, such that $u_i \to u$ in $W^{1,p}(\Omega)$ as $i \to \infty$ (exercise). This useful observation can be applied to show that a function belongs to a Sobolev space with zero boundary values. The advantage is that the approximating functions do not necessarily have to be smooth.

Theorem 1.25. Let $1 \leqslant p < \infty$ and let $\Omega \subset \mathbb{R}^n$ be an open set. Assume that $u \in W_0^{1,p}(\Omega)$ and let u_0 be the zero extension of u, that is,

$$u_0(x) = \begin{cases} u(x), & x \in \Omega, \\ 0, & x \in \mathbb{R}^n \setminus \Omega. \end{cases}$$

Then $u_0 \in W^{1,p}(\mathbb{R}^n)$ and

$$Du_0 = \begin{cases} Du & \text{a.e. in } \Omega, \\ 0 & \text{a.e. in } \mathbb{R}^n \setminus \Omega. \end{cases}$$

In particular, $\|u_0\|_{W^{1,p}(\mathbb{R}^n)} = \|u\|_{W^{1,p}(\Omega)}$.

T H E M O R A L : Functions in a Sobolev space with zero boundary values can always be extended by zero so that the obtained function belongs to the corresponding Sobolev space over the entire space.

Proof. Let $Du = (D_1u,\dots,D_nu)$ be the weak gradient of u in Ω, and let $f = (f_1,\dots,f_n)$ be the componentwise zero extension of Du. Since $u \in W_0^{1,p}(\Omega)$, there exist functions $u_i \in C_0^\infty(\Omega)$, $i = 1,2,\dots$, such that $u_i \to u$ in $W^{1,p}(\Omega)$ as $i \to \infty$. Let $\varphi \in C_0^\infty(\mathbb{R}^n)$. By Hölder's inequality we have

$$\begin{aligned}
\left| \int_\Omega u D_j\varphi\,dx - \int_\Omega u_i D_j\varphi\,dx \right| &\leqslant \int_\Omega |u - u_i||D_j\varphi|\,dx \\
&\leqslant \left(\int_\Omega |u-u_i|^p\,dx \right)^{\frac{1}{p}} \left(\int_\Omega |D_j\varphi|^{p'}\,dx \right)^{\frac{1}{p'}} \\
&\leqslant \left(\int_\Omega |u-u_i|^p\,dx \right)^{\frac{1}{p}} \left(\int_{\mathbb{R}^n} |D_j\varphi|^{p'}\,dx \right)^{\frac{1}{p'}} \xrightarrow{i\to\infty} 0, \quad j = 1,\dots,n,
\end{aligned}$$

where $\frac{1}{p} + \frac{1}{p'} = 1$. A similar argument shows that

$$\left| \int_\Omega D_j u_i \varphi\,dx - \int_\Omega D_j u \varphi\,dx \right| \leqslant \left(\int_\Omega |D_j u - D_j u_i|^p\,dx \right)^{\frac{1}{p}} \left(\int_{\mathbb{R}^n} |\varphi|^{p'}\,dx \right)^{\frac{1}{p'}} \xrightarrow{i\to\infty} 0$$

for every $j = 1,2,\dots,n$. Moreover, integration by parts for the smooth functions $u_i \in C_0^\infty(\Omega)$, $i = 1,2,\dots$ and $\varphi \in C_0^\infty(\mathbb{R}^n)$ gives

$$-\int_\Omega u_i D_j \varphi\,dx = \int_\Omega D_j u_i \varphi\,dx, \quad j = 1,\dots,n,$$

for every $i = 1,2,\dots$. It follows that

$$\begin{aligned}-\int_{\mathbb{R}^n} u_0 D_j \varphi\,dx &= -\int_\Omega u D_j \varphi\,dx = -\lim_{i\to\infty}\int_\Omega u_i D_j \varphi\,dx\\ &= \lim_{i\to\infty}\int_\Omega D_j u_i \varphi\,dx = \int_\Omega D_j u \varphi\,dx\\ &= \int_{\mathbb{R}^n} f_j \varphi\,dx, \quad j = 1,\dots,n,\end{aligned}$$

for every $\varphi \in C_0^\infty(\mathbb{R}^n)$. By the uniqueness of weak derivatives, see Lemma 1.4, we conclude that the weak gradient Du_0 of u_0 in $\mathbb{R}^n$ coincides almost everywhere with f. It follows that $u_0 \in W^{1,p}(\mathbb{R}^n)$ and $\|u_0\|_{W^{1,p}(\mathbb{R}^n)} = \|u\|_{W^{1,p}(\Omega)}$. □

Remark 1.26. Theorem 1.25 can be also proved by the Meyers-Serrin theorem, see Theorem 1.21. Since $u \in W_0^{1,p}(\Omega)$, there exist functions $u_i \in C_0^\infty(\Omega)$, $i = 1,2,\dots$, such that $u_i \to u$ in $W^{1,p}(\Omega)$ as $i \to \infty$. The zero extensions of u_i belong to $C^\infty(\mathbb{R}^n)$ and converge to the zero extension of u in $W^{1,p}(\mathbb{R}^n)$ as $i \to \infty$. The claims follow from this.

Lemma 1.27. If $u \in W^{1,p}(\Omega)$ and $\operatorname{supp} u$ is a compact subset of Ω, then $u \in W_0^{1,p}(\Omega)$.

Proof. Let $\eta \in C_0^\infty(\Omega)$ be a cutoff function such that $\eta = 1$ on the support of u.

Claim: If $u_i \in C^\infty(\Omega)$, $i = 1,2,\dots$, such that $u_i \to u$ in $W^{1,p}(\Omega)$, then $\eta u_i \in C_0^\infty(\Omega)$ converges to $\eta u = u$ in $W^{1,p}(\Omega)$.

Reason. We observe that

$$\begin{aligned}\|\eta u_i - \eta u\|_{W^{1,p}(\Omega)} &= \left(\|\eta u_i - \eta u\|_{L^p(\Omega)}^p + \|D(\eta u_i - \eta u)\|_{L^p(\Omega)}^p\right)^{\frac{1}{p}}\\ &\leqslant \|\eta u_i - \eta u\|_{L^p(\Omega)} + \|D(\eta u_i - \eta u)\|_{L^p(\Omega)},\end{aligned}$$

where

$$\begin{aligned}\|\eta u_i - \eta u\|_{L^p(\Omega)} &= \left(\int_\Omega |\eta u_i - \eta u|^p\,dx\right)^{\frac{1}{p}} = \left(\int_\Omega |\eta|^p |u_i - u|^p\,dx\right)^{\frac{1}{p}}\\ &\leqslant \|\eta\|_{L^\infty(\Omega)}\left(\int_\Omega |u_i - u|^p\,dx\right)^{\frac{1}{p}} \xrightarrow{i\to\infty} 0\end{aligned}$$

and by Lemma 1.14 (5)

$$\begin{aligned}
\|D(\eta u_i - \eta u)\|_{L^p(\Omega)} &= \left(\int_\Omega |D(\eta u_i - \eta u)|^p\,dx\right)^{\frac{1}{p}} \\
&= \left(\int_\Omega |(u_i - u)D\eta + (Du_i - Du)\eta|^p\,dx\right)^{\frac{1}{p}} \\
&\leqslant \left(\int_\Omega |(u_i - u)D\eta|^p\,dx\right)^{\frac{1}{p}} + \left(\int_\Omega |(Du_i - Du)\eta|^p\,dx\right)^{\frac{1}{p}} \\
&\leqslant \|D\eta\|_{L^\infty(\Omega)}\left(\int_\Omega |u_i - u|^p\,dx\right)^{\frac{1}{p}} + \|\eta\|_{L^\infty(\Omega)}\left(\int_\Omega |Du_i - Du|^p\,dx\right)^{\frac{1}{p}} \xrightarrow{i\to\infty} 0. \quad \blacksquare
\end{aligned}$$

Since $\eta u_i \in C_0^\infty(\Omega)$, $i = 1,2,\dots$, and $\eta u_i \to u$ in $W^{1,p}(\Omega)$, we conclude that $u \in W_0^{1,p}(\Omega)$. □

Since $W_0^{1,p}(\Omega) \subset W^{1,p}(\Omega)$, functions in these spaces have similar general properties and they will not be repeated here. Thus we shall focus on properties that are typical for Sobolev spaces with zero boundary values.

Lemma 1.28. $W^{1,p}(\mathbb{R}^n) = W_0^{1,p}(\mathbb{R}^n)$ with $1 \leqslant p < \infty$.

T H E M O R A L : The standard Sobolev space and the Sobolev space with zero boundary value coincide in the whole space.

W A R N I N G : $W^{1,p}(B(0,1)) \neq W_0^{1,p}(B(0,1))$, $1 \leqslant p < \infty$. Thus the spaces are not same in general.

Proof. Assume that $u \in W^{1,p}(\mathbb{R}^n)$. Let $\eta_k \in C_0^\infty(B(0,k+1))$ such that $\eta_k = 1$ on $B(0,k)$, $0 \leqslant \eta_k \leqslant 1$ and $|D\eta_k| \leqslant c$. Lemma 1.27 implies $u\eta_k \in W_0^{1,p}(\mathbb{R}^n)$.

Claim: $u\eta_k \to u$ in $W^{1,p}(\mathbb{R}^n)$ as $k \to \infty$.

Reason.

$$\begin{aligned}
\|u - u\eta_k\|_{W^{1,p}(\mathbb{R}^n)} &\leqslant \|u - u\eta_k\|_{L^p(\mathbb{R}^n)} + \|D(u - u\eta_k)\|_{L^p(\mathbb{R}^n)} \\
&= \left(\int_{\mathbb{R}^n} |u(1-\eta_k)|^p\,dx\right)^{\frac{1}{p}} + \left(\int_{\mathbb{R}^n} |D(u(1-\eta_k))|^p\,dx\right)^{\frac{1}{p}} \\
&= \left(\int_{\mathbb{R}^n} |u(1-\eta_k)|^p\,dx\right)^{\frac{1}{p}} + \left(\int_{\mathbb{R}^n} |(1-\eta_k)Du - uD\eta_k)|^p\,dx\right)^{\frac{1}{p}} \\
&\leqslant \left(\int_{\mathbb{R}^n} |u(1-\eta_k)|^p\,dx\right)^{\frac{1}{p}} + \left(\int_{\mathbb{R}^n} |(1-\eta_k)Du|^p\,dx\right)^{\frac{1}{p}} + \left(\int_{\mathbb{R}^n} |uD\eta_k|^p\,dx\right)^{\frac{1}{p}}.
\end{aligned}$$

We note that $\lim_{k\to\infty} u(1-\eta_k) = 0$ almost everywhere and $|u(1-\eta_k)|^p \leqslant |u|^p \in L^1(\mathbb{R}^n)$ will do as an integrable majorant. The dominated convergence theorem gives

$$\left(\int_{\mathbb{R}^n} |u(1-\eta_k)|^p\,dx\right)^{\frac{1}{p}} \xrightarrow{k\to\infty} 0.$$

A similar argument shows that

$$\left(\int_{\mathbb{R}^n} |(1-\eta_k)Du|^p \, dx\right)^{\frac{1}{p}} \xrightarrow{k\to\infty} 0.$$

Moreover, by the dominated convergence theorem

$$\begin{aligned}\left(\int_{\mathbb{R}^n} |uD\eta_k|^p \, dx\right)^{\frac{1}{p}} &\leqslant c\left(\int_{B(0,k+1)\setminus B(0,k)} |u|^p \, dx\right)^{\frac{1}{p}} \\ &= c\left(\int_{\mathbb{R}^n} |u|^p \chi_{B(0,k+1)\setminus B(0,k)} \, dx\right)^{\frac{1}{p}} \xrightarrow{k\to\infty} 0.\end{aligned}$$

Here $|u|^p \chi_{B(0,k+1)\setminus B(0,k)} \leqslant |u|^p \in L^1(\mathbb{R}^n)$ will do as an integrable majorant. ■

Since $u\eta_k \in W_0^{1,p}(\mathbb{R}^n)$, $i = 1,2,\ldots$, $u\eta_k \to u$ in $W^{1,p}(\mathbb{R}^n)$ as $k \to \infty$ and $W_0^{1,p}(\mathbb{R}^n)$ is complete, we conclude that $u \in W_0^{1,p}(\mathbb{R}^n)$. □

2 Methods and characterizations

2.1 Chain rule

We shall prove some useful results for the first order Sobolev spaces.

Lemma 2.1 (Chain rule). Let $u \in W^{1,p}(\Omega)$, $1 \leqslant p < \infty$, and $f \in C^1(\mathbb{R})$ such that $f' \in L^\infty(\mathbb{R})$ and $f(0) = 0$. Then $f \circ u \in W^{1,p}(\Omega)$ and

$$D_j(f \circ u) = f'(u) D_j u, \quad j = 1, 2, \dots, n$$

almost everywhere in Ω.

Proof. By Theorem 1.21, there exist a sequence of functions $u_i \in C^\infty(\Omega) \cap W^{1,p}(\Omega)$, $i = 1, 2, \dots$, such that $u_i \to u$ in $W^{1,p}(\Omega)$ as $i \to \infty$. Let $\varphi \in C_0^\infty(\Omega)$.

Claim: $\displaystyle\int_\Omega f(u) D_j \varphi \, dx = \lim_{i\to\infty} \int_\Omega f(u_i) D_j \varphi \, dx$, $j = 1, \dots, n$.

Reason. $\boxed{1 < p < \infty}$ By Hölder's inequality

$$\begin{aligned}
\left| \int_\Omega f(u) D_j \varphi \, dx - \int_\Omega f(u_i) D_j \varphi \, dx \right| &\leqslant \int_\Omega |f(u) - f(u_i)| |D\varphi| \, dx \\
&\leqslant \left(\int_\Omega |f(u) - f(u_i)|^p \, dx \right)^{\frac{1}{p}} \left(\int_\Omega |D\varphi|^{p'} \, dx \right)^{\frac{1}{p'}} \\
&\leqslant \|f'\|_\infty \left(\int_\Omega |u - u_i|^p \, dx \right)^{\frac{1}{p}} \left(\int_\Omega |D\varphi|^{p'} \, dx \right)^{\frac{1}{p'}} \xrightarrow{i\to\infty} 0.
\end{aligned}$$

In the last inequality, we used the fact that

$$|f(u) - f(u_i)| = \left| \int_{u_i}^u f'(t) \, dt \right| \leqslant \|f'\|_\infty |u - u_i|.$$

$\boxed{p=1}$ A similar argument as above gives

$$\begin{aligned}\left|\int_\Omega f(u)D_j\varphi\,dx-\int_\Omega f(u_i)D_j\varphi\,dx\right| &\leqslant \int_\Omega |f(u)-f(u_i)||D\varphi|\,dx\\ &\leqslant \|f'\|_\infty\|D\varphi\|_{L^\infty(\Omega)}\int_\Omega |u-u_i|\,dx \xrightarrow{i\to\infty} 0.\end{aligned}$$

■

Claim: $\lim\limits_{i\to\infty}\int_\Omega f'(u_i)D_ju_i\varphi\,dx=\int_\Omega f'(u)D_ju\varphi\,dx.$

Reason. $\boxed{1<p<\infty}$ By Hölder's inequality

$$\begin{aligned}&\left|\int_\Omega f'(u_i)D_ju_i\varphi\,dx-\int_\Omega f'(u)D_ju\varphi\,dx\right|\\ &\quad=\left|\int_\Omega f'(u_i)(D_ju_i-D_ju)\varphi\,dx-\int_\Omega (f'(u)-f'(u_i)D_ju\varphi\,dx\right|\\ &\quad\leqslant \|f'\|_\infty\int_\Omega |D_ju_i-D_ju||\varphi|\,dx+\int_\Omega |f'(u)-f'(u_i)||D_ju||\varphi|\,dx\\ &\quad\leqslant \|f'\|_\infty\left(\int_\Omega |D_ju_i-Du|^p\,dx\right)^{\frac1p}\left(\int_\Omega |\varphi|^{p'}\,dx\right)^{\frac1{p'}}\\ &\qquad+\|f'\|_\infty\left(\int_\Omega |u-u_i|^p\,dx\right)^{\frac1p}\left(\int_\Omega |D_ju\varphi|^{p'}\,dx\right)^{\frac1{p'}}\xrightarrow{i\to\infty}0.\end{aligned}$$

The estimate for the second integral on the right-hand side in the last inequality follows as in the proof of the first claim above.

$\boxed{p=1}$ A similar argument as above gives

$$\begin{aligned}&\left|\int_\Omega f'(u_i)D_ju_i\varphi\,dx-\int_\Omega f'(u)D_ju\varphi\,dx\right|\\ &\qquad\leqslant \|f'\|_\infty\|\varphi\|_{L^\infty(\Omega)}\int_\Omega |D_ju_i-Du|\,dx\\ &\qquad\quad+\|\varphi\|_{L^\infty(\Omega)}\int_\Omega |f'(u)-f'(u_i)||D_ju|\,dx.\end{aligned}$$

Since $u_i\to u$ in $L^1(\Omega)$, by passing to a subsequence, we may assume that $u_i\to u$ almost everywhere in Ω as $i\to\infty$. Since $f\in C^1(\mathbb{R})$, we we conclude that $f'(u_i)\to f'(u)$ almost everywhere in Ω as $i\to\infty$. We note that

$$|f'(u)-f'(u_i)|\leqslant|f'(u)|+|f'(u_i)|\leqslant 2\|f'\|_\infty,\quad i=1,2,\dots.$$

It follows that

$$\lim_{i\to\infty}|f'(u)-f'(u_i)||D_ju|=0$$

almost everywhere in Ω and

$$|f'(u)-f'(u_i)||D_ju|\leqslant 2\|f'\|_\infty|D_ju|\in L^1(\Omega),\quad i=1,2,\dots,$$

almost everywhere in Ω. By the dominated convergence theorem, we have

$$\lim_{i\to\infty}\int_\Omega |f'(u)-f'(u_i)||D_ju|\,dx=\int_\Omega\lim_{i\to\infty}|f'(u)-f'(u_i)||D_ju|\,dx=0.$$

■

Next, we use the claims above, integration by parts for smooth functions and the chain rule for smooth functions to obtain

$$\begin{aligned}\int_\Omega (f\circ u)D_j\varphi\,dx &= \lim_{i\to\infty}\int_\Omega f(u_i)D_j\varphi\,dx\\ &= -\lim_{i\to\infty}\int_\Omega D_j(f(u_i))\varphi\,dx\\ &= -\lim_{i\to\infty}\int_\Omega f'(u_i)D_ju_i\varphi\,dx\\ &= -\int_\Omega f'(u)D_ju\varphi\,dx\\ &= -\int_\Omega (f'\circ u)D_ju\varphi\,dx, \quad j=1,\dots,n,\end{aligned}$$

for every $\varphi\in C_0^\infty(\Omega)$.

Finally, we show that $f(u)$ and $f'(u)\frac{\partial u}{\partial x_j}$ belong to $L^p(\Omega)$. Since

$$|f(u)| = |f(u)-f(0)| = \left|\int_0^u f'(t)\,dt\right| \leqslant \|f'\|_\infty |u|,$$

we have

$$\left(\int_\Omega |f(u)|^p\,dx\right)^{1/p} \leqslant \|f'\|_\infty\left(\int_\Omega |u|^p\,dx\right)^{\frac{1}{p}} < \infty,$$

and similarly,

$$\left(\int_\Omega \left|f'(u)D_ju\right|^p\,dx\right)^{\frac{1}{p}} \leqslant \|f'\|_\infty\left(\int_\Omega |Du|^p\,dx\right)^{\frac{1}{p}} < \infty.$$

□

Remark 2.2. Lemma 2.1 also holds for $u\in W^{1,\infty}(\Omega)$, since then $u\in W^{1,p}_{\text{loc}}(\Omega)$ for $1\leqslant p<\infty$ (exercise).

2.2 Truncation

The truncation property is an important property of first order Sobolev spaces, which means that we can cut the functions at certain level and the truncated function is still in the same Sobolev space. Higher order Sobolev spaces do not enjoy this property, see Example 1.7.

Theorem 2.3. Let $u\in W^{1,p}(\Omega)$, $1\leqslant p<\infty$. Then $u^+=\max\{u,0\}\in W^{1,p}(\Omega)$, $u^-=-\min\{u,0\}\in W^{1,p}(\Omega)$, $|u|\in W^{1,p}(\Omega)$ and

$$Du^+ = \begin{cases} Du & \text{a.e. in } \Omega\cap\{u>0\},\\ 0 & \text{a.e. in } \Omega\cap\{u\leqslant 0\},\end{cases}$$

$$Du^- = \begin{cases} 0 & \text{a.e. in } \Omega\cap\{u\geqslant 0\},\\ -Du & \text{a.e. in } \Omega\cap\{u<0\},\end{cases}$$

and

$$D|u| = \begin{cases} Du & \text{a.e. in } \Omega \cap \{u > 0\}, \\ 0 & \text{a.e. in } \Omega \cap \{u = 0\}, \\ -Du & \text{a.e. in } \Omega \cap \{u < 0\}. \end{cases}$$

THE MORAL: In contrast with C^1, the Sobolev space $W^{1,p}$ are closed under taking absolute values.

Proof. Let $0 < \varepsilon < 1$ and let $f_\varepsilon : \mathbb{R} \to \mathbb{R}$, $f_\varepsilon(t) = \sqrt{t^2 + \varepsilon^2} - \varepsilon$. The function f_ε has the properties $f_\varepsilon \in C^1(\mathbb{R})$, $f_\varepsilon(0) = 0$,

$$\lim_{\varepsilon \to 0} f_\varepsilon(t) = |t| \quad \text{for every} \quad t \in \mathbb{R},$$

$$(f_\varepsilon)'(t) = \frac{1}{2}(t^2 + \varepsilon^2)^{-\frac{1}{2}} 2t = \frac{t}{\sqrt{t^2 + \varepsilon^2}} \quad \text{for every} \quad t \in \mathbb{R},$$

and $\|(f_\varepsilon)'\|_\infty \leqslant 1$ for every $\varepsilon > 0$. From Lemma 2.1, we conclude that $f_\varepsilon \circ u \in W^{1,p}(\Omega)$ and

$$\int_\Omega (f_\varepsilon \circ u) D_j \varphi \, dx = -\int_\Omega (f_\varepsilon)'(u) D_j u \varphi \, dx, \quad j = 1, \dots, n,$$

for every $\varphi \in C_0^\infty(\Omega)$. We note that

$$f_\varepsilon(t) = \sqrt{t^2 + \varepsilon^2} - \varepsilon \leqslant 2^{\frac{1}{2}}(|t| + \varepsilon) \leqslant \sqrt{2}(|t| + 1)$$

for every $t \in \mathbb{R}$ and $0 < \varepsilon < 1$ and thus

$$|(f_\varepsilon \circ u) D_j \varphi| \leqslant \sqrt{2} \|D\varphi\|_{L^\infty(\Omega)} (\|u\|_{L^\infty(\operatorname{supp}\varphi)} + 1) \chi_{\operatorname{supp}\varphi} \in L^1(\Omega)$$

for every $0 < \varepsilon < 1$ and $j = 1, \dots, n$. Moreover, we have $\|(f_\varepsilon)'\|_{L^\infty(\Omega)} \leqslant 1$ for every $0 < \varepsilon < 1$ and

$$\lim_{\varepsilon \to 0} (f_\varepsilon)'(t) = \begin{cases} 1, & t > 0, \\ 0, & t = 0, \\ -1, & t < 0. \end{cases}$$

Thus, by the dominated convergence theorem, we obtain

$$\begin{aligned} \int_\Omega |u| D_j \varphi \, dx &= \int_\Omega \lim_{\varepsilon \to 0} (f_\varepsilon \circ u) D_j \varphi \, dx \\ &= \lim_{\varepsilon \to 0} \int_\Omega (f_\varepsilon \circ u) D_j \varphi \, dx \\ &= -\lim_{\varepsilon \to 0} \int_\Omega (f_\varepsilon)'(u) D_j u \varphi \, dx \\ &= -\int_\Omega \lim_{\varepsilon \to 0} (f_\varepsilon)'(u) D_j u \varphi \, dx \\ &= -\int_\Omega D_j |u| \varphi \, dx, \quad j = 1, \dots, n, \end{aligned}$$

for every $\varphi \in C_0^\infty(\Omega)$, where $D_j|u|$ is as in the statement of the theorem.

The other claims follow by observing that

$$u^+ = \frac{1}{2}(u + |u|) \quad \text{and} \quad u^- = \frac{1}{2}(|u| - u). \qquad \square$$

Remarks 2.4:

(1) If $u, v \in W^{1,p}(\Omega)$, then $\max\{u,v\} \in W^{1,p}(\Omega)$ and $\min\{u,v\} \in W^{1,p}(\Omega)$. Moreover,

$$D\max\{u,v\} = \begin{cases} Du & \text{a.e. in } \Omega \cap \{u \geqslant v\}, \\ Dv & \text{a.e. in } \Omega \cap \{u \leqslant v\}, \end{cases}$$

and

$$D\min\{u,v\} = \begin{cases} Du & \text{a.e. in } \Omega \cap \{u \leqslant v\}, \\ Dv & \text{a.e. in } \Omega \cap \{u \geqslant v\}. \end{cases}$$

In particular, $Du = Dv$ almost everywhere in $\{x \in \Omega : u(x) = v(x)\}$.

Reason.

$$\max\{u,v\} = \frac{1}{2}(u + v + |u - v|) \quad \text{and} \quad \min\{u,v\} = \frac{1}{2}(u + v - |u - v|). \qquad \blacksquare$$

(2) If $u \in W^{1,p}(\Omega)$ and $\lambda \in \mathbb{R}$, then $Du = 0$ almost everywhere in $\{x \in \Omega : u(x) = \lambda\}$ (exercise).

(3) If $u \in W^{1,p}(\Omega)$ and $\lambda \in \mathbb{R}$, then $\min\{u,\lambda\} \in W^{1,p}_{\text{loc}}(\Omega)$ and

$$D\min\{u,\lambda\} = \begin{cases} Du & \text{a.e. in } \Omega \cap \{u < \lambda\}, \\ 0 & \text{a.e. in } \Omega \cap \{u \geqslant \lambda\}. \end{cases}$$

A similar claim also holds for $\max\{u,\lambda\}$. This implies that a function $u \in W^{1,p}(\Omega)$ can be approximated by the truncated functions

$$u_\lambda = \max\{-\lambda, \min\{u,\lambda\}\} = \begin{cases} \lambda & \text{a.e. in } \Omega \cap \{u \geqslant \lambda\}, \\ u & \text{a.e. in } \Omega \cap \{-\lambda < u < \lambda\}, \\ -\lambda & \text{a.e. in } \Omega \cap \{u \leqslant -\lambda\}, \end{cases}$$

in $W^{1,p}(\Omega)$. (Here $\lambda > 0$.)

Reason. By applying the dominated convergence theorem to

$$|u - u_\lambda|^p \leqslant 2^p(|u|^p + |u_\lambda|^p) \leqslant 2^{p+1}|u|^p \in L^1(\Omega),$$

we have

$$\lim_{\lambda\to\infty} \int_\Omega |u - u_\lambda|^p \, dx = \int_\Omega \lim_{\lambda\to\infty} |u - u_\lambda|^p \, dx = 0,$$

and by applying the dominated convergence theorem to

$$|Du - Du_\lambda|^p \leqslant |Du|^p \in L^1(\Omega),$$

we have

$$\lim_{\lambda\to\infty}\int_\Omega |Du-Du_\lambda|^p\,dx=\int_\Omega \lim_{\lambda\to\infty}|Du-Du_\lambda|^p\,dx=0.$$

■

THE MORAL: Bounded $W^{1,p}$ functions are dense in $W^{1,p}$.

We discuss a useful converge result which can be applied in proving truncation properties for Sobolev spaces with zero boundary values. The following slight extension of the dominated convergence theorem is useful in the proof.

Theorem 2.5. Let $f_i:\mathbb{R}^n\to[-\infty,\infty]$, $i=1,2,\dots$, be measurable functions such that $f_i\to f$ almost everywhere as $i\to\infty$. Assume that there exist integrable functions $g_i,h_i:\mathbb{R}^n\to[-\infty,\infty]$ such that $g_i\leqslant f_i\leqslant h_i$ almost everywhere for $i=1,2,\dots$, $g_i\to g$ and $h_i\to h$ almost everywhere as $i\to\infty$ and

$$\int_{\mathbb{R}^n} g\,dx=\lim_{i\to\infty}\int_{\mathbb{R}^n} g_i\,dx \quad\text{and}\quad \int_{\mathbb{R}^n} h\,dx=\lim_{i\to\infty}\int_{\mathbb{R}^n} h_i\,dx.$$

Then f is integrable and

$$\int_{\mathbb{R}^n} f\,dx=\lim_{i\to\infty}\int_{\mathbb{R}^n} f_i\,dx.$$

Proof. Exercise, see [4, Vol. 1, Theorem 2.8.8]. □

Theorem 2.6. Let $1\leqslant p<\infty$ and let $\Omega\subset\mathbb{R}^n$ be an open set. Assume that (u_i) and (v_i) are sequences of functions $u_i,v_i\in W^{1,p}(\Omega)$, $i=1,2,\dots$, such that $u_i\to u$ and $v_i\to v$ in $W^{1,p}(\Omega)$ as $i\to\infty$. Then $\min\{u_i,v_i\}\to\min\{u,v\}$ in $W^{1,p}(\Omega)$ as $i\to\infty$.

Proof. Let (u_i) be a sequence of functions $u_i\in C^\infty(\Omega)$, $i=1,2,\dots$, such that $u_i\to u$ in $W^{1,p}(\Omega)$ as $i\to\infty$ and let (v_i) be a sequence of functions $v_i\in C^\infty(\Omega)$, $i=1,2,\dots$, such that $v_i\to v$ in $W^{1,p}(\Omega)$ as $i\to\infty$. By passing to subsequences, we may assume that $Du_i\to Du$ and $Dv_i\to Dv$ almost everywhere in Ω as $i\to\infty$.

By Minkowski's inequality

$$\begin{aligned}
&\|\min\{u_i,v_i\}-\min\{u,v\}\|_{L^p(\Omega)}\\
&\quad\leqslant\|\min\{u_i,v_i\}-\min\{u,v_i\}\|_{L^p(\Omega)}+\|\min\{u,v_i\}-\min\{u,v\}\|_{L^p(\Omega)}\\
&\quad=I_1+I_2,\quad i=1,2,\dots.
\end{aligned}$$

We estimate I_1 and I_2 separately. For I_1 we have

$$\begin{aligned}
I_1^p &= \int_{\Omega\cap\{u\geqslant v_i\}} |\min\{u_i,v_i\}-v_i|^p\,dx + \int_{\Omega\cap\{u<v_i\}} |\min\{u_i,v_i\}-u|^p\,dx \\
&= \int_{\Omega\cap\{u\geqslant v_i\}\cap\{u_i\geqslant v_i\}} \underbrace{|v_i-v_i|^p}_{=0}\,dx + \int_{\Omega\cap\{u\geqslant v_i\}\cap\{u_i<v_i\}} |u_i-v_i|^p\,dx \\
&\quad + \int_{\Omega\cap\{u<v_i\}\cap\{u_i\geqslant v_i\}} |v_i-u|^p\,dx + \int_{\Omega\cap\{u<v_i\}\cap\{u_i<v_i\}} |u_i-u|^p\,dx \\
&\leqslant \int_{\Omega\cap\{u\geqslant v_i\}\cap\{u_i<v_i\}} (u-u_i)^p\,dx + \int_{\Omega\cap\{u<v_i\}\cap\{u_i\geqslant v_i\}} (u_i-u)^p\,dx \\
&\quad + \int_{\Omega\cap\{u<v_i\}\cap\{u_i<v_i\}} |u_i-u|^p\,dx \\
&\leqslant \int_{\Omega\cap\{u\geqslant v_i\}} |u-u_i|^p\,dx + \int_{\Omega\cap\{u<v_i\}} |u_i-u|^p\,dx \\
&\quad + \int_{\Omega\cap\{u<v_i\}\cap\{u_i<v_i\}} |u_i-u|^p\,dx \\
&\leqslant \int_{\Omega\cap\{u\geqslant v_i\}} |u-u_i|^p\,dx + \int_{\Omega\cap\{u<v_i\}} |u_i-u|^p\,dx \\
&= \int_{\Omega} |u-u_i|^p\,dx \xrightarrow{i\to\infty} 0.
\end{aligned}$$

For I_2, we have

$$\begin{aligned}
I_2^p &= \int_{\Omega\cap\{u\geqslant v_i\}} |v_i-\min\{u,v\}|^p\,dx + \int_{\Omega\cap\{u<v_i\}} |u-\min\{u,v\}|^p\,dx \\
&= \int_{\Omega\cap\{u\geqslant v_i\}\cap\{u\geqslant v\}} |v_i-v|^p\,dx + \int_{\Omega\cap\{u\geqslant v_i\}\cap\{u<v\}} |v_i-u|^p\,dx \\
&\quad + \int_{\Omega\cap\{u<v_i\}\cap\{u\geqslant v\}} |u-v|^p\,dx + \int_{\Omega\cap\{u<v_i\}\cap\{u<v\}} \underbrace{|u-u|^p}_{=0}\,dx \\
&\leqslant \int_{\Omega\cap\{u\geqslant v_i\}\cap\{u\geqslant v\}} |v_i-v|^p\,dx + \int_{\Omega\cap\{u\geqslant v_i\}\cap\{u<v\}} |v-v_i|^p\,dx \\
&\quad + \int_{\Omega\cap\{u<v_i\}\cap\{u\geqslant v\}} |v_i-v|^p\,dx \\
&\leqslant \int_{\Omega\cap\{u\geqslant v_i\}} |v_i-v|^p\,dx + \int_{\Omega\cap\{u<v_i\}} |v_i-v|^p\,dx \\
&= \int_{\Omega} |v-v_i|^p\,dx \xrightarrow{i\to\infty} 0.
\end{aligned}$$

This shows that $\min\{u_i,v_i\}\to\min\{u,v\}$ in $L^p(\Omega)$ as $i\to\infty$.

For the weak partial derivatives, we have

$$D_j\min\{u,v\} = \begin{cases} D_jv & \text{a.e. in } \Omega\cap\{u\geqslant v\}, \\ D_ju & \text{a.e. in } \Omega\cap\{u<v\}, \end{cases}$$

$j = 1,\dots,n$. By Minkowski's inequality

$$\begin{aligned}
&\|D_j \min\{u_i, v_i\} - D_j \min\{u, v\}\|_{L^p(\Omega)} \\
&\quad \le \|D_j \min\{u_i, v_i\} - D_j \min\{u, v_i\}\|_{L^p(\Omega)} + \|D_j \min\{u, v_i\} - D_j \min\{u, v\}\|_{L^p(\Omega)} \\
&\quad = J_1 + J_2, \quad i = 1,2,\dots, \quad j = 1,\dots,n.
\end{aligned}$$

For J_1, we have

$$\begin{aligned}
J_1^p &= \int_{\Omega \cap \{u \ge v_i\}} |D_j \min\{u_i, v_i\} - D_j v_i|^p \, dx + \int_{\Omega \cap \{u < v_i\}} |D_j \min\{u_i, v_i\} - D_j u|^p \, dx \\
&= K_1 + K_2.
\end{aligned}$$

For K_1, we have

$$\begin{aligned}
K_1 = &\int_\Omega |D_j u_i - D_j v_i|^p \chi_{\{u \ge v_i\}} \chi_{\{u_i < v_i\}} \, dx \\
&+ \int_\Omega \underbrace{|D_j v_i - D_j v_i|^p}_{=0} \chi_{\{u \ge v_i\}} \chi_{\{u_i \ge v_i\}} \, dx.
\end{aligned}$$

Since

$$\lim_{i\to\infty} |D_j u_i - D_j v_i|^p \chi_{\{u \ge v_i\}} \chi_{\{u_i < v_i\}} = 0$$

almost everywhere in Ω and

$$|D_j u_i - D_j v_i|^p \chi_{\{u \ge v_i\}} \chi_{\{u_i < v_i\}} \le |D_j u_i - D_j v_i|^p \xrightarrow{i\to\infty} |D_j u - D_j v|^p$$

in $L^1(\Omega)$, by Theorem 2.5, we have

$$\begin{aligned}
&\lim_{i\to\infty} \int_\Omega |D_j u_i - D_j v_i|^p \chi_{\{u \ge v_i\}} \chi_{\{u_i < v_i\}} \, dx \\
&\quad = \int_\Omega \lim_{i\to\infty} |D_j u_i - D_j v_i|^p \chi_{\{u \ge v_i\}} \chi_{\{u_i < v_i\}} \, dx = 0.
\end{aligned}$$

This shows that $K_1 \to 0$ as $i \to \infty$.

For K_2, we have

$$\begin{aligned}
K_2 = &\int_\Omega |D_j u_i - D_j u|^p \chi_{\{u < v_i\}} \chi_{\{u_i < v_i\}} \, dx \\
&+ \int_\Omega |D_j v_i - D_j u|^p \chi_{\{u < v_i\}} \chi_{\{u_i \ge v_i\}} \, dx,
\end{aligned}$$

where

$$\int_\Omega |D_j u_i - D_j u|^p \chi_{\{u < v_i\}} \chi_{\{u_i < v_i\}} \, dx \le \int_\Omega |D_j u_i - D_j u|^p \, dx \xrightarrow{i\to\infty} 0.$$

Since

$$\lim_{i\to\infty} |D_j v_i - D_j u|^p \chi_{\{u < v_i\}} \chi_{\{u_i \ge v_i\}} = 0$$

almost everywhere in Ω and

$$|D_j v_i - D_j u|^p \chi_{\{u < v_i\}} \chi_{\{u_i \ge v_i\}} \le |D_j v_i - D_j u|^p \xrightarrow{i\to\infty} |D_j v - D_j u|^p$$

in $L^1(\Omega)$, by Theorem 2.5, we have

$$\lim_{i\to\infty}\int_\Omega |D_j v_i - D_j u|^p \chi_{\{u<v_i\}}\chi_{\{u_i\geqslant v_i\}}\,dx$$
$$= \int_\Omega \lim_{i\to\infty} |D_j v_i - D_j u|^p \chi_{\{u<v_i\}}\chi_{\{u_i\geqslant v_i\}}\,dx = 0.$$

This shows that $K_2 \to 0$ as $i\to\infty$. It follows that $J_1 = K_1 + K_2 \to 0$ as $i\to\infty$.

For J_2, we have

$$J_2^p = \int_{\Omega\cap\{u\geqslant v_i\}} |D_j v_i - D_j \min\{u,v\}|^p\,dx + \int_{\Omega\cap\{u<v_i\}} |D_j u - D_j \min\{u,v\}|^p\,dx$$
$$= \int_\Omega |D_j v_i - D_j \min\{u,v\}|^p \chi_{\{u\geqslant v_i\}}\,dx + \int_\Omega |D_j u - D_j \min\{u,v\}|^p \chi_{\{u<v_i\}}\,dx.$$

Since

$$\lim_{i\to\infty} |D_j v_i - D_j \min\{u,v\}|^p \chi_{\{u\geqslant v_i\}} = 0$$

almost everywhere in Ω and

$$|D_j v_i - D_j \min\{u,v\}|^p \chi_{\{u\geqslant v_i\}} \leqslant |D_j v_i - D_j \min\{u,v\}|^p \xrightarrow{i\to\infty} |D_j v - D_j \min\{u,v\}|^p$$

in $L^1(\Omega)$, by Theorem 2.5, we have

$$\lim_{i\to\infty}\int_\Omega |D_j v_i - D_j \min\{u,v\}|^p \chi_{\{u\geqslant v_i\}}\,dx$$
$$= \int_\Omega \lim_{i\to\infty} |D_j v_i - D_j \min\{u,v\}|^p \chi_{\{u\geqslant v_i\}}\,dx = 0.$$

On other other hand, since

$$\lim_{i\to\infty} |D_j u - D_j \min\{u,v\}|^p \chi_{\{u<v_i\}} = 0$$

almost everywhere in Ω as $i\to\infty$ and

$$|D_j u - D_j \min\{u,v\}|^p \chi_{\{u<v_i\}} \leqslant |D_j u - D_j \min\{u,v\}|^p \in L^1(\Omega),$$

by the dominated convergence theorem

$$\lim_{i\to\infty}\int_\Omega |D_j u - D_j \min\{u,v\}|^p \chi_{\{u<v_i\}}\,dx$$
$$= \int_\Omega \lim_{i\to\infty} |D_j u - D_j \min\{u,v\}|^p \chi_{\{u<v_i\}}\,dx = 0.$$

It follows that $J_2 \to 0$ as $i\to\infty$.

Thus $\min\{u_i,v_i\} \to \min\{u,v\}$ and $D_j \min\{u_i,v_i\} \to D_j \min\{u,v\}$ in $L^p(\Omega)$ as $i\to\infty$, which implies that $\min\{u_i,v_i\}\to\min\{u,v\}$ in $W^{1,p}(\Omega)$ as $i\to\infty$.

Remark 2.7. We leave the proofs of the following results as exercises.

(1) The corresponding convergence result holds true for $\max\{u,v\}$ by a similar argument or by observing that $\max\{u,v\} = -\min\{-u,-v\}$.

(2) The corresponding convergence result holds true for u^+, u^- and $|u|$.

(3) If $u \in W_0^{1,p}(\Omega)$, then $u^+ = \max\{u,0\} \in W_0^{1,p}(\Omega)$, $u^- = -\min\{u,0\} \in W_0^{1,p}(\Omega)$, $|u| \in W_0^{1,p}(\Omega)$.

(4) If $u,v \in W_0^{1,p}(\Omega)$, then $\max\{u,v\} \in W_0^{1,p}(\Omega)$ and $\min\{u,v\} \in W_0^{1,p}(\Omega)$.

(5) Assume that $u \in W_0^{1,p}(\Omega)$. If $v \in W^{1,p}(\Omega)$ and $0 \leqslant v \leqslant u$ almost everywhere in Ω, then $v \in W_0^{1,p}(\Omega)$.

(6) Assume that $u \in W_0^{1,p}(\Omega)$. If $v \in W^{1,p}(\Omega)$ and $|v| \leqslant |u|$ almost everywhere in $\Omega \setminus K$, where K is a compact subset of Ω, then $v \in W_0^{1,p}(\Omega)$.

2.3 Weak convergence methods for Sobolev spaces

In this section we consider a function and its weak partial derivatives together and it is convenient to apply vector valued L^p spaces. Let $1 \leqslant p < \infty$ and let $\Omega \subset \mathbb{R}^n$ be an open set. Recall that $L^p(\Omega;\mathbb{R}^m)$ is the space of $\mathbb{R}^m$-valued functions $f : \Omega \to \mathbb{R}^m$, $f = (f_1, f_2, \dots, f_m)$ with $m \in \mathbb{N}$ for which

$$\|f\|_{L^p(\Omega;\mathbb{R}^m)} = \left(\sum_{j=1}^m \|f_j\|_{L^p(\Omega)}^p\right)^{\frac{1}{p}} < \infty.$$

It is clear that $f \in L^p(\Omega;\mathbb{R}^m)$ if and only if $f_j \in L^p(\Omega)$ for every $j = 1,2,\dots,m$. The norm above will be convenient for us, since if $f \in L^p(\Omega;\mathbb{R}^m)$, then

$$\|f\|_{L^p(\Omega;\mathbb{R}^m)} = \sup\left\{\int_\Omega f \cdot g\,dx : \|g\|_{L^{p'}(\Omega;\mathbb{R}^m)} = 1\right\},$$

where $f \cdot g = \sum_{j=1}^m f_j g_j$ is the Euclidean inner product.

Let $f \in L^p(\Omega;\mathbb{R}^m)$ and $g \in L^{p'}(\Omega;\mathbb{R}^m)$, where $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{p'} = 1$. By Hölder's inequality for functions and finite series, we have

$$\begin{aligned}\int_\Omega f \cdot g\,dx &\leqslant \left|\int_\Omega f \cdot g\,dx\right| = \left|\sum_{j=1}^m \int_\Omega f_j g_j\,dx\right| \leqslant \sum_{j=1}^m \int_\Omega |f_j||g_j|\,dx \\ &\leqslant \sum_{j=1}^m \|f_j\|_{L^p(\Omega)}\|g_j\|_{L^{p'}(\Omega)} \leqslant \left(\sum_{j=1}^m \|f_j\|_{L^p(\Omega)}^p\right)^{\frac{1}{p}} \left(\sum_{j=1}^m \|g_j\|_{L^{p'}(\Omega)}^{p'}\right)^{\frac{1}{p'}} \\ &= \|f\|_{L^p(\Omega;\mathbb{R}^m)}\|g\|_{L^{p'}(\Omega;\mathbb{R}^m)}.\end{aligned} \tag{2.8}$$

Hence, for every $g \in L^{p'}(\Omega;\mathbb{R}^m)$ with $\|g\|_{L^{p'}(\Omega;\mathbb{R}^m)} = 1$, we have

$$\int_\Omega f \cdot g\,dx \leqslant \|f\|_{L^p(\Omega;\mathbb{R}^m)}$$

and thus

$$\sup\left\{\int_\Omega f\cdot g\,dx : \|g\|_{L^{p'}(\Omega;\mathbb{R}^m)}=1\right\}\leqslant \|f\|_{L^p(\Omega;\mathbb{R}^m)}.$$

Next we show that the supremum above is attained, that is, there exists a function $g\in L^{p'}(\Omega;\mathbb{R}^m)$ with $\|g\|_{L^{p'}(\Omega;\mathbb{R}^m)}=1$ such that

$$\int_\Omega f\cdot g\,dx=\|f\|_{L^p(\Omega;\mathbb{R}^m)}.$$

Let

$$g_j=\frac{|f_j|^{\frac{p}{p'}}\operatorname{sgn} f_j}{\|f\|_{L^p(\Omega;\mathbb{R}^m)}^{\frac{p}{p'}}},\quad j=1,\dots,m.$$

Then

$$\begin{aligned}\|g_j\|_{L^{p'}(\Omega)}&=\left(\int_\Omega |g_j|^{p'}\,dx\right)^{\frac{1}{p'}}=\|f\|_{L^p(\Omega;\mathbb{R}^m)}^{-\frac{p}{p'}}\left(\int_\Omega |f_j|^p\,dx\right)^{\frac{1}{p'}}\\&=\|f\|_{L^p(\Omega;\mathbb{R}^m)}^{-\frac{p}{p'}}\|f_j\|_{L^p(\Omega)}^{\frac{p}{p'}},\quad j=1,\dots,m.\end{aligned}$$

Consequently

$$\begin{aligned}\|g\|_{L^{p'}(\Omega;\mathbb{R}^m)}&=\left(\sum_{j=1}^m\|g_j\|_{L^{p'}(\Omega)}^{p'}\right)^{\frac{1}{p'}}=\|f\|_{L^p(\Omega;\mathbb{R}^m)}^{-\frac{p}{p'}}\left(\sum_{j=1}^m\|f_j\|_{L^p(\Omega)}^p\right)^{\frac{1}{p'}}\\&=\|f\|_{L^p(\Omega;\mathbb{R}^m)}^{-\frac{p}{p'}+\frac{p}{p'}}=1\end{aligned}$$

and

$$\begin{aligned}\int_\Omega f\cdot g\,dx&=\sum_{j=1}^m\int_\Omega f_jg_j\,dx=\|f\|_{L^p(\Omega;\mathbb{R}^m)}^{-\frac{p}{p'}}\left(\sum_{j=1}^m\int_\Omega |f_j|^{1+\frac{p}{p'}}\,dx\right)\\&=\|f\|_{L^p(\Omega;\mathbb{R}^m)}^{-\frac{p}{p'}}\left(\sum_{j=1}^m\int_\Omega |f_j|^p\,dx\right)=\|f\|_{L^p(\Omega;\mathbb{R}^m)}^{-\frac{p}{p'}+p}=\|f\|_{L^p(\Omega;\mathbb{R}^m)}.\end{aligned}$$

This shows that

$$\|f\|_{L^p(\Omega;\mathbb{R}^m)}=\sup\left\{\int_\Omega f\cdot g\,dx : \|g\|_{L^{p'}(\Omega;\mathbb{R}^m)}=1\right\}.$$

The following version of the Riesz representation theorem will be useful.

Lemma 2.9. Let $1<p<\infty$. For every $L\in L^p(\Omega;\mathbb{R}^m)^*$ there exists a unique $g\in L^{p'}(\Omega;\mathbb{R}^m)$ such that

$$L(f)=\int_\Omega f\cdot g\,dx$$

for every $f\in L^p(\Omega;\mathbb{R}^m)$. Moreover, we have

$$\|L\|_{L^p(\Omega;\mathbb{R}^m)^*}=\|g\|_{L^{p'}(\Omega;\mathbb{R}^m)}.$$

Proof. Let $1 < p < \infty$ and let $L : L^p(\Omega;\mathbb{R}^{n+1}) \to \mathbb{R}$ be a bounded linear functional, that is, $L \in L^p(\Omega;\mathbb{R}^{n+1})^*$. Let $e_j = (0,\dots,1,\dots,0)$, $j = 0,1,\dots,n$, be the standard jth basis vector in $\mathbb{R}^{n+1}$. Then $L_j : L^p(\Omega) \to \mathbb{R}$,

$$L_j(f) = L(fe_j) = L((0,\dots,f,\dots,0)), \quad j = 0,1,\dots,n,$$

where f is in the jth slot, is a bounded linear functional on $L^p(\Omega)$, that is, $L_j \in L^p(\Omega)^*$, $j = 0,1,\dots,n$. To see this, we observe that

$$|L_j(f)| = |L(fe_j)| \leq \|L\|\|fe_j\|_{L^p(\Omega;\mathbb{R}^{n+1})} = \|L\|\|f\|_{L^p(\Omega)}$$

for every $f \in L^p(\Omega)$. The linearity of L_j, $j = 0,1,\dots,n$, follows immediately from the linearity of L.

By the Riesz representation theorem, there exists $g_j \in L^{p'}(\Omega)$ such that

$$L_j(f) = \int_\Omega f g_j\,dx, \quad j = 0,1,\dots,n,$$

for every $f \in L^p(\Omega)$.

Let $f = (f_0, f_1, \dots, f_n) = \sum_{j=0}^n f_j e_j$. Since L is linear, we have

$$\begin{aligned} L(f) &= L\left(\sum_{j=0}^n f_j e_j\right) = \sum_{j=0}^n L(f_j e_j) = \sum_{j=0}^n L_j(f_j) \\ &= \sum_{j=0}^n \int_\Omega f_j g_j\,dx = \int_\Omega \sum_{j=0}^n f_j g_j\,dx = \int_\Omega f \cdot g\,dx. \end{aligned}$$

By Hölder's inequality for functions and finite series, we have

$$\begin{aligned} \left|\int_\Omega f \cdot g\,dx\right| &= \left|\sum_{j=0}^n \int_\Omega f_j g_j\,dx\right| \leq \sum_{j=0}^n \int_\Omega |f_j||g_j|\,dx \\ &\leq \sum_{j=0}^n \|f_j\|_{L^p(\Omega)}\|g_j\|_{L^{p'}(\Omega)} \\ &\leq \left(\sum_{j=0}^n \|f_j\|^p_{L^p(\Omega)}\right)^{\frac{1}{p}} \left(\sum_{j=0}^n \|g_j\|^{p'}_{L^{p'}(\Omega)}\right)^{\frac{1}{p'}} \\ &= \|f\|_{L^p(\Omega;\mathbb{R}^{n+1})}\|g\|_{L^{p'}(\Omega;\mathbb{R}^{n+1})}. \end{aligned}$$

Hence

$$\begin{aligned} \|L\| &= \sup\left\{|L(f)| : \|f\|_{L^p(\Omega;\mathbb{R}^{n+1})} \leq 1\right\} \\ &= \sup\left\{\left|\int_\Omega f \cdot g\,dx\right| : \|f\|_{L^p(\Omega;\mathbb{R}^{n+1})} \leq 1\right\} \leq \|g\|_{L^{p'}(\Omega;\mathbb{R}^{n+1})}. \end{aligned}$$

On the other hand, let $f_j = |g_j|^{p'-1}\operatorname{sgn} g_j$, $j = 0,1,\dots,n$. Then $|f_j|^p = |g_j|^{p'}$ and

$$\begin{aligned} \|f\|_{L^p(\Omega;\mathbb{R}^{n+1})} &= \left(\sum_{j=0}^n \|f_j\|^p_{L^p(\Omega)}\right)^{\frac{1}{p}} \\ &= \left(\sum_{j=0}^n \|g_j\|^{p'}_{L^{p'}(\Omega)}\right)^{\frac{1}{p}} = \|g\|^{\frac{p'}{p}}_{L^{p'}(\Omega;\mathbb{R}^{n+1})} < \infty. \end{aligned}$$

It follows that

$$\begin{aligned}
\|g\|^{p'}_{L^{p'}(\Omega;\mathbb{R}^{n+1})} &= \sum_{j=0}^{n} \int_{\Omega} |g_j|^{p'}\,dx = \sum_{j=0}^{n} \int_{\Omega} |g_j|^{p'-1} g_j \operatorname{sgn} g_j\,dx \\
&= \sum_{j=0}^{n} \int_{\Omega} f_j g_j\,dx = \int_{\Omega} f \cdot g\,dx \\
&= L(f) \leqslant \|L\| \|f\|_{L^p(\Omega;\mathbb{R}^{n+1})} = \|L\| \|g\|^{\frac{p'}{p}}_{L^{p'}(\Omega;\mathbb{R}^{n+1})}.
\end{aligned}$$

This implies that

$$\|L\| \geqslant \|g\|^{p'-\frac{p'}{p}}_{L^{p'}(\Omega;\mathbb{R}^{n+1})} = \|g\|_{L^{p'}(\Omega;\mathbb{R}^{n+1})}.$$

Hence $\|L\| = \|g\|_{L^{p'}(\Omega;\mathbb{R}^{n+1})}$.

To show the uniqueness, we assume that there exist $g, h \in L^{p'}(\Omega;\mathbb{R}^{n+1})$ such that

$$L(f) = \int_{\Omega} f \cdot g\,dx \quad \text{and} \quad L(f) = \int_{\Omega} f \cdot h\,dx$$

for every $f \in L^p(\Omega;\mathbb{R}^{n+1})$. Then

$$\int_{\Omega} (f \cdot g - f \cdot h)\,dx = \int_{\Omega} f \cdot (g-h)\,dx = 0$$

for every $f \in L^p(\Omega;\mathbb{R}^{n+1})$. Let

$$f_j = |g_j - h_j|^{p'-1} \operatorname{sgn}(g_j - h_j), \quad j = 0, 1, \ldots, n.$$

Then $|f_j|^p = |g_j - h_j|^{p'}$ and

$$\begin{aligned}
\|f\|_{L^p(\Omega;\mathbb{R}^{n+1})} &= \left(\sum_{j=0}^{n} \|f_j\|^p_{L^p(\Omega)} \right)^{\frac{1}{p}} = \left(\sum_{j=0}^{n} \|g_j - h_j\|^{p'}_{L^{p'}(\Omega)} \right)^{\frac{1}{p}} \\
&= \|g-h\|^{\frac{p'}{p}}_{L^{p'}(\Omega;\mathbb{R}^{n+1})} < \infty.
\end{aligned}$$

Thus

$$\begin{aligned}
0 &= \int_{\Omega} f \cdot (g-h)\,dx = \int_{\Omega} \sum_{j=0}^{n} f_j (g_j - h_j)\,dx \\
&= \int_{\Omega} \sum_{j=0}^{n} |g_j - h_j|^{p'}\,dx = \sum_{j=0}^{n} \int_{\Omega} |g_j - h_j|^{p'}\,dx
\end{aligned}$$

which implies that $g_j = h_j$ for every $j = 0, 1, \ldots, n$. □

This section discusses weak convergence techniques for $L^p(\Omega;\mathbb{R}^m)$ even though most of the results hold for more general Banach spaces as well.

Definition 2.10. Let $1 < p < \infty$ and $m \in \mathbb{N}$, and let $\Omega \subset \mathbb{R}^n$ be an open set. A sequence $(f_i)_{i\in\mathbb{N}}$ of functions in $L^p(\Omega;\mathbb{R}^m)$ converges weakly in $L^p(\Omega;\mathbb{R}^m)$ to a function $f \in L^p(\Omega;\mathbb{R}^m)$, if

$$\lim_{i\to\infty} \int_{\Omega} f_i \cdot g\,dx = \int_{\Omega} f \cdot g\,dx$$

for every $g \in L^{p'}(\Omega;\mathbb{R}^m)$ with $p' = \frac{p}{p-1}$.

Next we show that weakly convergent sequences are bounded and that the L^p norm is lower semicontinuous with respect to the weak convergence.

Lemma 2.11. Let $1 < p < \infty$ and $m \in \mathbb{N}$, and let $\Omega \subset \mathbb{R}^n$ be an open set. If a sequence $(f_i)_{i\in\mathbb{N}}$ converges to f weakly in $L^p(\Omega;\mathbb{R}^m)$, then $(f_i)_{i\in\mathbb{N}}$ is a bounded sequence in $L^p(\Omega;\mathbb{R}^m)$. Moreover, we have

$$\|f\|_{L^p(\Omega;\mathbb{R}^m)} \leqslant \liminf_{i\to\infty} \|f_i\|_{L^p(\Omega;\mathbb{R}^m)}. \tag{2.12}$$

Proof. (1) First we show that $(f_i)_{i\in\mathbb{N}}$ is a bounded sequence in $L^p(\Omega;\mathbb{R}^m)$. Let $g \in L^{p'}(\Omega;\mathbb{R}^m)$ with $\|g\|_{L^{p'}(\Omega;\mathbb{R}^m)} = 1$. The definition of weak convergence and Hölder's inequality for functions and finite series imply

$$\begin{aligned}
\lim_{i\to\infty}\int_\Omega f_i \cdot g\,dx = \lim_{i\to\infty}\int_\Omega f_i \cdot g\,dx &= \int_\Omega f\cdot g\,dx \leqslant \left|\int_\Omega f\cdot g\,dx\right| \\
&\leqslant \int_\Omega |f\cdot g|\,dx \leqslant \sum_{j=1}^m \int_\Omega |f^{(j)}||g^{(j)}|\,dx \\
&\leqslant \sum_{j=1}^m \|f_i^{(j)}\|_{L^p(\Omega)}\|g^{(j)}\|_{L^{p'}(\Omega)} \\
&\leqslant \left(\sum_{j=1}^m \|f^{(j)}\|^p_{L^p(\Omega)}\right)^{\frac{1}{p}} \left(\sum_{j=1}^m \|g^{(j)}\|^{p'}_{L^{p'}(\Omega)}\right)^{\frac{1}{p'}} \\
&= \|f\|_{L^p(\Omega;\mathbb{R}^m)}\|g\|_{L^{p'}(\Omega;\mathbb{R}^m)} = \|f\|_{L^p(\Omega;\mathbb{R}^m)}.
\end{aligned}$$

Here we denote the coordinate functions as $g = (g^{(1)}, g^{(2)}, \dots, g^{(m)})$.

Let $g_i \in L^{p'}(\Omega;\mathbb{R}^m)$ with $\|g_i\|_{L^{p'}(\Omega;\mathbb{R}^m)} = 1$ and

$$\|f_i\|_{L^p(\Omega;\mathbb{R}^m)} = \int_\Omega f_i \cdot g_i\,dx, \quad i = 1,2,\dots.$$

By the estimate above we have

$$\limsup_{i\to\infty}\|f_i\|_{L^p(\Omega;\mathbb{R}^m)} = \limsup_{i\to\infty}\int_\Omega f_i\cdot g_i\,dx \leqslant \|f\|_{L^p(\Omega;\mathbb{R}^m)} < \infty$$

which implies that

$$\sup_{i\in\mathbb{N}}\|f_i\|_{L^p(\Omega;\mathbb{R}^m)} < \infty.$$

(2) In order to prove (2.12), let $g \in L^{p'}(\Omega;\mathbb{R}^m)$ with $\|g\|_{L^{p'}(\Omega;\mathbb{R}^m)} = 1$ and

$$\|f\|_{L^p(\Omega;\mathbb{R}^m)} = \int_\Omega f\cdot g\,dx.$$

The definition of weak convergence and Hölder's inequality for functions and finite series as above imply

$$\begin{aligned}
\|f\|_{L^p(\Omega;\mathbb{R}^m)} &= \int_\Omega f\cdot g\,dx = \lim_{i\to\infty}\int_\Omega f_i\cdot g\,dx\\
&\leqslant \liminf_{i\to\infty}\sum_{j=1}^m \int_\Omega |f_i^{(j)}||g^{(j)}|\,dx\\
&\leqslant \liminf_{i\to\infty}\sum_{j=1}^m \|f_i^{(j)}\|_{L^p(\Omega)}\|g^{(j)}\|_{L^{p'}(\Omega)}\\
&\leqslant \liminf_{i\to\infty}\left(\sum_{j=1}^m \|f_i^{(j)}\|_{L^p(\Omega)}^p\right)^{\frac{1}{p}}\left(\sum_{j=1}^m \|g^{(j)}\|_{L^{p'}(\Omega)}^{p'}\right)^{\frac{1}{p'}}\\
&= \liminf_{i\to\infty}\|f_i\|_{L^p(\Omega;\mathbb{R}^m)}\|g\|_{L^{p'}(\Omega;\mathbb{R}^m)}\\
&= \liminf_{i\to\infty}\|f_i\|_{L^p(\Omega;\mathbb{R}^m)}.
\end{aligned}$$

□

THE MORAL: The L^p-norm is lower semicontinuous with respect to the weak convergence.

A bounded sequence in $L^p(\Omega;\mathbb{R}^m)$ need not have a convergent subsequence. However, the following result shows that it always has a weakly convergent subsequence if $1 < p < \infty$. This will be important in our applications of weak convergence. The following result holds, since $L^p(\Omega;\mathbb{R}^m)$ is reflexive and separable when $1 < p < \infty$. Theorem 2.13 does not hold for $p = 1$. This can be seen by considering the standard mollifier that approximates the Dirac's delta.

Theorem 2.13. Let $1 < p < \infty$ and $m \in \mathbb{N}$, and let $\Omega \subset \mathbb{R}^n$ be an open set. Assume that $(f_i)_{i\in\mathbb{N}}$ is a bounded sequence in $L^p(\Omega;\mathbb{R}^m)$. There exists a subsequence $(f_{i_k})_{k\in\mathbb{N}}$ and a function $f \in L^p(\Omega;\mathbb{R}^m)$ such that $f_{i_k} \to f$ weakly in $L^p(\Omega;\mathbb{R}^m)$ as $k \to \infty$.

THE MORAL: This shows that L^p with $1 < p < \infty$ is weakly sequentially compact, that is, every bounded sequence in L^p has a weakly converging subsequence. One of the most useful applications of weak convergence is in compactness arguments. A bounded sequence in L^p does not need to have any convergent subsequence with convergence interpreted in the standard L^p sense. However, there exists a weakly converging subsequence.

Remark 2.14. Theorem 2.13 is equivalent to the fact that L^p spaces are reflexive for $1 < p < \infty$.

Weak convergence is often too weak mode of convergence and we need tools to upgrade it to stronger modes of convergence. We begin with the following result, which is related to Lemma 2.11. The next result holds, since $L^p(\Omega;\mathbb{R}^m)$ is a uniformly convex Banach space.

Lemma 2.15. Let $1 < p < \infty$ and $m \in \mathbb{N}$, and let $\Omega \subset \mathbb{R}^n$ be an open set. Assume that a sequence $(f_i)_{i\in\mathbb{N}}$ converges to f weakly in $L^p(\Omega;\mathbb{R}^m)$ and

$$\limsup_{i\to\infty} \|f_i\|_{L^p(\Omega;\mathbb{R}^m)} \leqslant \|f\|_{L^p(\Omega;\mathbb{R}^m)}. \tag{2.16}$$

Then $f_i \to f$ in $L^p(\Omega;\mathbb{R}^m)$ as $i \to \infty$.

Observe that, under the assumptions in Lemma 2.15, by (2.12) and (2.16) we have

$$\|f\|_{L^p(\Omega;\mathbb{R}^m)} \leqslant \liminf_{i\to\infty} \|f_i\|_{L^p(\Omega;\mathbb{R}^m)} \leqslant \limsup_{i\to\infty} \|f_i\|_{L^p(\Omega;\mathbb{R}^m)} \leqslant \|f\|_{L^p(\Omega;\mathbb{R}^m)},$$

which implies

$$\lim_{i\to\infty} \|f_i\|_{L^p(\Omega;\mathbb{R}^m)} = \|f\|_{L^p(\Omega;\mathbb{R}^m)}.$$

This means that the limit exists with an equality in (2.16).

THE MORAL: Weak convergence in $L^p(\Omega;\mathbb{R}^m)$ with $1 < p < \infty$ can be upgraded to strong convergence if $\|f_i\|_{L^p(\Omega;\mathbb{R}^m)} \to \|f\|_{L^p(\Omega;\mathbb{R}^m)}$ as $i \to \infty$.

Next we discuss another method to upgrade weak convergence to strong convergence. Mazur's lemma below asserts that a convex and closed subspace of a reflexive Banach space is weakly closed.

Theorem 2.17 (Mazur's lemma). Assume that X is a normed space and that $x_i \to x$ weakly in X as $i \to \infty$. Then there exists a sequence of convex combinations $\widetilde{x}_i = \sum_{j=i}^{m_i} a_{i,j} x_j$, with $a_{i,j} \geqslant 0$ and $\sum_{j=i}^{m_i} a_{i,j} = 1$, such that $\widetilde{x}_i \to x$ in the norm of X as $i \to \infty$.

THE MORAL: For every weakly converging sequence, there is a sequence of convex combinations that converges strongly. Thus weak convergence is upgraded to strong convergence for a sequence of convex combinations. Observe that some of the coefficients a_i may be zero so that the convex combination is essentially for a subsequence.

Remark 2.18. Since $L^p(\Omega;\mathbb{R}^m)$ is a uniformly convex Banach space, the Banach–Saks theorem asserts that a weakly convergent sequence has a subsequence whose arithmetic means converge in the norm. Let $1 < p < \infty$ and $m \in \mathbb{N}$, and let $\Omega \subset \mathbb{R}^n$ be an open set. Assume that a sequence $(f_i)_{i\in\mathbb{N}}$ converges to f weakly in $L^p(\Omega;\mathbb{R}^m)$ as $i \to \infty$. Then there exists a subsequence $(f_{i_k})_{k\in\mathbb{N}}$ for which the arithmetic mean $\frac{1}{k}\sum_{j=1}^{k} f_{i_j}$ converges to f in $L^p(\Omega;\mathbb{R}^m)$ as $k \to \infty$. The advantage of the Banach–Saks theorem compared to Mazur's lemma is that we can work with the arithmetic means instead of more general convex combinations

Remark 2.19. Mazur's lemma can be used to give a proof for (2.12) (exercise).

Theorem 2.20. Let $1 < p < \infty$. Assume that (u_i) is a bounded sequence in $W^{1,p}(\Omega)$. There exists a subsequence (u_{i_k}) and $u \in W^{1,p}(\Omega)$ such that $u_{i_k} \to u$ weakly in $L^p(\Omega)$ and $Du_{i_k} \to Du$ weakly in $L^p(\Omega)$ as $k \to \infty$. Moreover, if $u_i \in W_0^{1,p}(\Omega)$, $i = 1, 2\dots$, then $u \in W_0^{1,p}(\Omega)$.

THE MORAL: This shows that $W^{1,p}$ with $1 < p < \infty$ is weakly sequentially compact, that is, every bounded sequence in $W^{1,p}$ with $1 < p < \infty$ has a weakly converging subsequence. Note that there may exist several weakly converging subsequences and the limit may depend on the subsequence.

Proof. (1) Assume that $u \in W^{1,p}(\Omega)$. Denote

$$f_i = (u_i, Du_i) \in L^p(\Omega; \mathbb{R}^{n+1}), \quad i \in \mathbb{N}.$$

Then $(f_i)_{i\in\mathbb{N}}$ is a bounded sequence in $L^p(\Omega;\mathbb{R}^{n+1})$. By Theorem 2.13, there exist a subsequence $(f_{i_k})_{k\in\mathbb{N}}$ and a function $f \in L^p(\Omega;\mathbb{R}^{n+1})$ such that $f_{i_k} \to f$ weakly in $L^p(\Omega;\mathbb{R}^{n+1})$ as $k \to \infty$. Let $f = (u, v)$ with $u \in L^p(\Omega)$ and $v = (v_1, \dots, v_n) \in L^p(\Omega;\mathbb{R}^n)$. We claim that v is the weak gradient of u, that is, $v = Du$ in Ω.

(2) By using test functions of the form $(g_1, 0, \dots, 0)$ or $(0, g_2, \dots, g_{n+1})$ in the definition of weak convergence, we conclude that $u_{i_k} \to u$ weakly in $L^p(\Omega)$ and $Du_{i_k} \to v$ weakly in $L^p(\Omega;\mathbb{R}^n)$ as $k \to \infty$. Since $u_{i_k} \to u$ weakly in $L^p(\Omega)$ as $k \to \infty$ and $u_{i_k} \in W^{1,p}(\Omega)$, we have

$$\begin{aligned}\int_\Omega u D_j\varphi\,dx &= \lim_{k\to\infty}\int_\Omega u_{i_k} D_j\varphi\,dx \\ &= -\lim_{k\to\infty}\int_\Omega D_j u_{i_k}\varphi\,dx, \quad j = 1, \dots, n,\end{aligned}$$

for every $\varphi \in C_0^\infty(\Omega)$. On the other hand, since $Du_{i_k} \to v$ weakly in $L^p(\Omega;\mathbb{R}^n)$, by using the test function $(0, \dots, \varphi, \dots, 0) \in L^{p'}(\Omega;\mathbb{R}^n)$, where φ is in the jth position, we have

$$\lim_{k\to\infty}\int_\Omega D_j u_{i_k}\varphi\,dx = \int_\Omega v_j\varphi\,dx, \quad j = 1, \dots, n.$$

It follows that

$$\int_\Omega u D_j\varphi\,dx = -\int_\Omega v_j\varphi\,dx, \quad j = 1, \dots, n,$$

for every $\varphi \in C_0^\infty(\Omega)$. This shows that $D_j u = v_j$, $j = 1, \dots, n$. In particular, the weak partial derivatives $D_j u = v_j$, $j = 1, \dots, n$, exist. Moreover, $D_j u \in L^p(\Omega)$, $j = 1, \dots, n$, since $v = (v_1, \dots, v_n) \in L^p(\Omega;\mathbb{R}^n)$. It follows that $u \in W^{1,p}(\Omega)$. The argument above also shows that (u_{i_k}, Du_{i_k}) converges to (u, Du) weakly in $L^p(\Omega;\mathbb{R}^{n+1})$ as $k \to \infty$.

(3) Assume that $u_i \in W_0^{1,p}(\Omega)$ for every $i \in \mathbb{N}$ and that the sequence

$$(f_{i_k})_{k\in\mathbb{N}} = ((u_{i_k}, Du_{i_k}))_{k\in\mathbb{N}}$$

converges weakly to $f = (u, Du)$ in $L^p(\Omega;\mathbb{R}^{n+1})$. By Theorem 2.17, there exists a sequence of convex combinations

$$h_k = \sum_{j=k}^{m_k} a_{k,j} f_{i_j} = \sum_{j=k}^{m_k} a_{k,j}(u_{i_j}, Du_{i_j})$$

that converges to $f = (u, Du)$ in $L^p(\Omega;\mathbb{R}^{n+1})$ as $k \to \infty$. This implies

$$\sum_{j=k}^{m_k} a_{k,j} u_{i_j} \xrightarrow{k\to\infty} u \quad \text{and} \quad \sum_{j=k}^{m_k} a_{k,j} Du_{i_j} \xrightarrow{k\to\infty} Du$$

in $L^p(\Omega)$ and thus

$$\sum_{j=k}^{m_k} a_{k,j} u_{i_j} \xrightarrow{k\to\infty} u$$

in $W^{1,p}(\Omega)$. Moreover,

$$\sum_{j=k}^{m_k} a_{k,j} u_{i_j} \in W_0^{1,p}(\Omega)$$

for every $k \in \mathbb{N}$. Since $W_0^{1,p}(\Omega)$ is a closed subspace of $W^{1,p}(\Omega)$, it follows that $u \in W_0^{1,p}(\Omega)$. □

Remarks 2.21:

(1) Theorem 2.20 is equivalent to the fact that $W^{1,p}$ spaces are reflexive for $1 < p < \infty$.

(2) Since $u_{i_k} \to u$ weakly in $L^p(\Omega)$ and $Du_{i_k} \to Du$ weakly in $L^p(\Omega)$ as $k \to \infty$ in Theorem 2.20, by Lemma 2.11 we have

$$\|u\|_{W^{1,p}(\Omega)} \leqslant \liminf_{k\to\infty} \|u_{i_k}\|_{W^{1,p}(\Omega)}.$$

Thus the $W^{1,p}$-norm is lower semicontinuous with respect to the weak convergence in $W^{1,p}$.

(3) Another way to see that $W^{1,p}$ spaces are reflexive for $1 < p < \infty$ is to recall that a closed subspace of a reflexive space is reflexive. Thus it is enough to find an isomorphism between $W^{1,p}(\Omega)$ and a closed subspace of $L^p(\Omega,\mathbb{R}^{n+1}) = L^p(\Omega,\mathbb{R}^n) \times \cdots \times L^p(\Omega,\mathbb{R}^n)$. The mapping $u \mapsto (u, Du)$ will do for this purpose. This holds true for $W_0^{1,p}(\Omega)$ as well. This approach can be used to characterize elements in the dual space by the Riesz representation theorem, see [2, p. 62–65], [14, Section 11.4], [18, Section 4.3].

Example 2.22. The Sobolev space is not compact in the sense that every bounded sequence (u_i) in $W^{1,p}(\Omega)$ has a converging subsequence (u_{i_k}) and $u \in W^{1,p}(\Omega)$ such that $u_{i_k} \to u$ in $W^{1,p}(\Omega)$. For $i = 1,2,\dots$, consider $u_i : (0,2) \to \mathbb{R}$,

$$u_i(x) = \begin{cases} 0, & 0 < x \leqslant 1, \\ (x-1)i, & 1 \leqslant x \leqslant 1 + \frac{1}{i}, \\ 1, & 1 + \frac{1}{i} < x < 2. \end{cases}$$

Then $u_i \in W^{1,1}((0,2))$ and $\|u_i\|_{W^{1,1}((0,2))} \leqslant 2$ for every $i = 1,2,\dots$. However, there does not exist a subsequence that converges in $W^{1,1}((0,2))$. To conclude this, assume that there exists a subsequence (u_{i_k}) that converges in $W^{1,1}((0,2))$. In particular, the subsequence (u_{i_k}) converges in $L^1((0,2))$ and the limit function $u \in L^1((0,2))$ is

$$u(x) = \begin{cases} 0, & 0 < x \leqslant 1, \\ 1, & 1 < x < 2. \end{cases}$$

However, $u \notin W^{1,1}((0,2))$. This example also shows that Theorem 2.20 does not hold true in the case $p = 1$ (exercise).

Example 2.23. For $i = 1,2,\dots$, consider $u_i : (0,2) \to \mathbb{R}$,

$$u_i(x) = \begin{cases} 0, & 0 < x \leqslant 1, \\ (x-1)\sqrt{i}, & 1 \leqslant x \leqslant 1 + \frac{1}{i}, \\ \frac{1}{\sqrt{i}}, & 1 + \frac{1}{i} < x < 2. \end{cases}$$

Then $u_i \in W^{1,2}((0,2))$,

$$\|u_i\|^2_{L^2((0,2))} = \frac{1}{3i^2} + \frac{i-1}{i^2} = \frac{3i-2}{3i^2}, \quad \|Du_i\|^2_{L^2((0,2))} = 1,$$

for every $i = 1,2,\dots$ and $u_i \to u$ weakly in $W^{1,2}(\Omega)$ as $i \to \infty$, where $u = 0$ (exercise). Clearly

$$0 = \|u\|_{W^{1,2}((0,2))} < 1 \leqslant \liminf_{i\to\infty} \|u_i\|_{W^{1,2}((0,2))}.$$

This shows that norm is only lower semicontinous in the weak topology but not continuous. Observe carefully that $u_i \nrightarrow u$ in $W^{1,2}((0,2))$, since

$$\lim_{i\to\infty} \|Du_i\|^2_{L^2((0,2))} = 1 \neq 0.$$

For the proof of the next result we briefly discuss the convergence of a sequence of real numbers. Let $(a_i)_{i\in\mathbb{N}}$ be a sequence of real numbers and let $a \in \mathbb{R}$. It is easy to show that $a_i \to a$ as $i \to \infty$ if and only if every every subsequence $(a_{i_k})_{k\in\mathbb{N}}$ converges to a as $k \to \infty$. We state a refinement of this result.

Claim: $a_i \to a$ as $i \to \infty$ if and only if every subsequence $(a_{i_k})_{k\in\mathbb{N}}$ has a subsequence, denoted by $(a_{i_j})_{j\in\mathbb{N}}$, such that $a_{i_j} \to a$ as $j \to \infty$.

Reason. $\boxed{\Longrightarrow}$ Assume that $a_i \to a$ as $i \to \infty$. Then every subsequence $(a_{i_k})_{k\in\mathbb{N}}$ converges to a and we may take $(a_{i_j})_{j\in\mathbb{N}}$ be the subsequence $(a_{i_k})_{k\in\mathbb{N}}$ itself.

$\boxed{\Longleftarrow}$ Assume that every subsequence $(a_{i_k})_{k\in\mathbb{N}}$ has a subsequence, denoted by $(a_{i_j})_{j\in\mathbb{N}}$, such that $a_{i_j} \to a$ as $j \to \infty$. For a contraction, assume that a_i does not converge to a as $i \to \infty$. Then there exists $\varepsilon > 0$ such that for every $k \in \mathbb{N}$ there exists $i_k \geqslant k$ such that $|a_{i_k} - a| \geqslant \varepsilon$. Let $(a_{i_j})_{j\in\mathbb{N}}$ be a subsequence of $(a_{i_k})_{k\in\mathbb{N}}$. Then $|a_{i_j} - a| \geqslant \varepsilon$ for every $j \in \mathbb{N}$. In particular, there does not exist $n \in \mathbb{N}$ such that $|a_{i_j} - a| < \varepsilon$ for every $j \geqslant n$. This implies that $(a_{i_j})_{j\in\mathbb{N}}$ does not converge to a as $j \to \infty$. Thus $(a_{i_k})_{k\in\mathbb{N}}$ is a subsequence of $(a_i)_{i\in\mathbb{N}}$, but its every subsequence $(a_{i_j})_{j\in\mathbb{N}}$ does not converge to a as $j \to \infty$. This contradicts the assumption on the subsequences of $(a_i)_{i\in\mathbb{N}}$. It follows that $a_i \to a$ as $i \to \infty$ ■

Theorem 2.24. Let $1 < p < \infty$ and let $\Omega \subset \mathbb{R}^n$ be an open set. Assume that $(u_i)_{i\in\mathbb{N}}$ is a bounded sequence in $W^{1,p}(\Omega)$ such that $u_i \to u$ weakly in $L^p(\Omega)$ as $i \to \infty$ or that $u_i \to u$ almost everywhere in Ω as $i \to \infty$. Then $u \in W^{1,p}(\Omega)$, $u_i \to u$ weakly in $L^p(\Omega)$, and $Du_i \to Du$ weakly in $L^p(\Omega;\mathbb{R}^n)$ as $i \to \infty$. Moreover, if $u_i \in W_0^{1,p}(\Omega)$ for every $i \in \mathbb{N}$, then $u \in W_0^{1,p}(\Omega)$.

T H E M O R A L : Pointwise convergence implies weak convergence for a bounded sequence in $W^{1,p}$ (or L^p) with $1 < p < \infty$. In order to show that $u \in W^{1,p}(\Omega)$ it is enough to construct functions $u_i \in W^{1,p}(\Omega)$, $i = 1,2,\dots$, such that $u_i \to u$ almost everywhere in Ω as $i \to \infty$ and $\sup_i \|u_i\|_{W^{1,p}(\Omega)} < \infty$.

Proof. (1) We prove that $(u_i, Du_i) \to (u, Du)$ weakly in $L^p(\Omega;\mathbb{R}^{n+1})$ as $i \to \infty$ by showing that every subsequence $(u_{i_k})_{k\in\mathbb{N}}$ has a subsequence $(u_{i_j})_{j\in\mathbb{N}}$ such that $(u_{i_j}, Du_{i_j}) \to (u, Du)$ weakly in $L^p(\Omega;\mathbb{R}^{n+1})$ as $j \to \infty$.

To see this assume that every subsequence $(u_{i_k})_{k\in\mathbb{N}}$ has a subsequence $(u_{i_j})_{j\in\mathbb{N}}$ such that $(u_{i_j}, Du_{i_j}) \to (u, Du)$ weakly in $L^p(\Omega;\mathbb{R}^{n+1})$ as $j \to \infty$. We claim that $(u_i, Du_i) \to (u, Du)$ weakly in $L^p(\Omega;\mathbb{R}^{n+1})$ as $i \to \infty$. For a contradiction, assume that (u_i, Du_i) does not converge to (u, Du) weakly in $L^p(\Omega;\mathbb{R}^{n+1})$ as $i \to \infty$. Then there exists $g \in L^{p'}(\Omega;\mathbb{R}^{n+1})$ such that the sequence of real numbers

$$a_i = \int_\Omega (u_i, Du_i)\cdot g\,dx, \quad i = 1,2,\dots,$$

does not converge to

$$a = \int_\Omega (u, Du)\cdot g\,dx$$

as $i \to \infty$. Then $(a_i)_{i\in\mathbb{N}}$ has a subsequence $(a_{i_k})_{k\in\mathbb{N}}$ such that its all subsequences $(a_{i_j})_{k\in\mathbb{N}}$ with

$$a_{i_j} = \int_\Omega (u_{i_j}, Du_{i_j})\cdot g\,dx, \quad j = 1,2,\dots,$$

fail to converge to a. This is a contradiction with the assumption that every subsequence $(u_{i_k})_{k\in\mathbb{N}}$ has a subsequence $(u_{i_j})_{j\in\mathbb{N}}$ such that $(u_{i_j}, Du_{i_j}) \to (u, Du)$ weakly in $L^p(\Omega;\mathbb{R}^{n+1})$ as $j \to \infty$.

Let $(u_{i_k})_{k\in\mathbb{N}}$ be a subsequence of $(u_i)_{i\in\mathbb{N}}$. By Theorem 2.20, there exists a subsequence, again denoted by $(u_{i_k})_{k\in\mathbb{N}}$, and a function $v \in W^{1,p}(\Omega)$ such that

$$(u_{i_k}, Du_{i_k}) \xrightarrow{k\to\infty} (v, Dv)$$

weakly in $L^p(\Omega;\mathbb{R}^{n+1})$. We claim that $u = v$ almost everywhere in Ω.

If $u_i \to u$ weakly in $L^p(\Omega)$, then $u_{i_k} \to u$ weakly in $L^p(\Omega)$ and $u = v$ almost everywhere by the uniqueness of weak limit. This implies that $u \in W^{1,p}(\Omega)$.

It remains to consider the case $u_i \to u$ almost everywhere in Ω as $i \to \infty$. By Theorem 2.17, there exists a sequence of convex combinations

$$h_k = \sum_{j=k}^{m_k} a_{k,j}(u_{i_j}, Du_{i_j})$$

that converges to (v, Dv) in $L^p(\Omega;\mathbb{R}^{n+1})$ as $k\to\infty$. In particular,

$$h_{k,1} = \sum_{j=k}^{m_k} a_{k,j} u_{i_j} \xrightarrow{k\to\infty} v$$

in $L^p(\Omega)$. This implies that there exists a subsequence of $(h_{k,1})_{k\in\mathbb{N}}$ that converges to v almost everywhere in Ω. Since $u_{i_j}\to u$ almost everywhere in Ω as $j\to\infty$, we have

$$\lim_{k\to\infty} h_{k,1} = \lim_{k\to\infty}\sum_{j=k}^{m_k} a_{k,j} u_{i_j} = u$$

almost everywhere in Ω. This shows that $u=v$ almost everywhere in Ω, which implies that $u\in W^{1,p}(\Omega)$.

(2) If $u_i\in W_0^{1,p}(\Omega)$ for every $i\in\mathbb{N}$, then $u\in W_0^{1,p}(\Omega)$ by Theorem 2.20. □

Remark 2.25. Theorem 2.20 and Theorem 2.24 do not hold when $p=1$ (exercise).

As a final result in this section we show that pointwise uniform bounds are preserved under weak convergence.

Theorem 2.26. Let $1<p<\infty$ and $m\in\mathbb{N}$, and let $\Omega\subset\mathbb{R}^n$ be an open set. Assume that sequences $(f_i)_{i\in\mathbb{N}}$ and $(g_i)_{i\in\mathbb{N}}$ are such that f_i converges to f weakly in $L^p(\Omega;\mathbb{R}^m)$ and g_i converges to g weakly in $L^p(\Omega)$ as $i\to\infty$. If $|f_i(x)|\leqslant g_i(x)$ for almost every $x\in\Omega$, then $|f(x)|\leqslant g(x)$ for almost every $x\in\Omega$.

Proof. Let $x\in\Omega$ be a Lebesgue point of g and all components of f. Let $0<r<d(x,\partial\Omega)$ and assume that $\int_{B(x,r)} f(y)\,dy\neq 0$. Denote

$$e = \left|\fint_{B(x,r)} f(y)\,dy\right|^{-1}\fint_{B(x,r)} f(y)\,dy\in\mathbb{R}^m$$

and

$$h = \frac{\chi_{B(x,r)}}{|B(x,r)|} e\in L^{p'}(\Omega;\mathbb{R}^m).$$

By Cauchy–Schwarz's inequality and the assumptions, we have

$$\begin{aligned}
\left|\fint_{B(x,r)} f(y)\,dy\right| &= e\cdot\fint_{B(x,r)} f(y)\,dy = \int_\Omega f(y)\cdot h(y)\,dy\\
&= \lim_{i\to\infty}\int_\Omega f_i(y)\cdot h(y)\,dy \leqslant \liminf_{i\to\infty}\int_\Omega |f_i(y)||h(y)|\,dy\\
&\leqslant \liminf_{i\to\infty}\int_\Omega g_i(y)|h(y)|\,dy = \fint_{B(x,r)} g(y)\,dy.
\end{aligned}$$

This implies

$$\left|\fint_{B(x,r)} f(y)\,dy\right| \leqslant \fint_{B(x,r)} g(y)\,dy,$$

which clearly holds also if $\int_{B(x,r)} f(y)\,dy=0$. Since almost every point $x\in\Omega$ is a Lebesgue point of g and all components of f and the claim follows by passing $r\to 0$ on both sides of the inequality above. □

2.4 Dual spaces

Let X be a Banach space. A linear functional $L : X \to \mathbb{R}$ is bounded, if there exists a constant $M < \infty$ such that

$$|Lx| \leqslant M\|x\| \quad \text{for every} \quad x \in X.$$

The norm of L is

$$\|L\| = \sup_{\substack{x\in X,\\ \|x\|\neq 0}} \frac{|Lx|}{\|x\|} = \sup_{\substack{x\in X,\\ \|x\|\neq 0}} \frac{Lx}{\|x\|} = \sup_{\substack{x\in X,\\ \|x\|\leqslant 1}} |Lx| = \sup_{\substack{x\in X,\\ \|x\|=1}} |Lx|.$$

Recall, that for a linear functional L, we have

$$L : X \to \mathbb{R} \quad \text{is continuous} \quad \iff \quad L \text{ is bounded} \quad \iff \quad \|L\| < \infty.$$

The dual space X^* of a Banach space X is the collection of all bounded linear functionals on X. In this section we discuss the dual spaces of $W_0^{1,p}(\Omega)$ and $W^{1,p}(\Omega)$.

Let $1 \leqslant p < \infty$ and assume that $f_0, f_1, \dots, f_n \in L^{p'}(\Omega)$, where p' is the Hölder conjugate exponent of p with $\frac{1}{p} + \frac{1}{p'} = 1$. Then the functional $L : W_0^{1,p}(\Omega) \to \mathbb{R}$,

$$L(u) = \int_\Omega \left(f_0 u + \sum_{j=1}^n f_j D_j u \right) dx$$

belongs to $W_0^{1,p}(\Omega)^*$.

Reason. As in (2.8), we have

$$\begin{aligned}
|L(u)| &\leqslant \int_\Omega \left(f_0 u + \sum_{j=1}^n f_j D_j u \right) dx \\
&\leqslant \left(\sum_{j=0}^n \|f_j\|_{L^{p'}(\Omega)}^{p'} \right)^{\frac{1}{p'}} \left(\|u\|_{L^p(\Omega)}^p + \sum_{j=1}^n \|D_j u\|_{L^p(\Omega)}^p \right)^{\frac{1}{p}} \\
&= \|f\|_{L^p(\Omega;\mathbb{R}^{n+1})} \|u\|_{W^{1,p}(\Omega)}
\end{aligned}$$

for every $u \in W_0^{1,p}(\Omega)$. Here $f = (f_0, f_1, \dots, f_n)$ and

$$\|f\|_{L^p(\Omega;\mathbb{R}^{n+1})} = \left(\sum_{j=0}^n \|f_j\|_{L^p(\Omega)}^p \right)^{\frac{1}{p}} < \infty.$$

Thus

$$\|L\|_{W_0^{1,p}(\Omega)^*} = \sup_{\substack{u\in W_0^{1,p}(\Omega),\\ \|u\|_{W^{1,p}(\Omega)}\neq 0}} \frac{|L(u)|}{\|u\|_{W^{1,p}(\Omega)}} \leqslant \|f\|_{L^p(\Omega;\mathbb{R}^{n+1})} = \left(\sum_{j=0}^n \|f_j\|_{L^{p'}(\Omega)}^{p'} \right)^{\frac{1}{p'}}.$$

■

We apply the following version of the Riesz representation theorem in $L^p(\Omega;\mathbb{R}^{n+1})$.

Theorem 2.27. Let $1 \leqslant p < \infty$ and assume that $L : L^p(\Omega;\mathbb{R}^{n+1}) \to \mathbb{R}$ is a bounded linear functional, that is, $L \in L^p(\Omega;\mathbb{R}^{n+1})^*$. Then there exists a unique $g = (g_0, g_1, \dots, g_n) \in L^{p'}(\Omega;\mathbb{R}^{n+1})$ such that

$$L(f) = \int_\Omega \sum_{j=0}^n f_j g_j \, dx = \int_\Omega f \cdot g \, dx$$

for every $f = (f_0, f_1, \dots, f_n) \in L^p(\Omega;\mathbb{R}^{n+1})$. Moreover,

$$\|L\| = \|g\|_{L^{p'}(\Omega;\mathbb{R}^{n+1})} = \left(\sum_{j=0}^n \|g_j\|_{L^{p'}(\Omega)}^{p'} \right)^{\frac{1}{p'}}.$$

Proof. Apply the Riesz representation theorem in $L^p(\Omega)$ componentwise (exercise). □

The following result holds true for the dual of $W_0^{1,p}(\Omega)$ and $W^{1,p}(\Omega)$. We state it only for Sobolev spaces with zero boundary values.

Theorem 2.28. Let $1 \leqslant p < \infty$ and let $\Omega \subset \mathbb{R}^n$ be an open set. Then for every bounded linear functional $L \in W_0^{1,p}(\Omega)^*$ there exist $f_0, f_1, \dots, f_n \in L^{p'}(\Omega)$ such that

$$L(u) = \int_\Omega \left(f_0 u + \sum_{j=1}^n f_j D_j u \right) dx$$

for every $u \in W_0^{1,p}(\Omega)$ and

$$\|L\|_{W_0^{1,p}(\Omega)^*} = \left(\sum_{j=0}^n \|f_j\|_{L^{p'}(\Omega)}^{p'} \right)^{\frac{1}{p'}}.$$

Proof. Consider the embedding $T : W_0^{1,p}(\Omega) \to L^p(\Omega;\mathbb{R}^{n+1})$,

$$T(u) = (u, D_1 u, \dots, D_n u).$$

Then

$$\|T(u)\|_{L^p(\Omega;\mathbb{R}^{n+1})} = \left(\|u\|_{L^p(\Omega)}^p + \sum_{j=1}^n \|D_j u\|_{L^p(\Omega)}^p \right)^{\frac{1}{p}} = \|u\|_{W^{1,p}(\Omega)}$$

for every $u \in W_0^{1,p}(\Omega)$. Thus T is linear, one-to-one, bounded and norm preserving. Since $W_0^{1,p}(\Omega)$ is complete, we conclude that $X = T(W_0^{1,p}(\Omega))$ is a closed subspace of $L^p(\Omega;\mathbb{R}^{n+1})$. Let $L \in W_0^{1,p}(\Omega)^*$ and let $L_1 : X \to \mathbb{R}$,

$$L_1(g) = L(T^{-1}(g))$$

for every $g \in X$. Then L_1 is a bounded linear operator with

$$\begin{aligned}
\|L_1\|_{X^*} &= \sup_{\substack{g\in X,\\ \|g\|_X\neq 0}} \frac{|L_1(g)|}{\|g\|_X} = \sup_{\substack{g\in X,\\ \|g\|_X\neq 0}} \frac{|L(T^{-1}(g))|}{\|g\|_X} \\
&= \sup_{\substack{T(u)\in X,\\ \|T(u)\|_{L^p(\Omega;\mathbb{R}^{n+1})}\neq 0}} \frac{|L(u)|}{\|T(u)\|_{L^p(\Omega;\mathbb{R}^{n+1})}} \\
&= \sup_{\substack{u\in W_0^{1,p}(\Omega),\\ \|u\|_{W^{1,p}(\Omega)}\neq 0}} \frac{|L(u)|}{\|u\|_{W^{1,p}(\Omega)}} = \|L\|_{W_0^{1,p}(\Omega)^*}.
\end{aligned}$$

By the Hahn-Banach theorem we may extend L_1 to a bounded linear functional $\overline{L}_1 : L^p(\Omega;\mathbb{R}^{n+1}) \to \mathbb{R}$ such that

$$\|\overline{L}_1\|_{L^p(\Omega;\mathbb{R}^{n+1})^*} = \|L_1\|_{X^*} = \|L\|_{W_0^{1,p}(\Omega)^*}.$$

By Theorem 2.27 there exists a unique $g = (g_0, g_1, \dots, g_n) \in L^{p'}(\Omega;\mathbb{R}^{n+1})$ such that

$$\overline{L}_1(f) = \int_\Omega \sum_{j=0}^n f_j g_j \, dx$$

for every $f = (f_0, f_1, \dots, f_n) \in L^p(\Omega;\mathbb{R}^{n+1})$ and

$$\|L\|_{W_0^{1,p}(\Omega)^*} = \|\overline{L}_1\|_{L^p(\Omega;\mathbb{R}^{n+1})^*} = \|g\|_{L^{p'}(\Omega;\mathbb{R}^{n+1})} = \left(\sum_{j=0}^n \|g_j\|_{L^{p'}(\Omega)}^{p'}\right)^{\frac{1}{p'}}.$$

It follows that

$$L(u) = L_1(T(u)) = \overline{L}_1(T(u)) = \int_\Omega \left(f_0 u + \sum_{j=1}^n f_j D_j u\right) dx$$

for every $u \in W_0^{1,p}(\Omega)$. □

Remark 2.29. Note that the previous theorem does not imply that $W_0^{1,p}(\Omega)^* = L^{p'}(\Omega;\mathbb{R}^{n+1})$. We have shown that if $f = (f_0, f_1, \dots, f_n) \in L^{p'}(\Omega;\mathbb{R}^{n+1})$, then the functional $L : W_0^{1,p}(\Omega) \to \mathbb{R}$,

$$L(u) = \int_\Omega \left(f_0 u + \sum_{j=1}^n f_j D_j u\right) dx \tag{2.30}$$

belongs to $W_0^{1,p}(\Omega)^*$ with

$$\|L\|_{W_0^{1,p}(\Omega)^*} \leqslant \left(\sum_{j=0}^n \|f_j\|_{L^{p'}(\Omega)}^{p'}\right)^{\frac{1}{p'}}.$$

On the other hand, by Theorem 2.28, there exists $g = (g_0, g_1, \dots, g_n) \in L^{p'}(\Omega;\mathbb{R}^{n+1})$ such that

$$L(u) = \int_\Omega \left(g_0 u + \sum_{j=1}^n g_j D_j u\right) dx$$

for every $u \in W_0^{1,p}(\Omega)$ and

$$\|L\|_{W_0^{1,p}(\Omega)^*} = \left(\sum_{j=0}^{n} \|g_j\|_{L^{p'}(\Omega)}^{p'} \right)^{\frac{1}{p'}}.$$

If follows that

$$\|L\|_{W_0^{1,p}(\Omega)^*} = \min \left\{ \left(\sum_{j=0}^{n} \|f_j\|_{L^{p'}(\Omega)}^{p'} \right)^{\frac{1}{p'}} : f \in L^{p'}(\Omega;\mathbb{R}^{n+1}) \text{ such that (2.30) holds.} \right\}$$

Observe, that the representation (2.30) is not unique in general. For example, let $\Omega \subset \mathbb{R}^n$ be a bounded open set and assume that $h \in C^2(\mathbb{R}^n)$ be a harmonic function in $\mathbb{R}^n$, that is, a classical solution to the Laplace equation

$$\Delta h(x) = \sum_{j=1}^{n} \frac{\partial^2 h}{\partial x_j^2}(x)$$

for every $x \in \mathbb{R}^n$. Then

$$-\int_\Omega \sum_{j=1}^{n} D_j h D_j \varphi \, dx = \int_\Omega \sum_{j=1}^{n} \frac{\partial^2 h}{\partial x_j^2} \varphi \, dx = \int_\Omega \Delta h \varphi \, dx = 0$$

for every $\varphi \in C_0^\infty(\Omega)$. Let $u \in W_0^{1,p}(\Omega)$. Then there exist functions $\varphi_i \in C_0^\infty(\Omega)$, $i = 1, 2, \dots$, such that $\varphi_i \to u$ in $W^{1,p}(\Omega)$ as $i \to \infty$. Then

$$\begin{aligned}
\left| \int_\Omega D_j h D_j \varphi_i \, dx - \int_\Omega D_j h D_j u \, dx \right| &= \left| \int_\Omega (D_j h D_j \varphi_i - D_j h D_j u) \, dx \right| \\
&\leqslant \int_\Omega |D_j h| |D_j \varphi_i - D_j u| \, dx \\
&\leqslant \|D_j h\|_{L^{p'}(\Omega)} \|D_j \varphi_i - D_j u\|_{L^p(\Omega)} \\
&\leqslant \|Dh\|_{L^\infty(\overline{\Omega})} |\Omega|^{\frac{1}{p'}} \|D_j \varphi_i - D_j u\|_{L^p(\Omega)} \xrightarrow{i \to \infty} 0
\end{aligned}$$

for every $j = 1, \dots, n$. This shows that

$$\int_\Omega \sum_{j=1}^{n} D_j h D_j u \, dx = \lim_{i \to \infty} \int_\Omega \sum_{j=1}^{n} D_j h D_j \varphi_i \, dx = 0$$

for every $\varphi \in W_0^{1,p}(\Omega)$. Thus if $f = (f_0, f_1, \dots, f_n) \in L^{p'}(\Omega;\mathbb{R}^{n+1})$ such that (2.30) holds, then

$$\begin{aligned}
L(u) &= \int_\Omega \left(f_0 u + \sum_{j=1}^{n} f_j D_j u \right) dx \\
&= \int_\Omega \left(f_0 u + \sum_{j=1}^{n} (f_j + D_j h) D_j u \right) dx
\end{aligned}$$

for every $u \in W_0^{1,p}(\Omega)$.

2.5 Difference quotients

In this section we give a characterization of $W^{1,p}$, $1 < p < \infty$, in terms of difference quotients. This approach is useful in regularity theory for PDEs. Moreover, this characterization does not involve derivatives.

Definition 2.31. Let $u \in L^1_{\text{loc}}(\Omega)$ and $\Omega' \Subset \Omega$. The j^{th} difference quotient is

$$D_j^h u(x) = \frac{u(x+he_j)-u(x)}{h}, \quad j = 1,\dots,n,$$

for $x \in \Omega'$ and $h \in \mathbb{R}$ such that $0 < |h| < \operatorname{dist}(\Omega', \partial\Omega)$. We denote

$$D^h u = (D_1^h u, \dots, D_n^h u).$$

THE MORAL: Note that the definition of the difference quotient makes sense at every $x \in \Omega$ whenever $0 < |h| < \operatorname{dist}(x, \partial\Omega)$. If $\Omega = \mathbb{R}^n$, then the definition makes sense for every $h \neq 0$.

Theorem 2.32.

(1) Assume $u \in W^{1,p}(\Omega)$, $1 \leqslant p < \infty$. Then for every $\Omega' \Subset \Omega$, we have

$$\|D^h u\|_{L^p(\Omega')} \leqslant c\|Du\|_{L^p(\Omega)}$$

for some constant $c = c(n,p)$ and all $0 < |h| < \operatorname{dist}(\Omega', \partial\Omega)$.

(2) If $u \in L^p(\Omega')$, $1 < p < \infty$, and there is a constant c such that

$$\|D^h u\|_{L^p(\Omega')} \leqslant c$$

whenever $0 < |h| < \operatorname{dist}(\Omega', \partial\Omega)$, then $u \in W^{1,p}(\Omega')$ and $\|Du\|_{L^p(\Omega')} \leqslant c$.

(3) Let $1 < p < \infty$, and assume that $u \in L^p(\mathbb{R}^n)$ and that there exists a constant c such that

$$\|D^h u\|_{L^p(\mathbb{R}^n)} \leqslant c$$

for every $h \neq 0$. Then the weak derivative Du with respect to $\mathbb{R}^n$ exists, $u \in W^{1,p}(\mathbb{R}^n)$ and $\|Du\|_{L^p(\mathbb{R}^n)} \leqslant c$.

THE MORAL: Pointwise derivatives are defined as limit of difference quotients and Sobolev spaces can be characterized by integrated difference quotients.

WARNING: Claim (2) does not hold for $p = 1$ (exercise).

Proof. (1) First assume that $u \in C^\infty(\Omega) \cap W^{1,p}(\Omega)$. Then

$$\begin{aligned} u(x+he_j) - u(x) &= \int_0^h \frac{\partial}{\partial t}(u(x+te_j))\,dt \\ &= \int_0^h Du(x+te_j)\cdot e_j\,dt \\ &= \int_0^h \frac{\partial u}{\partial x_j}(x+te_j)\,dt, \quad j = 1,\dots,n, \end{aligned}$$

for all $x \in \Omega'$, $0 < |h| < \operatorname{dist}(\Omega', \partial\Omega)$. By Hölder's inequality

$$\begin{aligned} |D_j^h u(x)| &= \left| \frac{u(x+he_j) - u(x)}{h} \right| \\ &\leqslant \frac{1}{|h|} \int_{-|h|}^{|h|} \left| \frac{\partial u}{\partial x_j}(x+te_j) \right| dt \\ &\leqslant \frac{1}{|h|} \left(\int_{-|h|}^{|h|} \left| \frac{\partial u}{\partial x_j}(x+te_j) \right|^p dt \right)^{1/p} |2h|^{1-\frac{1}{p}}, \end{aligned}$$

which implies

$$|D_j^h u(x)|^p \leqslant \frac{2^{p-1}}{|h|} \int_{-|h|}^{|h|} \left| \frac{\partial u}{\partial x_j}(x+te_j) \right|^p dt$$

Next we integrate over Ω' and switch the order of integration by Fubini's theorem to conclude

$$\begin{aligned} \int_{\Omega'} |D_j^h u(x)|^p\,dx &\leqslant \frac{2^{p-1}}{|h|} \int_{\Omega'} \int_{-|h|}^{|h|} \left| \frac{\partial u}{\partial x_j}(x+te_j) \right|^p dt\,dx \\ &= \frac{2^{p-1}}{|h|} \int_{-|h|}^{|h|} \int_{\Omega'} \left| \frac{\partial u}{\partial x_j}(x+te_j) \right|^p dx\,dt \\ &\leqslant 2^p \int_{\Omega} \left| \frac{\partial u}{\partial x_j}(x) \right|^p dx. \end{aligned}$$

The last inequality follows from the fact that, for $0 < |h| < \operatorname{dist}(\Omega', \partial\Omega)$ and $|t| \leqslant |h|$, we have

$$\int_{\Omega'} \left| \frac{\partial u}{\partial x_j}(x+te_j) \right|^p dx \leqslant \int_{\Omega} \left| \frac{\partial u}{\partial x_j}(x) \right|^p dx.$$

Using the elementary inequality $(a_1 + \dots + a_n)^\alpha \leqslant n^\alpha (a_1^\alpha + \dots + a_n^\alpha)$, $a_i \geqslant 0$, $\alpha > 0$, we obtain

$$\begin{aligned} \int_{\Omega'} |D^h u(x)|^p\,dx &= \int_{\Omega'} \left(\sum_{j=1}^n |D_j^h u(x)|^2 \right)^{\frac{p}{2}} dx \leqslant n^{\frac{p}{2}} \int_{\Omega'} \sum_{j=1}^n |D_j^h u(x)|^p\,dx \\ &= n^{\frac{p}{2}} \sum_{j=1}^n \int_{\Omega'} |D_j^h u(x)|^p\,dx \leqslant 2^p n^{\frac{p}{2}} \sum_{j=1}^n \int_{\Omega} \left| \frac{\partial u}{\partial x_j}(x) \right|^p dx \\ &\leqslant 2^p n^{1+\frac{p}{2}} \int_{\Omega} |Du(x)|^p\,dx \end{aligned}$$

The general case $u \in W^{1,p}(\Omega)$ follows by an approximation, see Theorem 1.21. Let $u_i \in C^\infty(\Omega) \cap W^{1,p}(\Omega)$, $i \in \mathbb{N}$, such that $u_i \to u$ in $W^{1,p}(\Omega)$ as $i \to \infty$. By passing to a subsequence, if necessary, we may also assume that $u_i \to u$ pointwise almost everywhere in Ω as $i \to \infty$. Assume that $0 < |h| < \operatorname{dist}(\Omega', \partial\Omega)$. Then $D^h u_i(x) \to D^h u(x)$ for almost every $x \in \Omega'$ as $i \to \infty$. By Fatou's lemma and assumption we obtain

$$\begin{aligned}\int_{\Omega'} |D^h u(x)|^p \, dx &\leqslant \liminf_{i\to\infty} \int_{\Omega'} |D^h u_i(x)|^p \, dx \\ &\leqslant c(n) \liminf_{i\to\infty} \int_{\Omega} |Du_i(x)|^p \, dx \\ &= c(n) \int_{\Omega} |Du(x)|^p \, dx.\end{aligned}$$

(2) Let $\varphi \in C_0^\infty(\Omega')$. Then by a change of variables we see that, for $0 < |h| < \operatorname{dist}(\operatorname{supp}\varphi, \partial\Omega')$, we have

$$\int_{\Omega'} u(x) \frac{\varphi(x+he_j) - \varphi(x)}{h} \, dx = -\int_{\Omega'} \frac{u(x-he_j) - u(x)}{-h} \varphi(x) \, dx, \quad j = 1, \dots, n.$$

This shows that

$$\int_{\Omega'} u D_j^h \varphi \, dx = -\int_{\Omega} (D_j^{-h} u) \varphi \, dx, \quad j = 1, \dots, n.$$

By assumption

$$\sup_{0<|h|<\operatorname{dist}(\Omega',\partial\Omega)} \|D_j^{-h} u\|_{L^p(\Omega')} \leqslant c < \infty.$$

Since $1 < p < \infty$, by Theorem 2.13 there exists $f \in L^p(\Omega'; \mathbb{R}^n)$ and a sequence $(h_i)_{i\in\mathbb{N}}$ converging to zero such that $D^{-h_i} u \to f$ weakly in $L^p(\Omega'; \mathbb{R}^n)$ as $i \to \infty$. This implies

$$\begin{aligned}\int_{\Omega'} u \frac{\partial \varphi}{\partial x_j} \, dx &= \int_{\Omega'} u \left(\lim_{h_i \to 0} D_j^{h_i} \varphi \right) dx = \lim_{h_i \to 0} \int_{\Omega'} u D_j^{h_i} \varphi \, dx \\ &= -\lim_{h_i\to 0} \int_{\Omega'} (D_j^{-h_i} u) \varphi \, dx = -\int_{\Omega'} f_j \varphi \, dx\end{aligned}$$

for every $\varphi \in C_0^\infty(\Omega')$. Here the second equality follows from the dominated convergence theorem and the last equality is weak convergence the weak convergence tested with $g = (0, \dots, \varphi, \dots, 0)$, where φ is in the jth position. It follows that $Du = f$ in the weak sense in Ω' and thus $u \in W^{1,p}(\Omega')$. By (2.12),

$$\|Du\|_{L^p(\Omega';\mathbb{R}^n)} = \|f\|_{L^p(\Omega';\mathbb{R}^n)} \leqslant \liminf_{i\to\infty} \|D^{-h_i} u\|_{L^p(\Omega';\mathbb{R}^n)} \leqslant c.$$

(3) Let $\Omega_i = B(0, 2i)$ and $\Omega_i' = B(0, i)$ for every $i \in \mathbb{N}$. Assertion (2) and the assumption imply that $u_i = u|_{\Omega_i}$, $i \in \mathbb{N}$, has a weak derivative Du_i in Ω_i' and $\|Du_i\|_{L^p(\Omega_i';\mathbb{R}^n)} \leqslant c$. Since $Du_{i+1} = Du_i$ almost everywhere in Ω_i', we see that the limit

$$f(x) = \lim_{i\to\infty} \chi_{\Omega_i'}(x) Du_i(x)$$

exists for almost every $x \in \mathbb{R}^n$. The weak derivative of u with respect to $\mathbb{R}^n$ coincides with $f \in L^1_{\text{loc}}(\mathbb{R}^n;\mathbb{R}^n)$ and Fatou's lemma implies

$$\begin{aligned}\|Du\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} = \|f\|_{L^p(\mathbb{R}^n;\mathbb{R}^n)} &= \Big(\int_{\mathbb{R}^n} \lim_{i\to\infty} |\chi_{\Omega_i'} Du_i|^p\,dx\Big)^{\frac{1}{p}} \\ &\leqslant \liminf_{i\to\infty} \Big(\int_{\Omega_i'} |Du_i|^p\,dx\Big)^{\frac{1}{p}} \leqslant c.\end{aligned}$$

From this it also follows that $u \in W^{1,p}(\mathbb{R}^n)$. □

Remark 2.33. By the proof of the previous theorem $u \in W^{1,p}(\mathbb{R}^n)$, $1 < p < \infty$, if and only if $u \in L^p(\mathbb{R}^n)$ and

$$\limsup_{h\to 0} \|D^h u\|_{L^p(\mathbb{R}^n)} < \infty.$$

2.6 Absolute continuity on lines

In this section we relate weak derivatives to classical derivatives and give a characterization $W^{1,p}$ in terms of absolute continuity on lines.

Recall that a function $u : [a,b] \to \mathbb{R}$ is absolutely continuous, if for every $\varepsilon > 0$, there exists $\delta > 0$ such that if $a = x_1 < y_1 \leqslant x_2 < y_2 \leqslant \ldots \leqslant x_m < y_m = b$ is a partition of $[a,b]$ with

$$\sum_{i=1}^{m} (y_i - x_i) < \delta,$$

then

$$\sum_{i=1}^{m} |u(y_i) - u(x_i)| < \varepsilon.$$

Absolute continuity can be characterized in terms of the fundamental theorem of calculus.

Theorem 2.34. A function $u : [a,b] \to \mathbb{R}$ is absolutely continuous if and only if there exists a function $g \in L^1((a,b))$ such that

$$u(x) = u(a) + \int_a^x g(t)\,dt.$$

By the Lebesgue differentiation theorem $g = u'$ almost everywhere in (a,b).

THE MORAL: Absolutely continuous functions are precisely those functions for which the fundamental theorem of calculus holds true.

Examples 2.35:

(1) Every Lipchitz continuous function $u : [a,b] \to \mathbb{R}$ is absolutely continuous.

(2) The Cantor function u is continuous in $[0,1]$ and differentiable almost everywhere in $(0,1)$, but not absolutely continuous in $[0,1]$.

Reason.

$$u(1) = 1 \neq 0 = u(0) + \int_0^1 \underbrace{u'(t)}_{=0} \, dt.$$

■

The next result relates weak partial derivatives with the classical partial derivatives.

Theorem 2.36 (Nikodym, ACL characterization). Assume that $u \in W^{1,p}_{\mathrm{loc}}(\Omega)$, $1 \leqslant p \leqslant \infty$ and let $\Omega' \Subset \Omega$. Then there exists $u^* : \Omega \to [-\infty, \infty]$ such that

(1) $u^* = u$ almost everywhere in Ω,

(2) u^* is absolutely continuous on almost every (with respect to the $(n-1)$-dimensional Lebesgue measure) line segments in Ω', that are parallel to the coordinate axes and

(3) the classical partial derivatives of u^* coincide with the weak partial derivatives of u almost everywhere in Ω.

Conversely, if $u \in L^p_{\mathrm{loc}}(\Omega)$ and there exists u^* as above such that $D_i u^* \in L^p_{\mathrm{loc}}(\Omega)$, $i = 1, \ldots, n$, then $u \in W^{1,p}_{\mathrm{loc}}(\Omega)$.

THE MORAL: This is a very useful characterization of $W^{1,p}$, since many claims for weak derivatives can be reduced to the one-dimensional claims for absolute continuous functions. In addition, this gives a practical tool to show that a function belongs to a Sobolev space.

Remarks 2.37:

(1) Let $u \in W^{1,p}(\mathbb{R}^n)$. By the ACL characterization there exists a function u^* such that $u^* = u$ almost everywhere in $\mathbb{R}^n$, u^* is absolutely continuous on almost every line segments in $\mathbb{R}^n$ parallel to the coordinate axes and the classical partial derivatives of u^* coincide with the weak partial derivatives of u almost everywhere in $\mathbb{R}^n$.

(2) In the one-dimensional case we obtain the following characterization: $u \in W^{1,p}((a,b))$, $1 \leqslant p \leqslant \infty$, if u can be redefined on a set of measure zero in such a way that $u \in L^p((a,b))$ and u is absolutely continuous on every compact subinterval of (a,b) and the classical derivative exists and belongs to $u \in L^p((a,b))$. Moreover, the classical derivative equals to the weak derivative almost everywhere.

(3) A function $u \in W^{1,p}(\Omega)$ has a representative that has classical partial derivatives almost everywhere. However, this does not give any information concerning the total differentiability of the function. See Theorem 3.28.

Proof. Since the claims are local, we may assume that $\Omega = \mathbb{R}^n$ and that u has a compact support.

$\boxed{\Longrightarrow}$ Let $u_i = u_{\varepsilon_i}$, $i = 1,2,\dots$, be a sequence of standard convolution approximations of u such that $\operatorname{supp} u_i \subset B(0,R)$ for every $i = 1,2,\dots$ and

$$\|u_i - u\|_{W^{1,1}(\mathbb{R}^n)} < \frac{1}{2^i}, \quad i = 1,2,\dots$$

By Lemma 1.18 (2), the sequence of convolution approximations converges pointwise almost everywhere and thus the limit $\lim_{i\to\infty} u_i(x)$ exists for every $x \in \mathbb{R}^n \setminus E$ for some $E \subset \mathbb{R}^n$ with $|E| = 0$. We define

$$u^*(x) = \begin{cases} \lim\limits_{i\to\infty} u_i(x), & x \in \mathbb{R}^n \setminus E, \\ 0, & x \in E. \end{cases}$$

We fix a standard base vector in $\mathbb{R}^n$ and, without loss of generality, we may assume that it is $(0,\dots,0,1)$. Let

$$f_i(x_1,\dots,x_{n-1}) = \int_{\mathbb{R}} \left(|u_{i+1} - u_i| + \sum_{j=1}^{n} \left| \frac{\partial u_{i+1}}{\partial x_j} - \frac{\partial u_i}{\partial x_j} \right| \right) (x_1,\dots,x_n)\, dx_n$$

and

$$f(x_1,\dots,x_{n-1}) = \sum_{i=1}^{\infty} f_i(x_1,\dots,x_{n-1}).$$

By the monotone convergence theorem and Fubini's theorem

$$\begin{aligned} \int_{\mathbb{R}^{n-1}} f\, dx_1 \dots dx_{n-1} &= \int_{\mathbb{R}^{n-1}} \sum_{i=1}^{\infty} f_i\, dx_1 \dots dx_{n-1} \\ &= \sum_{i=1}^{\infty} \int_{\mathbb{R}^{n-1}} f_i\, dx_1 \dots dx_{n-1} \\ &= \sum_{i=1}^{\infty} \int_{\mathbb{R}^n} \left(|u_{i+1} - u_i| + \left| \frac{\partial u_{i+1}}{\partial x_j} - \frac{\partial u_i}{\partial x_j} \right| \right) dx \\ &< \sum_{i=1}^{\infty} \frac{1}{2^i} < \infty. \end{aligned}$$

This shows that $f \in L^1(\mathbb{R}^{n-1})$ and thus $f < \infty$ $(n-1)$-almost everywhere in $\mathbb{R}^{n-1}$. Let $\widehat{x} = (x_1,\dots,x_{n-1}) \in \mathbb{R}^{n-1}$ such that $f(\widehat{x}) < \infty$. Denote

$$g_i(t) = u_i(\widehat{x}, t) \quad \text{and} \quad g(t) = u^*(\widehat{x}, t).$$

Claim: (g_i) is a Cauchy sequence in $C(\mathbb{R})$.

Reason. Note that

$$g_i = g_1 + \sum_{k=1}^{i-1} (g_{k+1} - g_k), \quad i = 1,2,\dots,$$

where

$$
\begin{aligned}
|g_{k+1}(t)-g_k(t)| &= \left|\int_{-\infty}^{t}(g'_{k+1}-g'_k)(s)\,ds\right| \\
&\leqslant \int_{\mathbb{R}} |g'_{k+1}(s)-g'_k(s)|\,ds \\
&\leqslant \int_{\mathbb{R}} \left|\frac{\partial u_{k+1}}{\partial x_n}(\hat{x},s)-\frac{\partial u_k}{\partial x_n}(\hat{x},s)\right| ds \leqslant f_k(\hat{x}).
\end{aligned}
$$

Thus

$$
\begin{aligned}
\left|\sum_{k=1}^{i-1}(g_{k+1}(t)-g_k(t))\right| &\leqslant \sum_{k=1}^{i-1}|g_{k+1}(t)-g_k(t)| \\
&\leqslant \sum_{k=1}^{\infty} f_k(\hat{x}) = f(\hat{x}) < \infty
\end{aligned}
$$

for every $t\in\mathbb{R}$. This implies that (g_i) is a Cauchy sequence in $C(\mathbb{R})$. Since $C(\mathbb{R})$ is complete, there exists $g\in C(\mathbb{R})$ such that $g_i\to g$ uniformly in $\mathbb{R}$. It follows that $\{\hat{x}\}\times\mathbb{R}\subset\mathbb{R}^n\setminus E$. ■

Claim: (g_i) is a Cauchy sequence in $L^1(\mathbb{R})$.

Reason.

$$
\begin{aligned}
\int_{\mathbb{R}}\left|\sum_{k=1}^{i-1}(g_{k+1}(t)-g_k(t))\right| dt &\leqslant \sum_{k=1}^{i-1}\int_{\mathbb{R}}|g_{k+1}(t)-g_k(t)|\,dt \\
&\leqslant \sum_{k=1}^{i-1} f_k(\hat{x}) \leqslant f(\hat{x}) < \infty.
\end{aligned}
$$

This implies that (g_i) is a Cauchy sequence in $L^1(\mathbb{R})$. Since $L^1(\mathbb{R})$ is complete, there exists $\tilde{g}\in L^1(\mathbb{R})$ such that $g_i\to\tilde{g}$ in $L^1(\mathbb{R})$ as $i\to\infty$. ■

Claim: g is absolutely continuous in $\mathbb{R}$.

Reason.

$$
g(t)=\lim_{i\to\infty} g_i(t)=\lim_{i\to\infty}\int_{-\infty}^{t} g'_i(s)\,ds=\int_{-\infty}^{t}\tilde{g}'(s)\,ds
$$

This implies that g is absolutely continuous in $\mathbb{R}$ and $g'=\tilde{g}$ almost everywhere in $\mathbb{R}$. ■

Claim: $\tilde{g}$ is the weak derivative of g.

Reason. Let $\varphi\in C_0^\infty(\mathbb{R})$. Then

$$
\int_{\mathbb{R}} g\varphi'\,dt=\lim_{i\to\infty}\int_{\mathbb{R}} g_i\varphi'\,dt=-\lim_{i\to\infty}\int_{\mathbb{R}} g'_i\varphi\,dt=-\int_{\mathbb{R}}\tilde{g}\varphi\,dt. \qquad ■
$$

Thus for every $\varphi \in C_0^\infty(\mathbb{R}^n)$ we have

$$\int_{\mathbb{R}} u^*(\widehat{x},x_n)\frac{\partial \varphi}{\partial x_n}(\widehat{x},x_n)\,dx_n = -\int_{\mathbb{R}} \frac{\partial u^*}{\partial x_n}(\widehat{x},x_n)\varphi(\widehat{x},x_n)\,dx_n$$

and by Fubini's theorem

$$\int_{\mathbb{R}^n} u\frac{\partial \varphi}{\partial x_n}\,dx = -\int_{\mathbb{R}^n} \frac{\partial u^*}{\partial x_n}\varphi\,dx.$$

This shows that u^* has the classical partial derivatives almost everywhere in $\mathbb{R}^n$ and that they coincide with the weak partial derivatives of u almost everywhere in $\mathbb{R}^n$.

$\boxed{\Longleftarrow}$ Assume that u has a representative u^* as in the statement of the theorem. For every $\varphi \in C_0^\infty(\mathbb{R}^n)$, the function $u^*\varphi$ has the same absolute continuity properties as u^*. By the fundamental theorem of calculus

$$\int_{\mathbb{R}} \frac{\partial (u^*\varphi)}{\partial x_n}(\widehat{x},t)\,dt = 0$$

for $(n-1)$-almost every $\widehat{x} \in \mathbb{R}^{n-1}$. Thus

$$\int_{\mathbb{R}} u^*(\widehat{x},t)\frac{\partial \varphi}{\partial x_n}(\widehat{x},t)\,dt = -\int_{\mathbb{R}} \frac{\partial u^*}{\partial x_n}(\widehat{x},t)\varphi(\widehat{x},t)\,dt$$

and by Fubini's theorem

$$\int_{\mathbb{R}^n} u^*\frac{\partial \varphi}{\partial x_n}\,dx = -\int_{\mathbb{R}^n} \frac{\partial u^*}{\partial x_n}\varphi\,dx.$$

Since $u^* = u$ almost everywhere in $\mathbb{R}^n$, we see that $\frac{\partial u^*}{\partial x_n}$ is the nth weak partial derivative of u. The same argument applies to all other partial derivatives $\frac{\partial u^*}{\partial x_j}$, $j = 1,\dots,n$ as well. □

Remarks 2.38:

(1) The ACL characterization can be used to give a simple proof of Example 1.10 (exercise).

(2) The ACL characterization can be used to give a simple proof of the Leibniz rule. If $u \in W^{1,p}(\Omega)\cap L^\infty(\Omega)$ and $v \in W^{1,p}(\Omega)\cap L^\infty(\Omega)$, then $uv \in W^{1,p}(\Omega)$ and

$$D_j(uv) = vD_ju + uD_jv, \quad j = 1,\dots,n,$$

almost everywhere in Ω (exercise), compare to Lemma 1.14 (5).

(3) The ACL characterization can be used to give a simple proof for Lemma 2.1 and Theorem 2.3. The claim that if $u,v \in W^{1,p}(\Omega)$, then $\max\{u,v\} \in W^{1,p}(\Omega)$ and $\min\{u,v\} \in W^{1,p}(\Omega)$ follows also in a similar way (exercise).

(4) The ACL characterization can be used to show that if Ω is connected, $u \in W^{1,p}_{\mathrm{loc}}(\Omega)$ and $Du = 0$ almost everywhere in Ω, then u is a constant almost everywhere in Ω (exercise).

(5) The ACL characterization can be used to show that $u \in W^{1,\infty}_{\mathrm{loc}}(\Omega)$ if and only if u is locally Lipschitz (exercise). Compare with Theorem 3.31 below.

Example 2.39. Let $\mathbb{R}^n_+ = \{(x_1,\dots,x_n) \in \mathbb{R}^n : x_n > 0\}$ be the upper half space. Assume that $u \in W^{1,p}(\mathbb{R}^n_+)$, $1 \leqslant p \in \infty$. Let

$$\overline{u}(x) = \begin{cases} u(x_1,\dots,x_n), & x_n > 0, \\ u(x_1,\dots,-x_n), & x_n < 0. \end{cases}$$

Then $\overline{u} \in W^{1,p}(\mathbb{R}^n)$ and

$$D_j\overline{u}(x) = \begin{cases} D_j u(x_1,\dots,x_n), & x_n > 0, \\ (-1)^{\delta_{jn}} D_j u(x_1,\dots,-x_n), & x_n < 0, \end{cases}$$

$j = 1,\dots,n$, where δ_{in} is the Kronecker delta, that is $\delta_{jn} = 1$ if $j = n$ and $\delta_{jn} = 0$ otherwise (exercise). Moreover, we have

$$\|\overline{u}\|_{W^{1,p}(\mathbb{R}^n)} = \begin{cases} 2\|u\|_{W^{1,p}(\mathbb{R}^n_+)}, & 1 \leqslant p < \infty, \\ \|u\|_{W^{1,\infty}(\mathbb{R}^n_+)}, & p = \infty. \end{cases}$$

Thus there exists a bounded linear extension operator $E : W^{1,p}(\mathbb{R}^n_+) \to W^{1,p}(\mathbb{R}^n)$, $Eu = \overline{u}$ such that $(Eu)|_{\mathbb{R}^n_+} = u$ for every $W^{1,p}(\mathbb{R}^n_+)$.

Example 2.40. The radial projection $u : B(0,1) \to \partial B(0,1)$, $u(x) = \frac{x}{|x|}$ is discontinuous at the origin. However, the coordinate functions $\frac{x_j}{|x|}$, $j = 1,\dots,n$, are absolutely continuous on almost every lines. Moreover,

$$D_i\left(\frac{x_j}{|x|}\right) = \frac{\delta_{ij}|x| - \frac{x_i x_j}{|x|}}{|x|^2} \in L^p(B(0,1))$$

whenever $1 \leqslant p < n$. Here δ_{ij} is the Kronecker symbol. By the ACL characterization the coordinate functions of u belong to $W^{1,p}(B(0,1))$ whenever $1 \leqslant p < n$.

Remark 2.41. We say that a relatively closed $E \subset \Omega$ is be removable for $W^{1,p}(\Omega)$, if $|E| = 0$ and $W^{1,p}(\Omega \setminus E) = W^{1,p}(\Omega)$ in the sense that every function in $W^{1,p}(\Omega \setminus E)$ can be approximated by the restrictions of functions in $C^\infty(\Omega)$ in the norm $\|\cdot\|_{W^{1,p}(\Omega)}$. Theorem 2.36 implies the following removability theorem for $W^{1,p}\Omega)$: if $\mathcal{H}^{n-1}(E) = 0$, then E is removable for $W^{1,p}(\Omega)$. Observe, that if $\mathcal{H}^{n-1}(E) = 0$, then E is contained in a measure zero set of lines in a fixed direction (equivalently the projection of E onto a hyperplane also has $\mathcal{H}^{n-1}$-measure zero).

This result is quite sharp. For example, let $\Omega = B(0,1)$ and $E = \{x \in B(0,1) : x_2 = 0\}$. Then $0 < \mathcal{H}^{n-1}(E) < \infty$, but E is not removable since, using Theorem 2.36 again, it is easy to see that the function which is 1 on the upper half-plane and 0 on the lower half-plane does not belong to $W^{1,p}\Omega)$. With a little more work we can show that $E' = E \cap B(0,\frac{1}{2})$ is not removable for $W^{1,p}(B(0,1))$.

3

Sobolev inequalities

The term Sobolev inequalities refers to a variety of inequalities involving functions and their derivatives. As an example, we consider an inequality of the form

$$\left(\int_{\mathbb{R}^n} |u|^q \, dx\right)^{\frac{1}{q}} \leqslant c\left(\int_{\mathbb{R}^n} |Du|^p \, dx\right)^{\frac{1}{p}} \tag{3.1}$$

for every $u \in C_0^\infty(\mathbb{R}^n)$, where constant $0 < c < \infty$ and exponent $1 \leqslant q < \infty$ are independent of u. By density of smooth functions in Sobolev spaces, see Theorem 1.21, we may conclude that (3.1) holds for functions in $W^{1,p}(\mathbb{R}^n)$ as well. Let $u \in C_0^\infty(\mathbb{R}^n)$, $u \not\equiv 0$, $1 \leqslant p < n$ and consider $u_\lambda(x) = u(\lambda x)$ with $\lambda > 0$. Since $u \in C_0^\infty(\mathbb{R}^n)$, it follows that (3.1) holds true for every u_λ with $\lambda > 0$ with c and q independent of λ. Thus

$$\left(\int_{\mathbb{R}^n} |u_\lambda|^q \, dx\right)^{\frac{1}{q}} \leqslant c\left(\int_{\mathbb{R}^n} |Du_\lambda|^p \, dx\right)^{\frac{1}{p}}$$

for every $\lambda > 0$. By a change of variables $y = \lambda x$, $dx = \frac{1}{\lambda^n}\, dy$, we see that

$$\int_{\mathbb{R}^n} |u_\lambda(x)|^q \, dx = \int_{\mathbb{R}^n} |u(\lambda x)|^q \, dx = \int_{\mathbb{R}^n} |u(y)|^q \frac{1}{\lambda^n} \, dy = \frac{1}{\lambda^n} \int_{\mathbb{R}^n} |u(x)|^q \, dx$$

and

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

Thus

$$\frac{1}{\lambda^{\frac{n}{q}}}\left(\int_{\mathbb{R}^n} |u|^q \, dx\right)^{\frac{1}{q}} \leqslant c\frac{\lambda}{\lambda^{\frac{n}{p}}}\left(\int_{\mathbb{R}^n} |Du|^p \, dx\right)^{\frac{1}{p}}$$

for every $\lambda > 0$, and equivalently,

$$\|u\|_{L^q(\mathbb{R}^n)} \leqslant c\lambda^{1-\frac{n}{p}+\frac{n}{q}} \|Du\|_{L^p(\mathbb{R}^n)}.$$

Since this inequality has to be independent of λ, we have

$$1 - \frac{n}{p} + \frac{n}{q} = 0 \quad \Longleftrightarrow \quad q = \frac{np}{n-p}.$$

THE MORAL: There is only one possible exponent q for which inequality (3.1) may hold true for all compactly supported smooth functions.

For $1 \leqslant p < n$, the Sobolev conjugate exponent of p is

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

Observe that

(1) $p^* > p$,

(2) If $p \to n-$, then $p^* \to \infty$ and

(3) If $p = 1$, then $p^* = \frac{n}{n-1}$.

3.1 Gagliardo-Nirenberg-Sobolev inequality

The following generalized Hölder's inequality will be useful for us.

Lemma 3.2. Let $1 \leqslant p_1, \ldots, p_k \leqslant \infty$ with $\frac{1}{p_1} + \cdots + \frac{1}{p_k} = 1$ and assume $f_i \in L^{p_i}(\Omega)$, $i = 1, \ldots, k$. Then

$$\int_\Omega |f_1 \ldots f_k|\,dx \leqslant \prod_{i=1}^k \|f_i\|_{L^{p_i}(\Omega)}.$$

Proof. Induction and Hölder's inequality (exercise). □

Sobolev proved the following theorem in the case $p > 1$ and Nirenberg and Gagliardo in the case $p = 1$.

Theorem 3.3 (Gagliardo-Nirenberg-Sobolev). Let $1 \leqslant p < n$. There exists $c = c(n,p)$ such that

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

for every $u \in W^{1,p}(\mathbb{R}^n)$.

THE MORAL: The Sobolev-Gagliardo-Nirenberg inequality implies that $W^{1,p}(\mathbb{R}^n) \subset L^{p^*}(\mathbb{R}^n)$, when $1 \leqslant p < n$. More precisely, $W^{1,p}(\mathbb{R}^n)$ is continuously imbedded in $L^{p^*}(\mathbb{R}^n)$, when $1 \leqslant p < n$. This is the Sobolev embedding theorem for $1 \leqslant p < n$.

Proof. (1) We start by proving the estimate for $u \in C_0^\infty(\mathbb{R}^n)$. By the fundamental theorem of calculus

$$u(x_1,\dots,x_j,\dots,x_n) = \int_{-\infty}^{x_j} \frac{\partial u}{\partial x_j}(x_1,\dots,t_j,\dots,x_n)\,dt_j, \quad j=1,\dots,n.$$

This implies that

$$|u(x)| \leqslant \int_{\mathbb{R}} |Du(x_1,\dots,t_j,\dots,x_n)|\,dt_j, \quad j=1,\dots,n.$$

By taking product of the previous estimate for each $j=1,\dots,n$, we obtain

$$|u(x)|^{\frac{n}{n-1}} \leqslant \prod_{j=1}^n \left(\int_{\mathbb{R}} |Du(x_1,\dots,t_j,\dots,x_n)|\,dt_j\right)^{\frac{1}{n-1}}.$$

We integrate with respect to x_1 and then we use generalized Hölder's inequality for the product of $(n-1)$ terms to obtain

$$\begin{aligned}\int_{\mathbb{R}} |u|^{\frac{n}{n-1}}\,dx_1 &\leqslant \left(\int_{\mathbb{R}} |Du|\,dt_1\right)^{\frac{1}{n-1}} \int_{\mathbb{R}} \prod_{j=2}^n \left(\int_{\mathbb{R}} |Du|\,dt_j\right)^{\frac{1}{n-1}} dx_1\\ &\leqslant \left(\int_{\mathbb{R}} |Du|\,dt_1\right)^{\frac{1}{n-1}} \prod_{j=2}^n \left(\int_{\mathbb{R}}\int_{\mathbb{R}} |Du|\,dx_1\,dt_j\right)^{\frac{1}{n-1}}.\end{aligned}$$

Next we integrate with respect to x_2 and use again generalized Hölder's inequality

$$\begin{aligned}\int_{\mathbb{R}}\int_{\mathbb{R}} |u|^{\frac{n}{n-1}}\,dx_1\,dx_2 &\leqslant \int_{\mathbb{R}} \left[\left(\int_{\mathbb{R}} |Du|\,dt_1\right)^{\frac{1}{n-1}} \prod_{j=2}^n \left(\int_{\mathbb{R}}\int_{\mathbb{R}} |Du|\,dx_1\,dt_j\right)^{\frac{1}{n-1}}\right] dx_2\\
&= \left(\int_{\mathbb{R}}\int_{\mathbb{R}} |Du|\,dx_1\,dt_2\right)^{\frac{1}{n-1}}\\
&\qquad \cdot \int_{\mathbb{R}} \left[\left(\int_{\mathbb{R}} |Du|\,dt_1\right)^{\frac{1}{n-1}} \prod_{j=3}^n \left(\int_{\mathbb{R}}\int_{\mathbb{R}} |Du|\,dx_1\,dt_j\right)^{\frac{1}{n-1}}\right] dx_2\\
&\leqslant \left(\int_{\mathbb{R}}\int_{\mathbb{R}} |Du|\,dx_1\,dt_2\right)^{\frac{1}{n-1}}\\
&\qquad \cdot \left(\int_{\mathbb{R}}\int_{\mathbb{R}} |Du|\,dt_1\,dx_2\right)^{\frac{1}{n-1}} \prod_{j=3}^n \left(\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} |Du|\,dx_1\,dx_2\,dt_j\right)^{\frac{1}{n-1}}.\end{aligned}$$

Then we integrate with respect to $x_3,\dots,x_n$ and obtain

$$\begin{aligned}\int_{\mathbb{R}^n} |u|^{\frac{n}{n-1}}\,dx &\leqslant \prod_{j=1}^n \left(\int_{\mathbb{R}} \dots \int_{\mathbb{R}} |Du|\,dx_1 \dots dt_j \dots dx_n\right)^{\frac{1}{n-1}}\\ &= \left(\int_{\mathbb{R}^n} |Du|\,dx\right)^{\frac{n}{n-1}}.\end{aligned}$$

This is the required inequality for $p = 1$.

If $1 < p < n$, we apply the estimate above to

$$v = |u|^\gamma,$$

where $\gamma > 1$ is to be chosen later. Since $\gamma > 1$, we have $v \in C^1(\mathbb{R}^n)$. Hölder's inequality implies

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

Now we choose γ so that $|u|$ has the same power on both sides. Thus

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

This gives

$$\frac{\gamma n}{n-1} = \frac{p(n-1)}{n-p}\frac{n}{n-1} = \frac{pn}{n-p} = p^*$$

and consequently

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

This proves the claim for $u \in C_0^\infty(\mathbb{R}^n)$.

(2) Assume then that $u \in W^{1,p}(\mathbb{R}^n)$. By Lemma 1.28 we have $W^{1,p}(\mathbb{R}^n) = W_0^{1,p}(\mathbb{R}^n)$. Thus there exist $u_i \in C_0^\infty(\mathbb{R}^n)$, $i = 1,2,\dots$, such that $\|u_i - u\|_{W^{1,p}(\mathbb{R}^n)} \to 0$ as $i \to \infty$. In particular $\|u_i - u\|_{L^p(\mathbb{R}^n)} \to 0$, as $i \to \infty$. Thus there exists a subsequence (u_i) such that $u_i \to u$ almost everywhere in $\mathbb{R}^n$ and $u_i \to u$ in $L^p(\mathbb{R}^n)$.

Claim: (u_i) is a Cauchy sequence in $L^{p^*}(\mathbb{R}^n)$.

Reason. Since $u_i - u_j \in C_0^\infty(\mathbb{R}^n)$, we use the Sobolev-Gagliardo-Nirenberg inequality for compactly supported smooth functions and Minkowski's inequality to conclude that

$$\begin{aligned}\|u_i - u_j\|_{L^{p^*}(\mathbb{R}^n)} &\leqslant c\|Du_i - Du_j\|_{L^p(\mathbb{R}^n)}\\ &\leqslant c\left(\|Du_i - Du\|_{L^p(\mathbb{R}^n)} + \|Du - Du_j\|_{L^p(\mathbb{R}^n)}\right) \to 0.\end{aligned}$$

■

Since $L^{p^*}(\mathbb{R}^n)$ is complete, there exists $v \in L^{p^*}(\mathbb{R}^n)$ such that $u_i \to v$ in $L^{p^*}(\mathbb{R}^n)$ as $i \to \infty$.

Since $u_i \to u$ almost everywhere in $\mathbb{R}^n$ and $u_i \to v$ in $L^{p^*}(\mathbb{R}^n)$, we have $u = v$ almost everywhere in $\mathbb{R}^n$. This implies that $u_i \to u$ in $L^{p^*}(\mathbb{R}^n)$ and that $u \in L^{p^*}(\mathbb{R}^n)$.

Now we can apply Minkowski's inequality and the Sobolev-Gagliardo-Nirenberg inequality for compactly supported smooth functions to conclude that

$$
\begin{aligned}
\|u\|_{L^{p^*}(\mathbb{R}^n)} &\leqslant \|u-u_i\|_{L^{p^*}(\mathbb{R}^n)} + \|u_i\|_{L^{p^*}(\mathbb{R}^n)} \\
&\leqslant \|u-u_i\|_{L^{p^*}(\mathbb{R}^n)} + c\|Du_i\|_{L^p(\mathbb{R}^n)} \\
&\leqslant \|u-u_i\|_{L^{p^*}(\mathbb{R}^n)} + c\left(\|Du_i - Du\|_{L^p(\mathbb{R}^n)} + \|Du\|_{L^p(\mathbb{R}^n)}\right) \\
&\xrightarrow{i\to\infty} c\|Du\|_{L^p(\mathbb{R}^n)},
\end{aligned}
$$

since $u_i \to u$ in $L^{p^*}(\mathbb{R}^n)$ and $Du_i \to Du$ in $L^p(\mathbb{R}^n)$. This completes the proof. $\square$

Remarks 3.4:

(1) The Gagliardo-Nirenberg-Sobolev inequality shows that if $u \in W^{1,p}(\mathbb{R}^n)$ with $1 \leqslant p < n$, then $u \in L^p(\mathbb{R}^n) \cap L^{p^*}(\mathbb{R}^n)$, with $p^* > p$.

(2) The Gagliardo-Nirenberg-Sobolev inequality shows that if $u \in W^{1,p}(\mathbb{R}^n)$ with $1 \leqslant p < n$ and $Du = 0$ almost everywhere in $\mathbb{R}^n$, then $u = 0$ almost everywhere in $\mathbb{R}^n$.

(3) The Sobolev-Gagliardo-Nirenberg inequality holds for Sobolev spaces with zero boundary values in open subsets of $\mathbb{R}^n$ by considering the zero extensions. There exists $c = c(n,p) > 0$ such that

$$
\left(\int_\Omega |u|^{p^*}\,dx\right)^{\frac{1}{p^*}} \leqslant c\left(\int_\Omega |Du|^p\,dx\right)^{\frac{1}{p}}
$$

for every $u \in W_0^{1,p}(\Omega)$, $1 \leqslant p < n$. If $|\Omega| < \infty$, by Hölder's inequality

$$
\begin{aligned}
\left(\int_\Omega |u|^q\,dx\right)^{\frac{1}{q}} &\leqslant \left(\int_\Omega |u|^{p^*}\,dx\right)^{\frac{1}{p^*}} |\Omega|^{1-\frac{1}{p^*}} \\
&\leqslant c|\Omega|^{1-\frac{1}{p^*}}\left(\int_\Omega |Du|^p\,dx\right)^{\frac{1}{p}}
\end{aligned}
$$

whenever $1 \leqslant q \leqslant p^*$. Thus for sets with finite measure all exponents below the Sobolev exponent will do.

(4) The Sobolev-Gagliardo-Nirenberg inequality shows that $W^{1,p}_{\text{loc}}(\mathbb{R}^n) \subset L^{p^*}_{\text{loc}}(\mathbb{R}^n)$. To see this, let $\Omega \Subset \mathbb{R}^n$ and choose a cutoff function $\eta \in C_0^\infty(\mathbb{R}^n)$ such that $\eta = 1$ in Ω. Then $\eta u \in W_0^{1,p}(\mathbb{R}^n) = W^{1,p}(\mathbb{R}^n)$ and $\eta u = u$ in Ω and

$$
\|u\|_{L^{p^*}(\Omega)} \leqslant \|\eta u\|_{L^{p^*}(\mathbb{R}^n)} \leqslant c\|D(\eta u)\|_{L^p(\mathbb{R}^n)} < \infty.
$$

(5) The Sobolev-Gagliardo-Nirenberg inequality holds for higher order Sobolev spaces as well. Let $k \in \mathbb{N}$, $1 \leqslant p < \frac{n}{k}$ and $p^* = \frac{np}{n-kp}$. There exists $c = c(n,p,k)$ such that

$$
\left(\int_{\mathbb{R}^n} |u|^{p^*}\,dx\right)^{\frac{1}{p^*}} \leqslant c\left(\int_{\mathbb{R}^n} |D^k u|^p\,dx\right)^{\frac{1}{p}}
$$

for every $u \in W^{k,p}(\Omega)$. Here $|D^k u|^2$ is the sum of squares of all kth order partial derivatives of u (exercise).

The Sobolev–Gagliardo–Nirenberg inequality has the following consequences.

Corollary 3.5. Let $1 \leqslant p < n$ and $p \leqslant q \leqslant p^*$. Then there exists a constant $c = c(n,p)$ such that

$$\|u\|_{L^q(\mathbb{R}^n)} \leqslant (1+c)\|u\|_{W^{1,p}(\mathbb{R}^n)}$$

for every $u \in W^{1,p}(\mathbb{R}^n)$.

THE MORAL: The embedding $L : W^{1,p}(\mathbb{R}^n) \to L^q(\mathbb{R}^n)$, $Lu = u$, is a bounded linear operator.

Proof. The claim is clear if $q = p$ and if $q = p^*$ the claim follows from the Gagliardo-Nirenberg-Sobolev inequality in Theorem 3.3, Thus we may assume that $p < q < p^*$. Let $0 < \theta < 1$ such that $\frac{1}{q} = \frac{\theta}{p} + \frac{1-\theta}{p^*}$. Then

$$\begin{aligned}\|u\|_{L^q(\mathbb{R}^n)} &\leqslant \|u\|^{\theta}_{L^p(\mathbb{R}^n)}\|u\|^{1-\theta}_{L^{p^*}(\mathbb{R}^n)} \\ &\leqslant \|u\|_{L^p(\mathbb{R}^n)} + \|u\|_{L^{p^*}(\mathbb{R}^n)},\end{aligned}$$

where we applied Young's inequality with the exponents $\frac{1}{\theta}$ and $(\frac{1}{\theta})'$. By the Gagliardo-Nirenberg-Sobolev inequality in Theorem 3.3, we have

$$\begin{aligned}\|u\|_{L^q(\mathbb{R}^n)} &\leqslant \|u\|_{L^p(\mathbb{R}^n)} + \|u\|_{L^{p^*}(\mathbb{R}^n)} \\ &\leqslant \|u\|_{L^p(\mathbb{R}^n)} + c\|Du\|_{L^p(\mathbb{R}^n)} \\ &\leqslant (1+c)\|u\|_{W^{1,p}(\mathbb{R}^n)}\end{aligned}$$

for every $u \in W^{1,p}(\mathbb{R}^n)$. □

Corollary 3.6. Let $1 \leqslant p < n$ and let $\Omega \subset \mathbb{R}^n$ be an open set. Assume that $u \in W^{1,p}_0(\Omega)$ is such that $|Du| = 0$ almost everywhere in Ω. Then $u = 0$ almost everywhere in Ω.

Proof. Extend u as zero outside Ω. Then we have $|Du| = 0$ almost everywhere in $\mathbb{R}^n$. Theorem 3.3 implies

$$\|u\|_{L^{p^*}(\mathbb{R}^n)} \leqslant c\|Du\|_{L^p(\mathbb{R}^n)} = 0.$$

It follows that $u = 0$ almost everywhere in $\mathbb{R}^n$, and thus almost everywhere in Ω. □

Since $p^* = \frac{np}{n-p} \to \infty$ as $p \to n-$, one might expect that $W^{1,n}(\Omega)$ would be continuously embedded in $L^\infty(\Omega)$. This is false for $n > 1$. Let $\Omega = B(0,1) \subset \mathbb{R}^n$. The function

$$u(x) = \log\left(\log\left(1 + \frac{1}{|x|}\right)\right)$$

belongs to $W^{1,n}(\Omega)$ but not $L^\infty(\Omega)$ (exercise).

The following result is a version of the Sobolev inequality for the full range $1 \leqslant p < \infty$.

Corollary 3.7. Let $1 \leqslant p < \infty$, let $\Omega \subset \mathbb{R}^n$ be an open set with $|\Omega| < \infty$, and assume that $u \in W_0^{1,p}(\Omega)$. Let $1 \leqslant q \leqslant p^* = \frac{np}{n-p}$, for $1 \leqslant p < n$, and $1 \leqslant q < \infty$ for $n \leqslant p < \infty$. There exists a constant $c = c(n,p,q)$ such that

$$\Big(\int_\Omega |u|^q\,dx\Big)^{\frac{1}{q}} \leqslant c|\Omega|^{\frac{1}{n}-\frac{1}{p}+\frac{1}{q}}\Big(\int_\Omega |Du|^p\,dx\Big)^{\frac{1}{p}}.$$

THE MORAL: Let $\Omega \subset \mathbb{R}^n$ be an open set with $|\Omega| < \infty$. If $u \in W_0^{1,p}(\Omega)$ with $p \geqslant n$, then $u \in L^q(\Omega)$ for every q with $1 \leqslant q < \infty$.

Proof. Extend u as zero outside Ω. Then $Du(x) = 0$ for almost every $x \in \Omega^c$. Assume first that $1 \leqslant p < n$. Hölder's inequality and Theorem 3.3 imply

$$\begin{aligned}\Big(\int_\Omega |u|^q\,dx\Big)^{\frac{1}{q}} &\leqslant |\Omega|^{\frac{1}{n}-\frac{1}{p}+\frac{1}{q}}\Big(\int_\Omega |u|^{\frac{np}{n-p}}\,dx\Big)^{\frac{n-p}{np}}\\ &\leqslant c(n,p)|\Omega|^{\frac{1}{n}-\frac{1}{p}+\frac{1}{q}}\Big(\int_\Omega |Du|^p\,dx\Big)^{\frac{1}{p}}.\end{aligned}$$

Assume then that $n \leqslant p < \infty$. If $q > p$, choose $1 < \widetilde{p} < n$ satisfying $q = \frac{n\widetilde{p}}{n-\widetilde{p}}$. By the first part of the proof and Hölder's inequality, we obtain

$$\begin{aligned}\Big(\int_\Omega |u|^q\,dx\Big)^{\frac{1}{q}} &\leqslant c(n,p,q)|\Omega|^{\frac{1}{n}-\frac{1}{\widetilde{p}}+\frac{1}{q}}\Big(\int_\Omega |Du|^{\widetilde{p}}\,dx\Big)^{\frac{1}{\widetilde{p}}}\\ &\leqslant c(n,p,q)|\Omega|^{\frac{1}{n}-\frac{1}{p}+\frac{1}{q}}\Big(\int_\Omega |Du|^p\,dx\Big)^{\frac{1}{p}}.\end{aligned}$$

Finally, if $q \leqslant p$, the claim follows from the previous case for some $\widetilde{q} > q$ and Hölder's inequality on the left-hand side. □

Remark 3.8. Let $1 \leqslant p < n$ and let $\Omega \subset \mathbb{R}^n$ be an open set with $|\Omega| < \infty$. The proof of Corollary 3.7 shows that the Sobolev inequality

$$\Big(\int_\Omega |u(x)|^{\frac{np}{n-p}}\,dx\Big)^{\frac{n-p}{np}} \leqslant c(n,p)\Big(\int_\Omega |Du(x)|^p\,dx\Big)^{\frac{1}{p}}$$

holds for every $u \in W_0^{1,p}(\Omega)$.

Remark 3.9. When $p = 1$ the Sobolev-Gagliardo-Nirenberg inequality is related to the isoperimetric inequality. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with smooth boundary and set

$$u_\varepsilon(x) = \begin{cases} 1, & x \in \Omega,\\ 1 - \frac{\operatorname{dist}(x,\Omega)}{\varepsilon}, & 0 < \operatorname{dist}(x,\Omega) < \varepsilon,\\ 0, & \operatorname{dist}(x,\Omega) \geqslant \varepsilon.\end{cases}$$

Note that u can been considered as an approximation of the characteristic function of Ω. The Lipschitz constant of $x \mapsto \operatorname{dist}(x,\Omega)$ is one so that the Lipschitz constant of u_ε is ε^{-1} and thus this function belongs to $W^{1,1}(\mathbb{R}^n)$, for example, by the ACL characterization, see Theorem 2.36, we have

$$|Du_\varepsilon(x)| \leqslant \begin{cases} \frac{1}{\varepsilon}, & 0 < \operatorname{dist}(x,\Omega) < \varepsilon, \\ 0, & \text{otherwise.} \end{cases}$$

The Sobolev-Gagliardo-Nirenberg inequality with $p = 1$ gives

$$\begin{aligned} |\Omega|^{\frac{n-1}{n}} &= \left(\int_\Omega |u_\varepsilon|^{\frac{n}{n-1}} \, dx \right)^{\frac{n-1}{n}} \leqslant \left(\int_{\mathbb{R}^n} |u_\varepsilon|^{\frac{n}{n-1}} \, dx \right)^{\frac{n-1}{n}} \\ &\leqslant c \int_{\mathbb{R}^n} |Du_\varepsilon| \, dx \leqslant c \int_{\{0<\operatorname{dist}(x,\Omega)<\varepsilon\}} \frac{1}{\varepsilon} \, dx \\ &= c \frac{|\{x \in \mathbb{R}^n : 0 < \operatorname{dist}(x,\Omega) < \varepsilon\}|}{\varepsilon} \to cH^{n-1}(\partial\Omega) \end{aligned}$$

This implies

$$|\Omega|^{\frac{n}{n-1}} \leqslant c\mathscr{H}^{n-1}(\partial\Omega),$$

which is an isoperimetric inequality with the same constant c as in the Sobolev-Gagliardo-Nirenberg inequality. According to the classical isoperimetric inequality, if $\Omega \subset \mathbb{R}^n$ is a bounded domain with smooth boundary, then

$$|\Omega|^{\frac{n-1}{n}} \leqslant n^{-1}\Omega_n^{-\frac{1}{n}} \mathscr{H}^{n-1}(\partial\Omega),$$

where $\mathscr{H}^{n-1}(\partial\Omega)$ stands for the $(n-1)$-dimensional Hausdorff measure of the boundary $\partial\Omega$. The isoperimetric inequality is equivalent with the statement that among all smooth bounded domains with fixed volume, balls have the least surface area.

Conversely, the Sobolev-Gagliardo-Nirenberg inequality can be proved by the isoperimetric inequality, but we shall not consider this argument here. From these considerations it is relatively obvious that the best constant in the Sobolev-Gagliardo-Nirenberg when $p = 1$ should be the isoperimetric constant $n^{-1}\Omega_n^{-\frac{1}{n}}$. This also gives a geometric motivation for the Sobolev exponent in the case $p = 1$.

3.2 Sobolev-Poincaré inequalities

We begin with a Poincaré inequality for Sobolev functions with zero boundary values in open subsets.

Theorem 3.10 (Poincaré). Assume that $\Omega \subset \mathbb{R}^n$ is bounded and $1 \leqslant p < \infty$. Then there exists a constant $c = c(p)$ such that

$$\int_\Omega |u|^p \, dx \leqslant c \operatorname{diam}(\Omega)^p \int_\Omega |Du|^p \, dx$$

for every $u \in W_0^{1,p}(\Omega)$.

THE MORAL: This shows that $W_0^{1,p}(\Omega) \subset L^p(\Omega)$ when $1 \leqslant p < \infty$, if $\Omega \subset \mathbb{R}^n$ is bounded. The main difference compared to the Gagliardo-Nirenberg-Sobolev inequality is that this applies for the whole range $1 \leqslant p < \infty$ without the Sobolev exponent.

Remark 3.11. The Poincaré inequality above also shows that if $Du = 0$ almost everywhere, then $u = 0$ almost everywhere. For this it is essential that the function belongs to the Sobolev space with zero boundary values.

Proof. (1) First assume that $u \in C_0^\infty(\Omega)$. Let $y = (y_1, \dots, y_n) \in \Omega$. Then

$$\Omega \subset \prod_{j=1}^{n} [y_j - \operatorname{diam}(\Omega), y_j + \operatorname{diam}(\Omega)] = \prod_{j=1}^{n} [a_j, b_j],$$

where $a_j = y_j - \operatorname{diam}(\Omega)$ and $b_j = y_j + \operatorname{diam}(\Omega)$, $j = 1, \dots, n$. As the proof of Theorem 3.3, we obtain

$$\begin{aligned} |u(x)| &\leqslant \int_{a_j}^{b_j} |Du(x_1, \dots, t_j, \dots, x_n)|\, dt_j \\ &\leqslant (2\operatorname{diam}(\Omega))^{1-\frac{1}{p}} \left(\int_{a_j}^{b_j} |Du(x_1, \dots, t_j, \dots, x_n)|^p \, dt_j \right)^{\frac{1}{p}}, \quad j = 1, \dots, n. \end{aligned}$$

The second inequality follows from Hölder's inequality. Thus

$$\begin{aligned} \int_\Omega |u(x)|^p \, dx &= \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} |u(x)|^p \, dx_1 \dots dx_n \\ &\leqslant (2\operatorname{diam}(\Omega))^{p-1} \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} \int_{a_1}^{b_1} |Du(t_1, x_2, \dots, x_n)|^p \, dt_1 \, dx_1 \dots dx_n \\ &\leqslant (2\operatorname{diam}(\Omega))^p \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} |Du(t_1, x_2, \dots, x_n)|^p \, dt_1 \dots dx_n \\ &= (2\operatorname{diam}(\Omega))^p \int_\Omega |Du(x)|^p \, dx. \end{aligned}$$

(2) The case $u \in W_0^{1,p}(\Omega)$ follows by approximation (exercise). □

The Gagliardo-Nirenberg-Sobolev inequality in Theorem 3.3 and Poincaré's inequality in Theorem 3.10 do not hold for functions $u \in W^{1,p}(\Omega)$, at least when $\Omega \subset \mathbb{R}^n$ is an open set $|\Omega| < \infty$, since nonzero constant functions give obvious counterexamples. However, there are several ways to obtain appropriate local estimates also in this case.

Next we consider estimates in the case when Ω is a cube. Later we consider similar estimates for balls. The set

$$Q = [a_1, b_1] \times \dots \times [a_n, b_n], \quad b_1 - a_1 = \dots = b_n - a_n$$

is a cube in $\mathbb{R}^n$. The side length of Q is

$$l(Q) = b_1 - a_1 = b_j - a_j, \quad j = 1, \dots, n,$$

and

$$Q(x,l) = \left\{ y \in \mathbb{R}^n : |y_j - x_j| \leqslant \tfrac{l}{2}, j = 1, \dots, n \right\}$$

is the cube with center x and sidelength l. Clearly,

$$|Q(x,l)| = l^n \quad \text{and} \quad \operatorname{diam}(Q(x,l)) = \sqrt{n}\, l$$

The integral average of $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ over cube $Q(x,l)$ is denoted by

$$f_{Q(x,l)} = \fint_{Q(x,l)} f \, dy = \frac{1}{|Q(x,l)|} \int_{Q(x,l)} f(y) \, dy.$$

Same notation is used for integral averages over other sets as well.

Theorem 3.12 (Poincaré inequality on cubes). Let Ω be an open subset of $\mathbb{R}^n$. Assume that $u \in W^{1,p}_{\text{loc}}(\Omega)$ with $1 \leqslant p < \infty$. Then there exists a constant $c = c(n,p)$ such that

$$\left(\fint_{Q(x,l)} |u - u_{Q(x,l)}|^p \, dy \right)^{\frac{1}{p}} \leqslant c l \left(\fint_{Q(x,l)} |Du|^p \, dy \right)^{\frac{1}{p}}$$

for every cube $Q(x,l) \Subset \Omega$.

T H E   M O R A L : The Poincaré inequality shows that if the gradient is small in a cube, then the mean oscillation of the function is small in the same cube. In particular, if the gradient is zero, then the function is constant.

Proof. (1) First assume that $u \in C^\infty(\Omega)$. Let $z, y \in Q = Q(x,l) = [a_1, b_1] \times \dots \times [a_n, b_n]$. Then

$$\begin{aligned} |u(z) - u(y)| &\leqslant |u(z) - u(z_1, \dots, z_{n-1}, y_n)| + \ldots + |u(z_1, y_2, \dots, y_n) - u(y)| \\ &\leqslant \sum_{j=1}^n \int_{a_j}^{b_j} |Du(z_1, \dots, z_{j-1}, t, y_{j+1}, \dots, y_n)| \, dt \end{aligned}$$

By Hölder's inequality and the elementary inequality $(a_1 + \dots + a_n)^p \leqslant n^p (a_1^p + \dots + a_n^p)$, $a_i \geqslant 0$, we obtain

$$\begin{aligned} |u(z) - u(y)|^p &\leqslant \left(\sum_{j=1}^n \int_{a_j}^{b_j} |Du(z_1, \dots, z_{j-1}, t, y_{j+1}, \dots, y_n)| \, dt \right)^p \\ &\leqslant \left(\sum_{j=1}^n \left(\int_{a_j}^{b_j} |Du(z_1, \dots, z_{j-1}, t, y_{j+1}, \dots, y_n)|^p \, dt \right)^{\frac{1}{p}} (b_j - a_j)^{1 - \frac{1}{p}} \right)^p \\ &\leqslant n^p l^{p-1} \sum_{j=1}^n \int_{a_j}^{b_j} |Du(z_1, \dots, z_{j-1}, t, y_{j+1}, \dots, y_n)|^p \, dt. \end{aligned}$$

By Hölder's inequality and Fubini's theorem

$$\begin{aligned}
\int_Q |u(z)-u_Q|^p\,dz &= \int_Q \left| \fint_Q (u(z)-u(y))\,dy \right|^p dz \\
&\leqslant \int_Q \left(\fint_Q |u(z)-u(y)|\,dy \right)^p dz \leqslant \int_Q \fint_Q |u(z)-u(y)|^p\,dz\,dy \\
&\leqslant \frac{n^p l^{p-1}}{|Q|} \sum_{j=1}^n \int_Q \int_Q \int_{a_j}^{b_j} |Du(z_1,\dots,z_{j-1},t,y_{j+1},\dots,y_n)|^p\,dt\,dy\,dz \\
&\leqslant \frac{n^p l^{p-1}}{|Q|} \sum_{j=1}^n (b_j-a_j) \int_Q \int_Q |Du(z)|^p\,dz\,dw \\
&\leqslant n^{p+1} l^p \int_Q |Du(z)|^p\,dz.
\end{aligned}$$

(2) The case $u \in W^{1,p}_{\mathrm{loc}}(\Omega)$ follows by approximation. There exist $u_i \in C^\infty(\mathbb{R}^n)$, $i \in \mathbb{N}$, satisfying $u_i \to u$ in $W^{1,p}(Q)$ as $i \to \infty$. By passing to a subsequence, if necessary, we may in addition assume that $u_i \to u$ almost everywhere in Q. Moreover, it follows from Hölder's inequality and the L^p convergence that

$$|(u_i)_Q - u_Q| \leqslant \fint_Q |u_i(x)-u(x)|\,dx \leqslant \left(\fint_Q |u_i(x)-u(x)|^p\,dx \right)^{\frac{1}{p}} \to 0$$

and thus $(u_i)_Q \to u_Q$ as $i \to \infty$. Fatou's lemma and the first part of the proof for $u_i \in C^\infty(\Omega)$ give

$$\begin{aligned}
\left(\fint_Q |u-u_Q|^p\,dx \right)^{\frac{1}{p}} &\leqslant \liminf_{i\to\infty} \left(\fint_Q |u_i-(u_i)_Q|^p\,dx \right)^{\frac{1}{p}} \\
&\leqslant \liminf_{i\to\infty} c(n,p,q)\, l \left(\fint_Q |Du_i|^p\,dx \right)^{\frac{1}{p}} \\
&\leqslant c(n,p,q)\, l \left(\fint_Q |Du|^p\,dx \right)^{\frac{1}{p}},
\end{aligned}$$

and the proof is complete. □

Theorem 3.13 (Sobolev–Poincaré inequality on cubes). Let Ω be an open subset of $\mathbb{R}^n$. Assume that $u \in W^{1,p}_{\mathrm{loc}}(\Omega)$ with $1 \leqslant p < n$. Then there exists a constant $c = c(n,p)$ such that

$$\left(\fint_{Q(x,l)} |u-u_{Q(x,l)}|^{p^*}\,dy \right)^{\frac{1}{p^*}} \leqslant c\,l \left(\fint_{Q(x,2l)} |Du|^p\,dy \right)^{\frac{1}{p}}$$

for every cube $Q(x,2l) \Subset \Omega$.

THE MORAL: The Sobolev-Poincaré inequality shows that $W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n) \subset L^{p^*}_{\mathrm{loc}}(\mathbb{R}^n)$, when $1 \leqslant p < n$. This is a stronger version of the Poincaré inequality on cubes in which we have the Sobolev exponent on the left-hand side.

Proof. Let $\eta \in C_0^\infty(\mathbb{R}^n)$ be a cutoff function such that

$$0 \leqslant \eta \leqslant 1, \quad |D\eta| \leqslant \tfrac{c}{l}, \quad \operatorname{supp}\eta \subset Q(x,2l) \quad \text{and} \quad \eta = 1 \text{ in } Q(x,l).$$

Notice that the constant $c = c(n)$ does not depend on the cube. Then $(u - u_{Q(x,l)})\eta \in W^{1,p}(\mathbb{R}^n)$ and by the Gagliardo-Nirenberg-Sobolev inequality, see Theorem 3.3, and the Leibniz rule, see Theorem 1.14 (5), we have

$$\begin{aligned}
\Big(\int_{Q(x,l)} |u - u_{Q(x,l)}|^{p^*}\,dy\Big)^{\frac{1}{p^*}} &\leqslant \Big(\int_{\mathbb{R}^n} |(u - u_{Q(x,l)})\eta|^{p^*}\,dy\Big)^{\frac{1}{p^*}} \\
&\leqslant c\Big(\int_{\mathbb{R}^n} \big|D\big[(u - u_{Q(x,l)})\eta\big]\big|^p\,dy\Big)^{\frac{1}{p}} \\
&\leqslant c\Big(\int_{\mathbb{R}^n} \eta^p |Du|^p\,dy\Big)^{\frac{1}{p}} + c\Big(\int_{\mathbb{R}^n} |D\eta|^p |u - u_{Q(x,l)}|^p\,dy\Big)^{\frac{1}{p}} \\
&\leqslant c\Big(\int_{Q(x,2l)} |Du|^p\,dy\Big)^{\frac{1}{p}} + \frac{c}{l}\Big(\int_{Q(x,2l)} |u - u_{Q(x,l)}|^p\,dy\Big)^{\frac{1}{p}}.
\end{aligned}$$

By the Poincaré inequality on cubes, see Theorem 3.12, we obtain

$$\begin{aligned}
&\Big(\int_{Q(x,2l)} |u - u_{Q(x,l)}|^p\,dy\Big)^{\frac{1}{p}} \\
&\qquad\leqslant \Big(\int_{Q(x,2l)} |u - u_{Q(x,2l)}|^p\,dy\Big)^{\frac{1}{p}} + \Big(\int_{Q(x,2l)} |u_{Q(x,2l)} - u_{Q(x,l)}|^p\,dy\Big)^{\frac{1}{p}} \\
&\qquad\leqslant cl\Big(\int_{Q(x,2l)} |Du|^p\,dy\Big)^{\frac{1}{p}} + |u_{Q(x,2l)} - u_{Q(x,l)}||Q(x,2l)|^{\frac{1}{p}}.
\end{aligned}$$

By Hölder's inequality and Poincaré inequality on cubes, see Theorem 3.12, we have

$$\begin{aligned}
|u_{Q(x,2l)} - u_{Q(x,l)}||Q(x,2l)|^{\frac{1}{p}} &\leqslant (2l)^{\frac{n}{p}} \fint_{Q(x,l)} |u - u_{Q(x,2l)}|\,dy \\
&\leqslant (2l)^{\frac{n}{p}} \frac{|Q(x,2l)|}{|Q(x,l)|}\Big(\fint_{Q(x,2l)} |u - u_{Q(x,2l)}|^p\,dy\Big)^{\frac{1}{p}} \\
&\leqslant cl\Big(\int_{Q(x,2l)} |Du|^p\,dy\Big)^{\frac{1}{p}}.
\end{aligned}$$

By collecting the estimates above we obtain

$$\Big(\int_{Q(x,l)} |u - u_{Q(x,l)}|^{p^*}\,dy\Big)^{\frac{1}{p^*}} \leqslant c\Big(\int_{Q(x,2l)} |Du|^p\,dy\Big)^{\frac{1}{p}}.$$

□

Remark 3.14. The Sobolev-Poincaré inequality also holds in the form

$$\Big(\fint_{Q(x,l)} |u - u_{Q(x,l)}|^{p^*}\,dy\Big)^{\frac{1}{p^*}} \leqslant cl\Big(\fint_{Q(x,l)} |Du|^p\,dy\Big)^{\frac{1}{p}}.$$

Observe that there is the same cube on both sides. We shall return to this question later.

Remark 3.15. The Sobolev–Poincaré inequality in Theorem 3.13 holds with the same cubes on the both sides and it holds also for $p = 1$. We shall not consider these versions here.

Remark 3.16. In this remark we consider the case $p = n$.

(1) As Example 1.12 shows, functions in $W^{1,n}(\mathbb{R}^n)$ are not necessarily bounded.

(2) Assume that $u \in W^{1,n}(\mathbb{R}^n)$. The Poincaré inequality implies that

$$\begin{aligned}\fint_Q |u(y)-u_Q|\,dy &\leqslant \left(\fint_Q |u(y)-u_Q|^n\,dy\right)^{\frac{1}{n}}\\ &\leqslant cl\left(\fint_Q |Du(y)|^n\,dy\right)^{\frac{1}{n}}\\ &\leqslant c\|Du\|_{L^n(\mathbb{R}^n)} < \infty\end{aligned}$$

for every cube Q where $c = c(n)$. Thus if $u \in W^{1,n}(\mathbb{R}^n)$, then u is of bounded mean oscillation, denoted by $u \in \mathrm{BMO}(\mathbb{R}^n)$, and

$$\|u\|_* = \sup_{Q\subset\mathbb{R}^n} \fint_Q |u(y)-u_Q|\,dy \leqslant c\|Du\|_{L^n(\mathbb{R}^n)},$$

where $c = c(n)$.

(3) Assume that $u \in W^{1,n}(\mathbb{R}^n)$. The John-Nirenberg inequality for BMO functions gives

$$\fint_Q e^{\gamma|u(x)-u_Q|}\,dx \leqslant \frac{c_1\gamma\|u\|_*}{c_2-\gamma\|u\|_*}+1$$

for every cube Q in $\mathbb{R}^n$ with $0<\gamma<\frac{c_2}{\|u\|_*}$, where $c_1 = c_1(n)$ and $c_2 = c_2(n)$. By choosing $\gamma = \frac{c_2}{2\|u\|_*}$, we obtain

$$\fint_Q e^{c\frac{|u(x)-u_Q|}{\|Du\|_n}}\,dx \leqslant \fint_Q e^{c\frac{|u(x)-u_Q|}{\|u\|_*}}\,dx \leqslant c$$

for every cube Q in $\mathbb{R}^n$. In particular, this implies that $u \in L^p_{\mathrm{loc}}(\mathbb{R}^n)$ for every power p, with $1 \leqslant p < \infty$. This is the Sobolev embedding theorem in the borderline case when $p = n$.

In fact, there is a stronger result called Trudinger's inequality, which states that for small enough $c > 0$, we have

$$\fint_Q e^{\left(c\frac{|u(x)-u_Q|}{\|Du\|_n}\right)^{\frac{n}{n-1}}}\,dx \leqslant c$$

for every cube Q in $\mathbb{R}^n$, $n \geqslant 2$, but we shall not discuss this issue here.

T H E M O R A L : $W^{1,n}(\mathbb{R}^n) \subset L^p_{\mathrm{loc}}(\mathbb{R}^n)$ for every p, with $1 \leqslant p < \infty$. This is the Sobolev embedding theorem in the borderline case when $p = n$.

The next theorem gives a general Sobolev–Poincaré inequality for Sobolev functions.

Theorem 3.17. Let $1 < p < \infty$, let $\Omega \subset \mathbb{R}^n$ be an open set, and assume that $u \in W^{1,p}_{\mathrm{loc}}(\Omega)$. Let $1 \leqslant q \leqslant p^* = \frac{np}{n-p}$, for $1 < p < n$, and $1 \leqslant q < \infty$ for $n \leqslant p < \infty$. There exists a constant $c = c(n,p,q)$ such that

$$\left(\fint_{Q(x,l)} |u - u_{Q(x,l)}|^q \, dy \right)^{\frac{1}{q}} \leqslant cl \left(\fint_{Q(x,2l)} |Du|^p \, dy \right)^{\frac{1}{p}} \tag{3.18}$$

for every cube $Q(x,2l) \Subset \Omega$.

Proof. By Theorem 3.13, for $1 < p < n$, we obtain

$$\begin{aligned} \left(\fint_{Q(x,l)} |u - u_{Q(x,l)}|^{\frac{np}{n-p}} \, dy \right)^{\frac{n-p}{np}} &= c(n,p) l^{-\frac{n-p}{p}} \left(\int_{Q(x,l)} |u - u_{Q(x,l)}|^{\frac{np}{n-p}} \, dy \right)^{\frac{n-p}{np}} \\ &\leqslant c(n,p) l^{1-\frac{n}{p}} \left(\int_{Q(x,2l)} |Du|^p \, dy \right)^{\frac{1}{p}} \\ &= c(n,p) l \left(\fint_{Q(x,2l)} |Du|^p \, dy \right)^{\frac{1}{p}}. \end{aligned} \tag{3.19}$$

For $1 < p < n$, inequality (3.18) follows from (3.19) and Hölder's inequality on the left-hand side.

In the case $p \geqslant n$ we proceed as in the proof of Corollary 3.7. For $q > p \geqslant n$, there exists $1 < \tilde{p} < n$ such that $q = \frac{n\tilde{p}}{n-\tilde{p}}$, and (3.18) follows from (3.19) with exponent $\tilde{p}$ and an application of Hölder's inequality on the right-hand side. For $q \leqslant p$, the claim follows from the previous case and Hölder's inequality on the left-hand side. □

The next remark shows that it is possible to obtain a Poincaré inequality on cubes without the integral average also for functions that do not have zero boundary values. However, the functions have to vanish in a large subset.

Remark 3.20. Assume $u \in W^{1,p}(\mathbb{R}^n)$ and $u = 0$ in a set $A \subset Q(x,l) = Q$ satisfying

$$|A| \geqslant \gamma |Q| \quad \text{for some} \quad 0 < \gamma \leqslant 1.$$

This means that $u = 0$ in a large portion of Q. By the Poincaré inequality there exists $c = c(n,p)$ such that

$$\begin{aligned} \left(\fint_Q |u|^p \, dy \right)^{\frac{1}{p}} &\leqslant \left(\fint_Q |u - u_Q|^p \, dy \right)^{\frac{1}{p}} + \left(\fint_Q |u_Q|^p \, dy \right)^{\frac{1}{p}} \\ &\leqslant cl \left(\fint_Q |Du|^p \, dy \right)^{\frac{1}{p}} + |u_Q|, \end{aligned}$$

where

$$\begin{aligned} |u_Q| = \left| \fint_Q u(y) \, dy \right| &\leqslant \fint_Q \chi_{Q \setminus A}(y) |u(y)| \, dy \\ &\leqslant \left(\frac{|Q \setminus A|}{|Q|} \right)^{1-\frac{1}{p}} \left(\fint_Q |u(y)|^p \, dy \right)^{\frac{1}{p}} \\ &\leqslant (1-\gamma)^{1-\frac{1}{p}} \left(\fint_Q |u(y)|^p \, dy \right)^{\frac{1}{p}}. \end{aligned}$$

Since $0 \leqslant (1-\gamma)^{1-\frac{1}{p}} < 1$, we may absorb the integral average to the left hand side and obtain

$$(1-(1-\gamma)^{1-\frac{1}{p}})\left(\fint_Q |u|^p\,dy\right)^{\frac{1}{p}} \leqslant cl\left(\fint_Q |Du|^p\,dy\right)^{\frac{1}{p}}.$$

It follows that there exists $c = c(n,p,\gamma)$ such that

$$\left(\fint_Q |u|^p\,dy\right)^{\frac{1}{p}} \leqslant cl\left(\fint_Q |Du|^p\,dy\right)^{\frac{1}{p}}.$$

A similar argument can be done with the Sobolev-Poincaré inequality on cubes (exercise).

3.3 Morrey's inequality

Let $A \subset \mathbb{R}^n$. A function $u : A \to \mathbb{R}$ is Hölder continuous with exponent $0 < \alpha \leqslant 1$, if there exists a constant c such that

$$|u(x)-u(y)| \leqslant c|x-y|^\alpha$$

for every $x, y \in A$. We define the space $C^{0,\alpha}(A)$ to be the space of all bounded functions that are Hölder continuous with exponent α with the norm

$$\|u\|_{C^{0,\alpha}(A)} = \sup_{x\in A}|u(x)| + \sup_{x,y\in A, x\neq y}\frac{|u(x)-u(y)|}{|x-y|^\alpha}. \tag{3.21}$$

Remarks 3.22:

(1) Every function that is Hölder continuous with exponent $\alpha > 1$ in the whole space is constant (exercise).

(2) There are Hölder continuous functions that are not differentiable at any point. Thus Hölder continuity does not imply any differentiability properties.

(3) $C^{0,\alpha}(A)$ is a Banach space with the norm defined above (exercise).

(4) Every Hölder continuous function on $A \subset \mathbb{R}^n$ can be extended to a Hölder continuous function on $\mathbb{R}^n$ with the same exponent and same constant. Moreover, if A is bounded, we may assume that the Hölder continuous extension to $\mathbb{R}^n$ is bounded (exercise).

The next result shows that every function in $W^{1,p}(\mathbb{R}^n)$ with $p > n$ has a $\left(1-\frac{n}{p}\right)$-Hölder continuous representative up to a set of measure zero.

Theorem 3.23 (Morrey). Assume that $u \in W^{1,p}(\mathbb{R}^n)$ with $p > n$. Then there exists a constant $c = c(n,p)$ such that

$$|u(z)-u(y)| \leqslant c|z-y|^{1-\frac{n}{p}}\|Du\|_{L^p(\mathbb{R}^n)}$$

for almost every $z, y \in \mathbb{R}^n$.

Proof. (1) Assume first that $u \in C^\infty(\mathbb{R}^n) \cap W^{1,p}(\mathbb{R}^n)$. Let $z, y \in Q(x,l)$. Then

$$u(z) - u(y) = \int_0^1 \frac{\partial}{\partial t}\big(u(tz + (1-t)y)\big)\,dt = \int_0^1 Du(tz + (1-t)y)\cdot(z-y)\,dt$$

and

$$\begin{aligned}
|u(y) - u_{Q(x,l)}| &= \left|\fint_{Q(x,l)} (u(z) - u(y))\,dz\right| \\
&= \left|\fint_{Q(x,l)} \int_0^1 Du(tz + (1-t)y)\cdot(z-y)\,dt\,dz\right| \\
&\leqslant \sum_{j=1}^n \frac{1}{l^n} \int_{Q(x,l)} \int_0^1 \left|\frac{\partial u}{\partial x_j}(tz + (1-t)y)\right| |z_j - y_j|\,dt\,dz \\
&\leqslant \sum_{j=1}^n \frac{1}{l^{n-1}} \int_0^1 \int_{Q(x,l)} \left|\frac{\partial u}{\partial x_j}(tz + (1-t)y)\right| dz\,dt \\
&= \sum_{j=1}^n \frac{1}{l^{n-1}} \int_0^1 \frac{1}{t^n} \int_{Q(tx+(1-t)y,tl)} \left|\frac{\partial u}{\partial x_j}(w)\right| dw\,dt.
\end{aligned}$$

Here we used the fact that $|z_j - y_j| \leqslant l$, Fubini's theorem and finally the change of variables $w = tz + (1-t)y \Longleftrightarrow z = \frac{1}{t}(w - (1-t)y)$, $dz = \frac{1}{t^n}\,dw$. By Hölder's inequality

$$\begin{aligned}
&\sum_{j=1}^n \frac{1}{l^{n-1}} \int_0^1 \frac{1}{t^n} \int_{Q(tx+(1-t)y,tl)} \left|\frac{\partial u}{\partial x_j}(w)\right| dw\,dt \\
&\quad\leqslant \sum_{j=1}^n \frac{1}{l^{n-1}} \int_0^1 \frac{1}{t^n} \left(\int_{Q(tx+(1-t)y,tl)} \left|\frac{\partial u}{\partial x_j}(w)\right|^p dw\right)^{\frac{1}{p}} |Q(tx + (1-t)y, tl)|^{\frac{1}{p'}}\,dt \\
&\quad\leqslant n\|Du\|_{L^p(Q(x,l))} \frac{l^{n(1-\frac{1}{p})}}{l^{n-1}} \int_0^1 \frac{t^{n(1-\frac{1}{p})}}{t^n}\,dt \qquad (Q(tx + (1-t)y, tl) \subset Q(x,l))) \\
&\quad= \frac{np}{p-n} l^{1-\frac{n}{p}} \|Du\|_{L^p(Q(x,l))}.
\end{aligned}$$

Thus

$$\begin{aligned}
|u(z) - u(y)| &\leqslant |u(z) - u_{Q(x,l)}| + |u_{Q(x,l)} - u(y)| \\
&\leqslant 2\frac{np}{p-n} l^{1-\frac{n}{p}} \|Du\|_{L^p(Q(x,l))}
\end{aligned} \tag{3.24}$$

for every $z, y \in Q(x,l)$.

For every $z, y \in \mathbb{R}^n$, there exists a cube $Q(x,l) \ni z, y$ such that $l = |z-y|$. For example, we may choose $x = \frac{z+y}{2}$. Thus

$$|u(z) - u(y)| \leqslant c|z-y|^{1-\frac{n}{p}} \|Du\|_{L^p(Q(x,l))} \leqslant c|z-y|^{1-\frac{n}{p}} \|Du\|_{L^p(\mathbb{R}^n)}$$

for every $z, y \in \mathbb{R}^n$.

(2) Assume then that $u \in W^{1,p}(\mathbb{R}^n)$. Let u_ε be the standard mollification of u. Then

$$|u_\varepsilon(z) - u_\varepsilon(y)| \leqslant c|z-y|^{1-\frac{n}{p}} \|Du_\varepsilon\|_{L^p(\mathbb{R}^n)}.$$

Now by Lemma 1.18 (2) and by Theorem 1.19, we obtain

$$|u(z)-u(y)| \leqslant c|z-y|^{1-\frac{n}{p}} \|Du\|_{L^p(\mathbb{R}^n)}.$$

when z and y are Lebesgue points of u. The claim follows from the fact that almost every point of a locally integrable function is a Lebesgue point. □

Remarks 3.25:

(1) Morrey's inequality implies that u can be extended uniquely to $\mathbb{R}^n$ as a Hölder continuous function $\overline{u}$ such that

$$|\overline{u}(x)-\overline{u}(y)| \leqslant c|x-y|^{1-\frac{n}{p}} \|Du\|_{L^p(\mathbb{R}^n)} \quad \text{for all } x,y \in \mathbb{R}^n.$$

Reason. Let N be a set of zero measure such that Morrey's inequality holds for all points in $\mathbb{R}^n \setminus N$. Now for any $x \in \mathbb{R}^n$, choose a sequence of points (x_i) such that $x_i \in \mathbb{R}^n \setminus N$, $i = 1,2\dots$, and $x_i \to x$ as $i \to \infty$. By Morrey's inequality $(u(x_i))$ is a Cauchy sequence in $\mathbb{R}$ and thus we can define

$$\overline{u}(x) = \lim_{i\to\infty} u(x_i).$$

Now it is easy to check that $\overline{u}$ satisfies Morrey's inequality in every pair of points by considering sequences of points in $\mathbb{R}^n \setminus N$ converging to the pair of points. ■

(2) If $u \in W^{1,p}(\mathbb{R}^n)$ with $p > n$, then u is essentially bounded.

Reason. Let $y \in Q(x,1)$. Then Morrey's and Hölder's inequality imply

$$\begin{aligned}
|u(z)| &\leqslant |u(z)-u_{Q(x,1)}| + |u_{Q(x,1)}| \\
&\leqslant \fint_{Q(x,1)} |u(z)-u(y)|\,dy + \int_{Q(x,1)} |u(y)|\,dy \\
&\leqslant c\|Du\|_{L^p(\mathbb{R}^n)} + \left(\int_{Q(x,1)} |u(y)|^p\,dy\right)^{\frac{1}{p}} \\
&\leqslant c\|u\|_{W^{1,p}(\mathbb{R}^n)}
\end{aligned}$$

for almost every $z \in \mathbb{R}^n$. Thus $\|u\|_{L^\infty(\mathbb{R}^n)} \leqslant c\|u\|_{W^{1,p}(\mathbb{R}^n)}$. ■

This implies that

$$\|\overline{u}\|_{C^{0,1-\frac{n}{p}}(\mathbb{R}^n)} \leqslant c\|u\|_{W^{1,p}(\mathbb{R}^n)}, \quad c = c(n,p),$$

where $\overline{u}$ is the Hölder continuous representative of u. Hence $W^{1,p}(\mathbb{R}^n)$ is continuously embedded in $C^{0,1-\frac{n}{p}}(\mathbb{R}^n)$, when $p > n$.

(3) The proof of Theorem 3.23, see (3.24), shows that if Ω is an open subset of $\mathbb{R}^n$ and $u \in W^{1,p}_{\mathrm{loc}}(\Omega)$, $p > n$, then there is $c = c(n,p)$ such that

$$|u(z) - u(y)| \leqslant c|z-y|^{1-\frac{n}{p}} \|Du\|_{L^p(Q(x,l))}$$

for every $z, y \in Q(x,l)$, $Q(x,l) \Subset \Omega$. This is a local version of Morrey's inequality.

THE MORAL: $W^{1,p}(\mathbb{R}^n) \subset C^{0,1-\frac{n}{p}}(\mathbb{R}^n)$, when $p > n$. More precisely, $W^{1,p}(\mathbb{R}^n)$ is continuously embedded in $C^{0,1-\frac{n}{p}}(\mathbb{R}^n)$, when $p > n$. This is the Sobolev embedding theorem for $p > n$.

Definition 3.26. A function $u : \mathbb{R}^n \to \mathbb{R}$ is differentiable at $x \in \mathbb{R}^n$ if there exists a linear mapping $L : \mathbb{R}^n \to \mathbb{R}$ such that

$$\lim_{y \to x} \frac{|u(y) - u(x) - L(x-y)|}{|x-y|} = 0. \tag{3.27}$$

If such a linear mapping L exists at x, it is unique and we denote $L = Du(x)$ and call $Du(x)$ the derivative of u at x. If the derivative Du exists, it is unique and satisfies

$$Du(y-x) = Du(x) \cdot (y-x)$$

for every $y \in \mathbb{R}^n$, where

$$Du(x) = \left(\frac{\partial u}{\partial x_1}(x), \dots, \frac{\partial u}{\partial x_n}(x) \right)$$

is the pointwise gradient of u at x.

Theorem 3.28. If $u \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n)$, $n < p \leqslant \infty$, then u is differentiable almost everywhere and its derivative equals its weak derivative almost everywhere.

THE MORAL: By the ACL characterization, see Theorem 2.36, we know that every function in $W^{1,p}$, $1 \leqslant p \leqslant \infty$ has classical partial derivatives almost everywhere. If $p > n$, then every function in $W^{1,p}$ is also differentiable almost everywhere.

Proof. Since $W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}^n) \subset W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n)$, we may assume $n < p < \infty$. By the Lebesgue differentiation theorem

$$\lim_{l \to 0} \fint_{Q(x,l)} |Du(z) - Du(x)|^p \, dz = 0$$

for almost every $x \in \mathbb{R}^n$. Let x be such a point and denote

$$v(y) = u(y) - u(x) - Du(x) \cdot (y-x),$$

where $y \in Q(x,l)$. Observe that $v \in W^{1,p}_{\text{loc}}(\mathbb{R}^n)$ with $n < p < \infty$. By (3.24) in the proof of Morrey's inequality, there is $c = c(n,p)$ such that

$$|v(y) - v(x)| \leqslant cl \left(\fint_{Q(x,l)} |Dv(z)|^p \, dz \right)^{\frac{1}{p}}$$

for almost every $y \in Q(x,l)$, where $l = |x - y|$. Since $v(x) = 0$ and $Dv(z) = Du(z) - Du(x)$, we obtain

$$\frac{|u(y) - u(x) - Du(x) \cdot (y - x)|}{|y - x|} \leqslant c \left(\fint_{Q(x,l)} |Du(z) - Du(x)|^p \, dz \right)^{\frac{1}{p}} \to 0$$

as $y \to x$. □

3.4 Lipschitz functions and $W^{1,\infty}$

Let $A \subset \mathbb{R}^n$ and $0 \leqslant L < \infty$. A function $f : A \to \mathbb{R}$ is called Lipschitz continuous with constant L, or an L-Lipschitz function, if

$$|f(x) - f(y)| \leqslant L|x - y|$$

for every $x, y \in \mathbb{R}^n$. Observe that a function is Lipschitz continuous if it is Hölder continuous with exponent one. Moreover, $C^{0,1}(A)$ is the space of all bounded Lipschitz continuous functions with the norm (3.21).

Examples 3.29:

(1) For every $y \in \mathbb{R}^n$ the function $x \mapsto |x - y|$ is Lipschitz continuous with constant one. Note that this function is not smooth.

(2) For every nonempty set $A \subset \mathbb{R}^n$ the function $x \mapsto \operatorname{dist}(x, A)$ is Lipschitz continuous with constant one. Note that this function is not smooth when $A \neq \mathbb{R}^n$ (exercise).

(3) By considering the zero extension of $u \in C^1_0(\Omega)$, we may assume that $u \in C^1_0(\mathbb{R}^n)$. Let $x, y \in \mathbb{R}^n$, $x \neq y$. By the mean value theorem, there exists z in the line-segment between x and y such that

$$|u(x) - u(y)| = |Du(z) \cdot (x - y)| \leqslant \|Du\|_{L^\infty(\mathbb{R}^n)} |x - y|.$$

This shows that u is L-Lipschitz with $L = \|Du\|_{L^\infty(\mathbb{R}^n)}$.

Example 3.30. Let $x \in \mathbb{R}^n$ and $r > 0$. Define

$$u(y) = \max\left\{0, 1 - \tfrac{1}{r} \operatorname{dist}(y, B(x,r))\right\},$$

for $y \in \mathbb{R}^n$. The function u is $\frac{1}{r}$-Lipschitz in $\mathbb{R}^n$, $u = 1$ in $\overline{B}(x,r)$, and $u = 0$ in $\mathbb{R}^n \setminus B(x, 2r)$. This kind of function is used as a cutoff to localize estimates.

The next theorem describes the relation between Lipschitz functions and Sobolev functions.

Theorem 3.31. A function $u \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ has a representative that is bounded and Lipschitz continuous if and only if $u \in W^{1,\infty}(\mathbb{R}^n)$.

THE MORAL: The Sobolev embedding theorem for $p > n$ shows that $W^{1,p}(\mathbb{R}^n) \subset C^{0,1-\frac{n}{p}}(\mathbb{R}^n)$. In the limiting case $p = \infty$ we have $W^{1,\infty}(\mathbb{R}^n) = C^{0,1}(\mathbb{R}^n)$. This is the Sobolev embedding theorem for $p = \infty$.

Proof. $\boxed{\Longleftarrow}$ Assume that $u \in W^{1,\infty}(\mathbb{R}^n)$. Then $u \in L^\infty(\mathbb{R}^n)$ and $u \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n)$ for every $p > n$ and thus by Remark 3.25 we may assume that u is a bounded continuous function. Moreover, we may assume that the support of u is compact. By Lemma 1.18 (3) and by Theorem 1.19, the standard mollification $u_\varepsilon \in C_0^\infty(\mathbb{R}^n)$ for every $\varepsilon > 0$, $u_\varepsilon \to u$ uniformly in $\mathbb{R}^n$ as $\varepsilon \to 0$ and

$$\|Du_\varepsilon\|_{L^\infty(\mathbb{R}^n)} \leqslant \|Du\|_{L^\infty(\mathbb{R}^n)}$$

for every $\varepsilon > 0$. Thus

$$\begin{aligned}|u_\varepsilon(x) - u_\varepsilon(y)| &= \left| \int_0^1 Du_\varepsilon(tx + (1-t)y) \cdot (x-y)\,dt \right| \\ &\leqslant \|Du_\varepsilon\|_{L^\infty(\mathbb{R}^n)}|x-y| \\ &\leqslant \|Du\|_{L^\infty(\mathbb{R}^n)}|x-y|\end{aligned}$$

for every $x, y \in \mathbb{R}^n$. By letting $\varepsilon \to 0$, we obtain

$$|u(x) - u(y)| \leqslant \|Du\|_{L^\infty(\mathbb{R}^n)}|x-y|$$

for every $x, y \in \mathbb{R}^n$.

$\boxed{\Longrightarrow}$ Assume that u is Lipschitz continuous. Then there exists L such that

$$|u(x) - u(y)| \leqslant L|x-y|$$

for every $x, y \in \mathbb{R}^n$. This implies that

$$|D_j^{-h}u(x)| = \left| \frac{u(x - he_j) - u(x)}{h} \right| \leqslant L$$

for every $x \in \mathbb{R}^n$ and $h \neq 0$. This means that

$$\|D_j^{-h}u\|_{L^\infty(\mathbb{R}^n)} \leqslant L$$

for every $h \neq 0$ and thus

$$\|D_j^{-h}u\|_{L^2(\Omega)} \leqslant \|D_j^{-h}u\|_{L^\infty(\mathbb{R}^n)}|\Omega|^{\frac{1}{2}} \leqslant L|\Omega|^{\frac{1}{2}},$$

where $\Omega \subset \mathbb{R}^n$ is bounded and open.

As in the proof of Theorem 2.32, by Theorem 2.13, there exists $g \in L^2(\Omega';\mathbb{R}^n)$ and a sequence $(h_i)_{i\in\mathbb{N}}$ converging to zero such that $D^{-h_i}u \to g$ weakly in $L^p(\Omega';\mathbb{R}^n)$ as $i \to \infty$. This implies

$$\begin{aligned}\int_\Omega u\frac{\partial\varphi}{\partial x_j}\,dx &= \int_\Omega u\left(\lim_{h_i\to 0} D_j^{h_i}\varphi\right)dx = \lim_{h_i\to 0}\int_{\Omega'} uD_j^{h_i}\varphi\,dx\\ &= -\lim_{h_i\to 0}\int_\Omega (D_j^{-h_i}u)\varphi\,dx = -\int_\Omega g_j\varphi\,dx\end{aligned}$$

for every $\varphi \in C_0^\infty(\Omega')$. It follows that $Du = g$ in the weak sense in Ω and thus $u \in W^{1,2}(\Omega)$.

Claim: $D_j u \in L^\infty(\Omega)$, $j = 1,\dots,n$,

Reason. Let $f_i = D_j^{-h_i}u$, $i = 1,2\dots$. Since $f_i \to D_j u$ weakly in $L^2(\Omega)$ as $i\to\infty$, by Mazur's lemma as in the proof of Theorem 2.20, there exists a sequence of convex combinations such that

$$\widetilde{f}_i = \sum_{k=i}^{m_i} a_{i,k}f_k \to D_j u$$

in $L^p(\Omega)$ as $i \to \infty$. Observe that

$$\|\widetilde{f}_i\|_{L^\infty(\Omega)} = \left\|\sum_{k=i}^{m_i} a_{i,k}f_k\right\|_{L^\infty(\Omega)} \leqslant \sum_{k=i}^{m_i} a_{i,k}\left\|D_j^{-h_k}u(x)\right\|_{L^\infty(\Omega)} \leqslant L.$$

Since there exists a subsequence that converges almost everywhere, we conclude that

$$\left|D_j u(x)\right| \leqslant L, \quad j = 1,\dots,n,$$

for almost every $x \in \Omega$. ■

This shows that $Du \in L^\infty(\Omega)$, with $\|Du\|_{L^\infty(\Omega)} \leqslant L$. As u is bounded, this implies $u \in W^{1,\infty}(\Omega)$ for all bounded subsets $\Omega \subset \mathbb{R}^n$. Since the norm does not depend on Ω, we conclude that $u \in W^{1,\infty}(\mathbb{R}^n)$. □

A direct combination of Theorem 3.31 and Theorem 3.28 gives a proof for Rademacher's theorem.

Corollary 3.32 (Rademacher). Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz continuous. Then f is differentiable almost everywhere.

W A R N I N G : For an open subset Ω of $\mathbb{R}^n$, Morrey's inequality and the characterization of Lipschitz continuous functions holds only locally, that is, $W^{1,p}(\Omega) \subset C_{\mathrm{loc}}^{0,1-\frac{n}{p}}(\Omega)$, when $p > n$ and $W^{1,\infty}(\Omega) \subset C_{\mathrm{loc}}^{0,1}(\Omega)$.

Example 3.33. Let

$$\Omega = \{x \in \mathbb{R}^2 : 1 < |x| < 2\} \setminus \{(x_1,0) \in \mathbb{R}^2 : 1 < x_1 < 2\} \subset \mathbb{R}^2.$$

Then there exists a function such that $u \in W^{1,\infty}(\Omega)$, but $u \notin C^{0,\alpha}(\Omega)$, for example, by defining $u(x) = \theta$, where $0 < \theta < 2\pi$ is the argument of x in polar coordinates. Then $u \in W^{1,\infty}(\Omega)$, but u is not Lipschitz continuous in Ω. However, it is locally Lipschitz continuous in Ω.

Instead of the gradient Du, we are often interested in $|Du|$, for which we have the following representation.

Lemma 3.34. Assume that $u : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz continuous. Then

$$|Du(x)| = \lim_{r\to 0} \sup_{y\in B(x,r)} \frac{|u(y)-u(x)|}{r} \tag{3.35}$$

for almost every $x \in \mathbb{R}^n$.

Proof. By Corollary 3.32 u is differentiable at almost every $x \in \mathbb{R}^n$. Let $x \in \mathbb{R}^n$ be such a point and let $r > 0$. Then

$$\sup_{y\in B(x,r)} \frac{|Du(x)\cdot(y-x)|}{r} \leqslant \sup_{y\in B(x,r)} \frac{|Du(x)||y-x|}{r} \leqslant |Du(x)|.$$

On the other hand, we choose $z = x + r\frac{Du(x)}{|Du(x)|}$ if $Du(x) \neq 0$, and $z = x$ if $Du(x) = 0$. Since $z \in \overline{B}(x,r)$, we have

$$\begin{aligned}\sup_{y\in B(x,r)} \frac{|Du(x)\cdot(y-x)|}{r} &= \sup_{y\in \overline{B}(x,r)} \frac{|Du(x)\cdot(y-x)|}{r}\\ &\geqslant \frac{|Du(x)\cdot(z-x)|}{r} = |Du(x)|.\end{aligned}$$

Combining the estimates above we obtain

$$|Du(x)| = \sup_{y\in B(x,r)} \frac{|Du(x)\cdot(y-x)|}{r}, \tag{3.36}$$

for every $r > 0$.

Since u is differentiable at x, it follows from (3.27) that

$$\lim_{r\to 0} \sup_{y\in B(x,r)} \frac{|u(y)-u(x)-Du(x)\cdot(y-x)|}{r} = 0. \tag{3.37}$$

Let $r > 0$ and $y \in B(x,r)$, and write

$$\begin{aligned} a(y,r) &= \frac{Du(x)\cdot(y-x)}{r},\\ b(y,r) &= \frac{u(y)-u(x)-Du(x)\cdot(y-x)}{r},\\ c(y,r) &= \frac{u(y)-u(x)}{r}.\end{aligned}$$

Then $c(y,r) = a(y,r) + b(y,r)$ and hence

$$\begin{aligned}|a(y,r)| - \sup_{z\in B(x,r)} |b(z,r)| &\leqslant |a(y,r)| - |b(y,r)| \leqslant |c(y,r)|\\ &\leqslant |a(y,r)| + |b(y,r)| \leqslant |a(y,r)| + \sup_{z\in B(x,r)} |b(z,r)|.\end{aligned}$$

By taking supremums over all $y \in B(x,r)$, we obtain

$$\begin{aligned}\sup_{y\in B(x,r)} |a(y,r)| - \sup_{y\in B(x,r)} |b(y,r)| &\leqslant \sup_{y\in B(x,r)} |c(y,r)| \\ &\leqslant \sup_{y\in B(x,r)} |a(y,r)| + \sup_{y\in B(x,r)} |b(y,r)|.\end{aligned}$$

The claim follows by taking $r \to 0$ and using (3.36) and (3.37). □

Remark 3.38. From (3.35) we see that if u is an L-Lipschitz function, then $|Du(x)| \leqslant L$ for almost every $x \in \mathbb{R}^n$.

The following locality property is a useful consequence of (3.35).

Lemma 3.39. Assume that $u : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz continuous and $t \in \mathbb{R}$. Then $Du = 0$ almost everywhere in the set $\{x \in \mathbb{R}^n : u(x) = t\}$.

Proof. Let $A = \{x \in \mathbb{R}^n : u(x) = t\}$ and let $x \in A$ be such that (3.35) holds. By the Lebesgue density theorem, we may assume that

$$\lim_{r\to 0} \frac{|A \cap B(x,r)|}{|B(x,r)|} = 1. \tag{3.40}$$

Let

$$d(r) = \sup_{y\in B(x,r)} d(y,A) + r^2 > 0$$

for $r > 0$. By (3.40) we have

$$\lim_{r\to 0} \frac{d(r)}{r} = 0.$$

Let $r > 0$ and $y \in B(x,r)$. There exists a point $z \in A \cap B(y,d(r))$. Since u is an L-Lipschitz function for some constant $L > 0$, we have

$$\frac{|u(y)-u(x)|}{r} = \frac{|u(y)-u(z)|}{r} \leqslant \frac{L|y-z|}{r} \leqslant \frac{Ld(r)}{r}.$$

This implies

$$|Du(x)| = \lim_{r\to 0} \sup_{y\in B(x,r)} \frac{|u(y)-u(x)|}{r} \leqslant \lim_{r\to 0} \frac{Ld(r)}{r} = 0,$$

and the proof is complete. □

3.5 Summary of the Sobolev embeddings

We summarize the results related to Sobolev embeddings below. Assume that Ω is an open subset of $\mathbb{R}^n$.

$\boxed{1 \leqslant p < n}$ $W^{1,p}(\mathbb{R}^n) \subset L^{p^*}(\mathbb{R}^n)$, $W^{1,p}_{\text{loc}}(\Omega) \subset L^{p^*}_{\text{loc}}(\Omega)$, $p^* = \frac{np}{n-p}$ (Theorem 3.3 and Theorem 3.13).

$\boxed{p=n}$ $W^{1,n}(\mathbb{R}^n)\subset \mathrm{BMO}(\mathbb{R}^n)$, $W^{1,n}_{\mathrm{loc}}(\Omega)\subset L^p_{\mathrm{loc}}(\Omega)$ for every p, with $1\leqslant p<\infty$ (Remark 3.16 (3)).

$\boxed{n<p<\infty}$ $W^{1,p}(\mathbb{R}^n)\subset C^{0,1-\frac{n}{p}}(\mathbb{R}^n)$, $W^{1,p}_{\mathrm{loc}}(\Omega)\subset C^{0,1-\frac{n}{p}}_{\mathrm{loc}}(\Omega)$ (Theorem 3.23).

$\boxed{p=\infty}$ $W^{1,\infty}(\mathbb{R}^n)=C^{0,1}(\mathbb{R}^n)$, $W^{1,\infty}_{\mathrm{loc}}(\Omega)=C^{0,1}_{\mathrm{loc}}(\Omega)$ (Theorem 3.31).

For Sobolev embeddings in higher order spaces $W^{k,m}(\mathbb{R}^n)$, we refer to [7, Section 5.6.3].

We close this section with a useful remark.

Remark 3.41. Let $1\leqslant p<\infty$. If $u\in W^{k,p}_{\mathrm{loc}}(\Omega)$ for every $k=1,2,\dots$, then $u\in C^\infty(\Omega)$.

Reason. If $p>n$, then $W^{1,p}_{\mathrm{loc}}(\Omega)\subset C^{0,\alpha}_{\mathrm{loc}}(\Omega)$ and thus $W^{k,p}_{\mathrm{loc}}(\Omega)\subset C^{k-1,\alpha}_{\mathrm{loc}}(\Omega)$, $k=2,3,\dots$. It follows that $u\in C^\infty(\Omega)$. Then assume that $1\leqslant p\leqslant n$. Since $W^{k,q}_{\mathrm{loc}}(\Omega)\subset W^{k,p}_{\mathrm{loc}}(\Omega)$, $1\leqslant p<q$, we may assume that $1\leqslant p<\frac{n}{2}$. Let $k=2,3,\dots$ be such that $kp<n<(k+1)p$. Then

$$W^{1,p}_{\mathrm{loc}}(\Omega)\subset L^{\frac{np}{n-p}}_{\mathrm{loc}}(\Omega)$$

and recursively

$$W^{k+1,p}_{\mathrm{loc}}(\Omega)\subset W^{k,\frac{np}{n-p}}_{\mathrm{loc}}(\Omega)\subset W^{k-1,\frac{np}{n-2p}}_{\mathrm{loc}}(\Omega)\dots\subset W^{1,\frac{np}{n-kp}}_{\mathrm{loc}}(\Omega)\subset C^{0,\alpha}_{\mathrm{loc}}(\Omega),$$

since $\frac{np}{n-kp}>n$. Again it follows that $u\in C^\infty(\Omega)$. ■

3.6 Compactness

Let X and Y be Banach spaces. Recall that a bounded linear operator $L:X\to Y$ is compact, if every bounded sequence (x_i), $x_i\in X$, $i=1,2,\dots$, has a subsequence (x_{i_k}) such that the sequence (Tx_{i_k}) converges in Y.

Definition 3.42. Let $1\leqslant p\leqslant\infty$. An open set $\Omega\subset\mathbb{R}^n$ is called an extension domain, if there exists a linear operator $E:W^{1,p}(\Omega)\to W^{1,p}(\mathbb{R}^n)$ such that $Eu|_\Omega=u$ for every $u\in W^{1,p}(\Omega)$ and there exists a constant $c=c(n,p,\Omega)$ such that

$$\|Eu\|_{W^{1,p}(\mathbb{R}^n)}\leqslant c\|u\|_{W^{1,p}(\Omega)}$$

for every $u\in W^{1,p}(\Omega)$.

T H E M O R A L : If $\Omega\subset\mathbb{R}^n$ is an extension domain, every function in $W^{1,p}(\Omega)$ can be extended to a function in $W^{1,p}(\mathbb{R}^n)$ with uniform bounds for the norms.

Example 2.39 shows that an upper half space is an extension domain. It can be shown that open sets with Lipschitz boundary are extension domains, see [8, Section 4.4] and [14, Section 13.1]. Observe that every open set $\Omega\subset\mathbb{R}^n$ is an extension domain for $W^{1,p}_0(\Omega)$, since we may consider the zero extension to $\mathbb{R}^n\setminus\Omega$.

Next we show that the Gagliardo-Nirenberg-Sobolev inequality in Theorem 3.3, see also Corollary 3.5, does not only hold in the entire space but also in extension domains.

Theorem 3.43. Let $1 \leqslant p < n$, $p \leqslant q \leqslant p^*$ and assume that $\Omega \subset \mathbb{R}^n$ is an extension domain. Then there exists a constant $c = c(n,p,\Omega)$ such that

$$\|u\|_{L^q(\Omega)} \leqslant c\|u\|_{W^{1,p}(\Omega)}$$

for every $u \in W^{1,p}(\Omega)$.

THE MORAL: If $\Omega \subset \mathbb{R}^n$ is an extension domain, the embedding $L : W^{1,p}(\Omega) \to L^q(\Omega)$, $Lu = u$, is a bounded linear operator.

Proof. By Corollary 3.5, we have

$$\begin{aligned} \|u\|_{L^q(\Omega)} = \|Eu\|_{L^q(\Omega)} &\leqslant \|Eu\|_{L^q(\mathbb{R}^n)} \\ &\leqslant \|Eu\|_{W^{1,p}(\mathbb{R}^n)} \leqslant c\|u\|_{W^{1,p}(\Omega)} \end{aligned}$$

for every $u \in W^{1,p}(\Omega)$. □

Extension domains have certain compactness results that are useful, for example, in the existence theory for PDEs. For the proof, see [8, Section 4.6] and [14, Theorem 12.18]. Moreover, if $u_i \in W_0^{1,p}(\Omega)$, $i = 1,2\dots$, then $u \in W_0^{1,p}(\Omega)$ for every open set $\Omega \subset \mathbb{R}^n$.

Theorem 3.44 (Rellich–Kondrachov). Let $1 < p < n$ and assume that $\Omega \subset \mathbb{R}^n$ is a bounded extension domain. Assume that (u_i) is a bounded sequence of functions $u_i \in W^{1,p}(\Omega)$, $i = 1,2,\dots$. Then there exists a subsequence (u_{i_k}) and $u \in W^{1,p}(\Omega)$ such that $u_{i_k} \to u$ in $L^q(\Omega)$ as $k \to \infty$ for every $1 \leqslant q < p^*$.

THE MORAL: The embedding $L : W^{1,p}(\mathbb{R}^n) \to L^q(\mathbb{R}^n)$, $Lu = u$, is a compact operator.

Proof. (1) Let Ω' be a bounded open set such that $\Omega \Subset \Omega'$. Since Ω is an extension domain, we may extend every $u_i \in W^{1,p}(\Omega)$ to $\overline{u}_i \in W^{1,p}(\mathbb{R}^n)$, $i = 1,2,\dots$, with

$$\sup_{i\in\mathbb{N}} \|\overline{u}_i\|_{W^{1,p}(\mathbb{R}^n)} \leqslant c \sup_{i\in\mathbb{N}} \|u_i\|_{W^{1,p}(\Omega)} < \infty,$$

where $c = c(n,p,\Omega)$. Let $\eta \in C_0^\infty(\Omega')$ be a cutoff function with $0 \leqslant \eta \leqslant 1$ and $\eta = 1$ on

Ω. Then $\operatorname{supp}(\eta\overline{u}_i) \subset \Omega'$ and $\eta\overline{u}_i \in W_0^{1,p}(\Omega') \subset W^{1,p}(\mathbb{R}^n)$ with

$$\begin{aligned}
\|\eta\overline{u}_i\|_{W^{1,p}(\mathbb{R}^n)} &= \left(\|\eta\overline{u}_i\|^p_{L^p(\mathbb{R}^n)} + \sum_{j=1}^n \left\|D_j(\eta\overline{u}_i)\right\|^p_{L^p(\mathbb{R}^n)} \right)^{\frac{1}{p}} \\
&\leqslant \|\eta\overline{u}_i\|_{L^p(\mathbb{R}^n)} + \sum_{j=1}^n \left\|D_j(\eta\overline{u}_i)\right\|_{L^p(\mathbb{R}^n)} \\
&\leqslant \|\overline{u}_i\|_{L^p(\mathbb{R}^n)} + \sum_{j=1}^n \left\|\overline{u}_i D_j\eta + \eta D_j\overline{u}_i\right\|_{L^p(\mathbb{R}^n)} \\
&\leqslant \|\overline{u}_i\|_{L^p(\mathbb{R}^n)} + \sum_{j=1}^n \left\|\overline{u}_i D_j\eta\right\|_{L^p(\mathbb{R}^n)} + \sum_{j=1}^n \left\|\eta D_j\overline{u}_i\right\|_{L^p(\mathbb{R}^n)} \\
&\leqslant \|\overline{u}_i\|_{L^p(\mathbb{R}^n)} + \max_{j=1,\dots,n} \|D_j\eta\|_{L^\infty(\mathbb{R}^n)} \sum_{j=1}^n \|\overline{u}_i\|_{L^p(\mathbb{R}^n)} + \sum_{j=1}^n \left\|D_j\overline{u}_i\right\|_{L^p(\mathbb{R}^n)} \\
&\leqslant c\|\overline{u}_i\|_{W^{1,p}(\mathbb{R}^n)}, \quad i = 1,2,\dots,
\end{aligned}$$

where $c = c(n,\Omega')$. By replacing $\overline{u}_i$ with $\eta\overline{u}_i$, $i = 1,2,\dots$, we may conclude that $\overline{u}_i \in W^{1,p}(\mathbb{R}^n)$, $\operatorname{supp}\overline{u}_i \subset \Omega'$, $i = 1,2,\dots$, with

$$M = \sup_{i\in\mathbb{N}} \|\overline{u}_i\|_{W^{1,p}(\mathbb{R}^n)} \leqslant c \sup_{i\in\mathbb{N}} \|u_i\|_{W^{1,p}(\Omega)} < \infty,$$

where $c = c(n,p,\Omega,\Omega')$.

(2) Let $(\overline{u}_i)_\varepsilon = \overline{u}_i * \varphi_\varepsilon$, $\varepsilon > 0$, be the standard mollification of $\overline{u}_i$, $i = 1,2,\dots$.

Claim: There exists a constant $c = c(n,p,\Omega,\Omega')$ such that

$$\sup_{i\in\mathbb{N}} \|(\overline{u}_i)_\varepsilon - \overline{u}_i\|_{L^p(\mathbb{R}^n)} \leqslant c\varepsilon.$$

Reason. First assume that $\overline{u}_i \in C^\infty(\mathbb{R}^n)$, $i = 1,2,\dots$. Then

$$\begin{aligned}
|(\overline{u}_i)_\varepsilon(x) - \overline{u}_i(x)| &= \left| \int_{B(0,\varepsilon)} \phi_\varepsilon(y)\overline{u}_i(x-y)\,dy - \overline{u}_i(x) \int_{B(0,\varepsilon)} \phi_\varepsilon(y)\,dy \right| \\
&= \left| \int_{B(0,\varepsilon)} \phi_\varepsilon(y)(\overline{u}_i(x-y) - \overline{u}_i(x)\,dy \right| \\
&= \left| \int_{B(0,1)} \phi(z)(\overline{u}_i(x-\varepsilon z) - \overline{u}_i(x))\,dz \right| \\
&\leqslant \int_{B(0,1)} \phi(z)|\overline{u}_i(x-\varepsilon z) - \overline{u}_i(x)|\,dz \\
&= \int_{B(0,1)} \phi(z)|\overline{u}_i(x-\varepsilon z) - \overline{u}_i(x)|\,dz
\end{aligned}$$

Since

$$\begin{aligned}
|\overline{u}_i(x-\varepsilon z) - \overline{u}_i(x)| &= \left| \int_0^1 \frac{\partial}{\partial t}(\overline{u}_i(x-\varepsilon tz))\,dt \right| \leqslant \int_0^1 \left| \frac{\partial}{\partial t}(\overline{u}_i(x-\varepsilon tz)) \right| dt \\
&= \int_0^1 |D\overline{u}_i(x-\varepsilon tz)\cdot\varepsilon z|\,dt \leqslant \int_0^1 |D\overline{u}_i(x-\varepsilon tz)||\varepsilon z|\,dt \\
&\leqslant \varepsilon \int_0^1 |D\overline{u}_i(x-\varepsilon tz)|\,dt
\end{aligned}$$

for every $z \in B(0,1)$, by applying Hölder's inequality twice, we obtain

$$\begin{aligned}
|(\overline{u}_i)_\varepsilon(x) - \overline{u}_i(x)|^p &\leqslant \left(\int_{B(0,1)} \phi(z)|\overline{u}_i(x-\varepsilon z) - \overline{u}_i(x)|\,dz\right)^p \\
&\leqslant \varepsilon^p \left(\int_{B(0,1)} \phi(z)\left(\int_0^1 |D\overline{u}_i(x-\varepsilon tz)|\,dt\right)dz\right)^p \\
&= \varepsilon^p \left(\int_{B(0,1)} \phi(z)^{\frac{1}{p'}}\left(\phi(z)^{\frac{1}{p}}\int_0^1 |D\overline{u}_i(x-\varepsilon tz)|\,dt\right)dz\right)^p \\
&\leqslant \varepsilon^p \left(\int_{B(0,1)} \phi(z)\,dz\right)^{\frac{1}{p'}} \int_{B(0,1)} \phi(z)\left(\int_0^1 |D\overline{u}_i(x-\varepsilon tz)|\,dt\right)^p dz \\
&\leqslant \varepsilon^p \int_{B(0,1)} \phi(z)\left(\int_0^1 |D\overline{u}_i(x-\varepsilon tz)|^p\,dt\right)dz.
\end{aligned}$$

By Fubini's theorem we have

$$\begin{aligned}
\|(\overline{u}_i)_\varepsilon - \overline{u}_i\|_{L^p(\mathbb{R}^n)}^p &= \int_{\mathbb{R}^n} |(\overline{u}_i)_\varepsilon(x) - \overline{u}_i(x)|^p\,dx \\
&\leqslant \varepsilon^p \int_{\mathbb{R}^n}\left(\int_{B(0,1)} \phi(z)\left(\int_0^1 |D\overline{u}_i(x-\varepsilon tz)|^p\,dt\right)dz\right)dx \\
&= \varepsilon^p \int_{B(0,1)} \phi(z)\int_0^1\left(\int_{\mathbb{R}^n} |D\overline{u}_i(x-\varepsilon tz)|^p\,dx\right)dt\,dz \\
&= \varepsilon^p \int_{B(0,1)} \phi(z)\int_0^1\left(\int_{\mathbb{R}^n} |D\overline{u}_i(x)|^p\,dx\right)dt\,dz \\
&= \varepsilon^p \|D\overline{u}_i\|_{L^p(\mathbb{R}^n)}.
\end{aligned}$$

■

The general case $u_i \in W^{1,p}(\Omega)$ follows by approximation (exercise).

(3) **Claim:** For every $\varepsilon > 0$ the sequence $((\overline{u}_i)_\varepsilon)$ is bounded and equicontinuous in $\mathbb{R}^n$.

Reason. By Hölder's inequality, we have

$$\begin{aligned}
|(\overline{u}_i)_\varepsilon(x)| &= \left|\int_{B(x,\varepsilon)} \overline{u}_i(y)\phi_\varepsilon(x-y)\,dy\right| \leqslant \int_{B(x,\varepsilon)} \phi_\varepsilon(x-y)|\overline{u}_i(y)|\,dy \\
&\leqslant \left(\int_{B(x,\varepsilon)} \phi_\varepsilon(x-y)^{p'}\,dy\right)^{\frac{1}{p'}}\left(\int_{B(x,\varepsilon)} |\overline{u}_i(y)|^p\,dy\right)^{\frac{1}{p}},
\end{aligned}$$

where

$$\begin{aligned}
\left(\int_{B(x,\varepsilon)} \phi_\varepsilon(x-y)^{p'}\,dy\right)^{\frac{1}{p'}} &= \left(\int_{B(0,\varepsilon)} \phi_\varepsilon(y)^{p'}\,dy\right)^{\frac{1}{p'}} = \left(\int_{B(0,\varepsilon)} \left(\frac{1}{\varepsilon^n}\phi\left(\frac{y}{\varepsilon}\right)\right)^{p'} dy\right)^{\frac{1}{p'}} \\
&= \frac{1}{\varepsilon^n}\left(\int_{B(0,\varepsilon)} \phi\left(\frac{y}{\varepsilon}\right)^{p'} dy\right)^{\frac{1}{p'}} \leqslant \frac{1}{\varepsilon^n}\|\phi\|_{L^\infty(\mathbb{R}^n)}|B(0,\varepsilon)|^{\frac{1}{p'}} \\
&= \frac{1}{\varepsilon^n}\|\phi\|_{L^\infty(\mathbb{R}^n)}(\Omega_n\varepsilon^n)^{\frac{1}{p'}} = \varepsilon^{-\frac{n}{p}}\Omega_n^{\frac{1}{p'}}\|\phi\|_{L^\infty(\mathbb{R}^n)}.
\end{aligned}$$

This shows that

$$|(\overline{u}_i)_\varepsilon(x)| \leqslant \varepsilon^{-\frac{n}{p}} \Omega_n^{\frac{1}{p'}} \|\phi\|_{L^\infty(\mathbb{R}^n)} \|\overline{u}_i\|_{L^p(\mathbb{R}^n)} \leqslant c\varepsilon^{-\frac{n}{p}}$$

for every $x \in \mathbb{R}^n$ and for every $i = 1,2,\dots$. Since $(\overline{u}_i)$ is a bounded sequence in $W^{1,p}(\mathbb{R}^n)$, we conclude that there exists a constant $c = c(n,p,M)$ such that

$$\sup_{i\in\mathbb{N}} \|(\overline{u}_i)_\varepsilon\|_{L^\infty(\mathbb{R}^n)} \leqslant c\varepsilon^{-\frac{n}{p}}.$$

Similarly, for the partial derivatives, we have

$$\begin{aligned}|D_j(\overline{u}_i)_\varepsilon(x)| &= |D_j(\overline{u}_i * \phi_\varepsilon)(x)| = |(\overline{u}_i * D_j\phi_\varepsilon)(x)|\\ &= \left|\int_{B(x,\varepsilon)} \overline{u}_i(y) D_j\phi_\varepsilon(x-y)\,dy\right| \leqslant \int_{B(x,\varepsilon)} |D_j\phi_\varepsilon(x-y)||\overline{u}_i(y)|\,dy\\ &\leqslant \left(\int_{B(x,\varepsilon)} |D_j\phi_\varepsilon(x-y)|^{p'}\,dy\right)^{\frac{1}{p'}} \left(\int_{B(x,\varepsilon)} |\overline{u}_i(y)|^p\,dy\right)^{\frac{1}{p}}\end{aligned}$$

for every $j = 1,2,\dots,n$. We note that

$$D_j\phi_\varepsilon(y) = D_j\left(\frac{1}{\varepsilon^n}\phi\left(\frac{y}{\varepsilon}\right)\right) = \frac{1}{\varepsilon^n} D_j\left(\phi\left(\frac{y}{\varepsilon}\right)\right) = \frac{1}{\varepsilon^{n+1}} D_j\phi\left(\frac{y}{\varepsilon}\right), \quad j = 1,2,\dots,n.$$

Thus we have

$$\begin{aligned}\left(\int_{B(x,\varepsilon)} |D_j\phi_\varepsilon(x-y)|^{p'}\,dy\right)^{\frac{1}{p'}} &= \left(\int_{B(0,\varepsilon)} |D_j\phi_\varepsilon(y)|^{p'}\,dy\right)^{\frac{1}{p'}}\\ &= \frac{1}{\varepsilon^{n+1}} \left(\int_{B(0,\varepsilon)} \left|D_j\phi\left(\frac{y}{\varepsilon}\right)\right|^{p'} dy\right)^{\frac{1}{p'}} \leqslant \frac{1}{\varepsilon^{n+1}} \|D_j\phi\|_{L^\infty(\mathbb{R}^n)} |B(0,\varepsilon)|^{\frac{1}{p'}}\\ &= \frac{1}{\varepsilon^{n+1}} \|D_j\phi\|_{L^\infty(\mathbb{R}^n)} (\Omega_n\varepsilon^n)^{\frac{1}{p'}} = \varepsilon^{-1-\frac{n}{p}} \Omega_n^{\frac{1}{p'}} \|D\phi\|_{L^\infty(\mathbb{R}^n)}\end{aligned}$$

for every $j = 1,2,\dots,n$. This shows that

$$|D(\overline{u}_i)_\varepsilon(x)| \leqslant \varepsilon^{-1-\frac{n}{p}} \Omega_n^{\frac{1}{p'}} \|D\phi\|_{L^\infty(\mathbb{R}^n)} \|\overline{u}_i\|_{L^p(\mathbb{R}^n)}$$

for every $x \in \mathbb{R}^n$ and for every $i = 1,2,\dots$. Since $(\overline{u}_i)$ is a bounded sequence in $W^{1,p}(\mathbb{R}^n)$, we conclude that there exists a constant $c = c(n,p,M)$ such that

$$\sup_{i\in\mathbb{N}} \|D(\overline{u}_i)_\varepsilon\|_{L^\infty(\mathbb{R}^n)} \leqslant c\varepsilon^{-1-\frac{n}{p}}$$

Since

$$\begin{aligned}|(\overline{u}_i)_\varepsilon(x) - (\overline{u}_i)_\varepsilon(y)| &\leqslant \|D(\overline{u}_i)_\varepsilon\|_{L^\infty(\mathbb{R}^n)} |x-y|\\ &\leqslant c\varepsilon^{-1-\frac{n}{p}} |x-y|, \quad i = 1,2,\dots,\end{aligned}$$

for every $x, y \in \mathbb{R}^n$, we conclude that, for every $\varepsilon > 0$, the sequence $((\overline{u}_i)_\varepsilon)$ is uniformly bounded sequence of Lipschitz continuous functions with uniformly bounded Lipschitz constants. In particular, it is bounded and equicontinuous in $\mathbb{R}^n$. ∎

(4) **Claim:** For every $\delta > 0$ there exists a subsequence (u_{i_k}) of (u_i) such that

$$\limsup_{k,l\to\infty} \|u_{i_k} - u_{i_l}\|_{L^p(\Omega)} \leqslant \delta.$$

In other words, (u_{i_k}) is a Cauchy sequence in $L^p(\Omega)$.

Reason. By step (2) there exists a constant $c = c(n, p, \Omega, \Omega')$ such that

$$\sup_{i\in\mathbb{N}} \|(\overline{u}_i)_\varepsilon - \overline{u}_i\|_{L^p(\mathbb{R}^n)} \leqslant \frac{\delta}{3}.$$

By step (3) and the Arzela-Ascoli theorem, there exists a subsequence $((\overline{u}_{i_k})_\varepsilon)$ of $((\overline{u}_i)_\varepsilon)$ which converges uniformly in $\mathbb{R}^n$. Thus there exists $k_\delta \in \mathbb{N}$ such that

$$\begin{aligned}
\|(\overline{u}_{i_k})_\varepsilon - (\overline{u}_{i_l})_\varepsilon\|_{L^p(\mathbb{R}^n)} &= \|(\overline{u}_{i_k})_\varepsilon - (\overline{u}_{i_l})_\varepsilon\|_{L^p(\Omega'_\varepsilon)} \\
&\leqslant \|(\overline{u}_{i_k})_\varepsilon - (\overline{u}_{i_l})_\varepsilon\|_{L^\infty(\Omega'_\varepsilon)} |\Omega'_\varepsilon| \\
&\leqslant \|(\overline{u}_{i_k})_\varepsilon - (\overline{u}_{i_l})_\varepsilon\|_{L^\infty(\mathbb{R}^n)} |\Omega'_\varepsilon| \leqslant \frac{\delta}{3}
\end{aligned}$$

for every $k, l \geqslant k_\delta$. It follows that

$$\begin{aligned}
\|u_{i_k} - u_{i_l}\|_{L^p(\Omega)} &\leqslant \|\overline{u}_{i_k} - \overline{u}_{i_l}\|_{L^p(\mathbb{R}^n)} \\
&\leqslant \|\overline{u}_{i_k} - (\overline{u}_{i_k})_\varepsilon\|_{L^p(\mathbb{R}^n)} + \|(\overline{u}_{i_k})_\varepsilon - (\overline{u}_{i_l})_\varepsilon\|_{L^p(\mathbb{R}^n)} + \|(\overline{u}_{i_l})_\varepsilon - \overline{u}_{i_l}\|_{L^p(\mathbb{R}^n)} \\
&\leqslant \|\overline{u}_{i_k} - (\overline{u}_{i_k})_\varepsilon\|_{L^p(\mathbb{R}^n)} + \|(\overline{u}_{i_l})_\varepsilon - \overline{u}_{i_l}\|_{L^p(\mathbb{R}^n)} + \frac{\delta}{3}
\end{aligned}$$

for every $k, l \geqslant k_\delta$. Since

$$\|\overline{u}_{i_k} - (\overline{u}_{i_k})_\varepsilon\|_{L^p(\mathbb{R}^n)} \to 0 \quad \text{and} \quad \|(\overline{u}_{i_l})_\varepsilon - \overline{u}_{i_l}\|_{L^p(\mathbb{R}^n)} \to 0$$

as $k, l \to \infty$, we conclude that there exists $k'_\delta \in \mathbb{N}$ such that

$$\|u_{i_k} - u_{i_l}\|_{L^p(\Omega)} \leqslant \frac{\delta}{3} + \frac{\delta}{3} + \frac{\delta}{3} = \delta$$

for every $k, l \geqslant k'_\delta$. ■

(5) Since $L^p(\Omega)$ is complete, the sequence (u_{i_k}) converges in $L^p(\Omega)$ and thus there exists $u \in L^p(\Omega)$ such that

$$\|u_{i_k} - u\|_{L^p(\Omega)} \to 0$$

as $k \to \infty$. Theorem 2.24 implies that $u \in W^{1,p}(\Omega)$ with

$$\|u\|_{W^{1,p}(\Omega)} \leqslant \liminf_{k\to\infty} \|u_{i_k}\|_{W^{1,p}(\Omega)} \leqslant c \sup_{i\in\mathbb{N}} \|u_i\|_{W^{1,p}(\Omega)} < \infty.$$

Let $1 \leqslant q < p^*$ and $0 < \theta < 1$ such that $\frac{1}{q} = \frac{\theta}{p} + \frac{1-\theta}{p^*}$. Then

$$\|u_{i_k} - u\|_{L^q(\Omega)} \leqslant \|u_{i_k} - u\|^\theta_{L^p(\Omega)} \|u_{i_k} - u\|^{1-\theta}_{L^{p^*}(\Omega)}.$$

By Theorem 3.43 there exists a constant $c = c(n,p,\Omega)$ such that

$$\begin{aligned}\|u_{i_k} - u\|_{L^{p^*}(\Omega)} &\leqslant c\|u_{i_k} - u\|_{W^{1,p}(\Omega)}\\ &\leqslant c(\|u_{i_k}\|_{W^{1,p}(\Omega)} + \|u\|_{W^{1,p}(\Omega)})\\ &\leqslant c\sup_{i\in\mathbb{N}}\|u_i\|_{W^{1,p}(\Omega)} < \infty\end{aligned}$$

for every $k = 1,2,\dots$. It follows that

$$\|u_{i_k} - u\|_{L^q(\Omega)} \to 0$$

as $k \to \infty$. □

Remark 3.45. Let $1 < p < n$ and assume that $\Omega \subset \mathbb{R}^n$ is a bounded extension domain. Assume that (u_i) is a bounded sequence of functions $u_i \in W^{1,p}(\Omega)$, $i = 1,2,\dots$. By Theorem 3.44 there exists a subsequence (u_{i_k}) and $u \in W^{1,p}(\Omega)$ such that $u_{i_k} \to u$ in $L^q(\Omega)$ as $k \to \infty$ for every $1 \leqslant q < p^*$. Since $u_{i_k} \to u$ in $L^q(\Omega)$ as $k \to \infty$ implies that there exists a further subsequence denoted again by (u_{i_k}) such that $u_{i_k} \to u$ almost everywhere in Ω as $k \to \infty$.

Example 3.46. Let $\Omega = B(0,1) \subset \mathbb{R}^n$ and $1 \leqslant p < n$. Consider $u_i : \Omega \to \mathbb{R}$,

$$u_i(x) = \begin{cases} i^{\frac{n-p}{p}}(1 - i|x|), & |x| < \frac{1}{i},\\ 0, & |x| \geqslant \frac{1}{i}.\end{cases}$$

Then (u_i) is a bounded sequence in $W^{1,p}(\Omega)$, but it does not have any converging subsequence in $L^{p^*}(\Omega)$ (exercise).

THE MORAL: The Rellich-Kondrachev theorem does not hold with $q = p^*$.

As an application of the Rellich-Kondrachev theorem, we obtain a general version of the Sobolev-Poincaré inequality.

Theorem 3.47. Let $1 < p < n$ and assume that $\Omega \subset \mathbb{R}^n$ is a bounded and connected extension domain. Then there exist a constant $c = c(n,p,\Omega)$ such that

$$\left(\int_\Omega |u - u_\Omega|^{p^*}\,dx\right)^{\frac{1}{p^*}} \leqslant c\left(\int_\Omega |Du|^p\,dx\right)^{\frac{1}{p}}$$

where $p^* = \frac{pn}{n-p}$.

Proof. Let $v = u - u_\Omega$. Then $v \in W^{1,p}(\Omega)$, $Dv = Du$ and $\int_\Omega v\,dx = 0$. By Theorem 3.43 there exists a constant $c = c(n,p,\Omega)$ such that

$$\|u - u_\Omega\|_{L^{p^*}(\Omega)} = \|v\|_{L^{p^*}(\Omega)} \leqslant c\|v\|_{W^{1,p}(\Omega)} \leqslant c(\|v\|_{L^p(\Omega)} + \|Du\|_{L^p(\Omega)}).$$

It is enough to prove that there exists a constant c such that

$$\|w\|_{L^p(\Omega)} \leqslant c\|Dw\|_{L^p(\Omega)}$$

for every function $w \in W^{1,p}(\Omega)$ with $\int_\Omega w\,dx = 0$. Moreover, we may assume that $\|w\|_{L^p(\Omega)} > 0$, since otherwise the claim is clear.

For a contradiction, assume that the inequality above is not true. Then, for every $i = 1,2,\dots$, there exists a function $w_i \in W^{1,p}(\Omega)$ such that $\|w_i\|_{L^p(\Omega)} > 0$, $\int_\Omega w_i\,dx = 0$ and

$$\|w_i\|_{L^p(\Omega)} \geqslant i\|Dw_i\|_{L^p(\Omega)}.$$

We may replace w_i with $\frac{w_i}{\|w_i\|_{L^p(\Omega)}}$ and assume that $\|w_i\|_{L^p(\Omega)} = 1$. It follows that

$$\|Dw_i\|_{L^p(\Omega)} \leqslant \frac{1}{i}, \quad i = 1,2,\dots,$$

and thus (w_i) is a bounded sequence in $W^{1,p}(\Omega)$. By Theorem 3.44 there exists a subsequence (w_{i_k}) and $w \in W^{1,p}(\Omega)$ such that $w_{i_k} \to w$ in $L^p(\Omega)$ as $k \to \infty$. Moreover, we have

$$\|Dw\|_{L^p(\Omega)} = \lim_{k\to\infty} \|Dw_{i_k}\|_{L^p(\Omega)} \leqslant \lim_{k\to\infty} \frac{1}{i_k} = 0.$$

Thus $Dw = 0$ almost everywhere in Ω and $w = w_\Omega$ almost everywhere in Ω. On the other hand, we have

$$\left|\int_\Omega w_{i_k}\,dx - \int_\Omega w\,dx\right| \leqslant \int_\Omega |w_{i_k} - w|\,dx \xrightarrow{k\to\infty} 0,$$

which implies that

$$\int_\Omega w\,dx = \lim_{k\to\infty} \int_\Omega w_{i_k}\,dx = 0.$$

It follows that $w = w_\Omega = 0$ almost everywhere in Ω. This is a contradiction, since

$$\|w\|_{L^p(\Omega)} = \lim_{k\to\infty} \|w_i\|_{L^p(\Omega)} = 1.$$

□

4

Pointwise behaviour of Sobolev functions

In this chapter we study fine properties of Sobolev functions. By definition, Sobolev functions are defined only up to Lebesgue measure zero and thus it is not always clear how to use their pointwise properties to give meaning, for example, to boundary values.

4.1 Sobolev capacity

Capacities are needed to understand pointwise behavior of Sobolev functions. They also play an important role in studies of solutions of partial differential equations.

Definition 4.1. For $1 < p < \infty$, the Sobolev p-capacity of a set $E \subset \mathbb{R}^n$ is defined by

$$\begin{aligned}\operatorname{cap}_p(E) &= \inf_{u \in \mathscr{A}(E)} \|u\|^p_{W^{1,p}(\mathbb{R}^n)} \\ &= \inf_{u \in \mathscr{A}(E)} \left(\|u\|^p_{L^p(\mathbb{R}^n)} + \|Du\|^p_{L^p(\mathbb{R}^n)} \right) \\ &= \inf_{u \in \mathscr{A}(E)} \int_{\mathbb{R}^n} \left(|u|^p + |Du|^p \right) dx,\end{aligned}$$

where

$$\mathscr{A}(E) = \left\{ u \in W^{1,p}(\mathbb{R}^n) : u \geqslant 1 \text{ almost everywhere in a neighbourhood of } E \right\}.$$

If $\mathscr{A}(E) = \emptyset$, we set $\operatorname{cap}_p(E) = \infty$. Functions in $\mathscr{A}(E)$ are called admissible functions for E.

THE MORAL: Capacity measures the size of exceptional sets for Sobolev functions. Lebesgue measure is the natural measure for functions in $L^p(\mathbb{R}^n)$ and the Sobolev p-capacity is the natural outer measure for functions in $W^{1,p}(\mathbb{R}^n)$.

Remark 4.2. In the definition of capacity we can restrict ourselves to the admissible functions u for which $0 \leqslant u \leqslant 1$. Thus

$$\operatorname{cap}_p(E) = \inf_{u \in \mathscr{A}'(E)} \|u\|^p_{W^{1,p}(\mathbb{R}^n)},$$

where

$$\mathscr{A}'(E) = \big\{u \in W^{1,p}(\mathbb{R}^n) : 0 \leqslant u \leqslant 1, \\ u = 1 \text{ almost everywhere in a neighbourhood of } E\big\}.$$

Reason. (1) Since $\mathscr{A}'(E) \subset \mathscr{A}(E)$, we have

$$\operatorname{cap}_p(E) \leqslant \inf_{u \in \mathscr{A}'(E)} \|u\|^p_{W^{1,p}(\mathbb{R}^n)}.$$

(2) For the reverse inequality, let $\varepsilon > 0$ and let $u \in \mathscr{A}(E)$ such that

$$\|u\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant \operatorname{cap}_p(E) + \varepsilon.$$

Then $v = \max\{0, \min\{u, 1\}\} \in \mathscr{A}'(E)$, $|v| \leqslant |u|$ and by Remark 2.4 we have $|Dv| \leqslant |Du|$ almost everywhere. Thus

$$\inf_{u \in \mathscr{A}'(E)} \|u\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant \|v\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant \|u\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant \operatorname{cap}_p(E) + \varepsilon$$

and by letting $\varepsilon \to 0$ we obtain

$$\inf_{u \in \mathscr{A}'(E)} \|u\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant \operatorname{cap}_p(E).$$

Remarks 4.3:

(1) There are several alternative definitions for capacity and, in general, it does not matter which one we choose. For example, when $1 < p < n$, we may consider the definition

$$\operatorname{cap}_p(E) = \inf \int_{\mathbb{R}^n} |Du|^p \, dx,$$

where the infimum is taken over all $u \in L^{p^*}(\mathbb{R}^n)$ with $|Du| \in L^p(\mathbb{R}^n)$, $u \geqslant 0$ and $u \geqslant 1$ on a neighbourhood of E. Some estimates and arguments may become more transparent with this definition, but we stick to our original definition.

(2) The definition of Sobolev capacity applies also for $p = 1$, but we shall not discuss this case here.

The Sobolev p-capacity enjoys many desirable properties, one of the most important of which says that it is an outer measure.

Theorem 4.4. The Sobolev p-capacity is an outer measure, that is,

(1) $\operatorname{cap}_p(\emptyset) = 0$,

(2) if $E_1 \subset E_2$, then $\operatorname{cap}_p(E_1) \leqslant \operatorname{cap}_p(E_2)$ and

(3) $\operatorname{cap}_p\left(\bigcup_{i=1}^{\infty} E_i\right) \leqslant \sum_{i=1}^{\infty} \operatorname{cap}_p(E_i)$ whenever $E_i \subset \mathbb{R}^n$, $i = 1, 2, \dots$.

THE MORAL: Capacity is an outer measure, but measure theory is useless since there are few measurable sets.

Proof. (1) Clearly $\operatorname{cap}_p(\emptyset) = 0$.

(2) $\mathscr{A}(E_2) \subset \mathscr{A}(E_1)$ implies $\operatorname{cap}_p(E_1) \leqslant \operatorname{cap}_p(E_2)$.

(3) Let $\varepsilon > 0$. We may assume that $\sum_{i=1}^{\infty} \operatorname{cap}_p(E_i) < \infty$. Choose $u_i \in \mathscr{A}(E_i)$ so that

$$\|u_i\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant \operatorname{cap}_p(E_i) + \varepsilon 2^{-i}, \quad i = 1, 2, \dots.$$

Claim: $v = \sup_i u_i$ is admissible for $\bigcup_{i=1}^{\infty} E_i$.

Reason. First we show that $v \in W^{1,p}(\mathbb{R}^n)$. Let

$$v_k = \max_{1 \leqslant i \leqslant k} u_i, \quad k = 1, 2, \dots.$$

Then (v_k) is an increasing sequence such that $v_k \to v$ pointwise as $k \to \infty$. Moreover

$$|v_k| = |\max_{1 \leqslant i \leqslant k} u_i| \leqslant |\sup_i u_i| = |v|, \quad k = 1, 2, \dots,$$

and by Remark 2.4

$$|Dv_k| \leqslant \max_{1 \leqslant i \leqslant k} |Du_i| \leqslant \sup_i |Du_i|, \quad k = 1, 2, \dots.$$

We show that (v_k) is a a bounded sequence in $W^{1,p}(\mathbb{R}^n)$. To conclude this, we observe that

$$\begin{aligned}
\|v_k\|^p_{W^{1,p}(\mathbb{R}^n)} &= \int_{\mathbb{R}^n} |v_k|^p \, dx + \int_{\mathbb{R}^n} |Dv_k|^p \, dx \\
&\leqslant \int_{\mathbb{R}^n} \sup_i |u_i|^p \, dx + \int_{\mathbb{R}^n} \sup_i |Du_i|^p \, dx \\
&\leqslant \int_{\mathbb{R}^n} \sum_{i=1}^{\infty} |u_i|^p \, dx + \int_{\mathbb{R}^n} \sum_{i=1}^{\infty} |Du_i|^p \, dx \\
&= \sum_{i=1}^{\infty} \left(\int_{\mathbb{R}^n} |u_i|^p \, dx + \int_{\mathbb{R}^n} |Du_i|^p \, dx \right) \\
&\leqslant \sum_{i=1}^{\infty} (\operatorname{cap}_p(E_i) + \varepsilon 2^{-i}) \\
&\leqslant \sum_{i=1}^{\infty} \operatorname{cap}_p(E_i) + \varepsilon < \infty, \quad k = 1, 2, \dots.
\end{aligned}$$

Since $v_k \to v$ almost everywhere, by weak compactness of Sobolev spaces, see Theorem 2.24, we conclude that $v \in W^{1,p}(\mathbb{R}^n)$. Since $u_i \in \mathscr{A}(E_i)$, there exists an open set $O_i \supset E_i$ such that $u_i \geqslant 1$ on O_i for every $i = 1,2,\dots$. It follows that $v = \sup_i u_i \geqslant 1$ on $\bigcup_{i=1}^{\infty} O_i$, which is a neighbourhood of $\bigcup_{i=1}^{\infty} E_i$. ■

We conclude that

$$\operatorname{cap}_p\Big(\bigcup_{i=1}^{\infty} E_i\Big) \leqslant \|v\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant \sum_{i=1}^{\infty} \|u_i\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant \sum_{i=1}^{\infty} \operatorname{cap}_p(E_i) + \varepsilon.$$

The claim follows by letting $\varepsilon \to 0$. □

Remark 4.5. The Sobolev p-capacity is outer regular, that is,

$$\operatorname{cap}_p(E) = \inf\{\operatorname{cap}_p(O) : E \subset O, O \text{ open}\}.$$

Reason. (1) By monotonicity,

$$\operatorname{cap}_p(E) \leqslant \inf\{\operatorname{cap}_p(O) : E \subset O, O \text{ open}\}.$$

(2) To see the inequality in the other direction, let $\varepsilon > 0$ and take $u \in \mathscr{A}(E)$ such that

$$\|u\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant \operatorname{cap}_p(E) + \varepsilon.$$

Since $u \in \mathscr{A}(E)$ there is an open set O containing E such that $u \geqslant 1$ on O, which implies

$$\operatorname{cap}_p(O) \leqslant \|u\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant \operatorname{cap}_p(E) + \varepsilon.$$

The claim follows by letting $\varepsilon \to 0$. ■

THE MORAL: The capacity of a set is completely determined by the capacities of open sets containing the set. The same applies to the Lebesgue outer measure.

Next we discuss monotone convergence theorems for capacity. Note that these results do not immediately follow from the corresponding results in measure theory, since there are few measurable sets for capacity. We begin with monotone convergence for an increasing sequence of sets.

Theorem 4.6. Let $1 < p < \infty$ and let $E_i \subset \mathbb{R}^n$, $i = 1,2,\dots$, be arbitrary sets such that $E_i \subset E_{i+1}$ for every $i = 1,2,\dots$. Then

$$\operatorname{cap}_p\Big(\bigcup_{i=1}^{\infty} E_i\Big) = \lim_{i\to\infty} \operatorname{cap}_p(E_i).$$

Proof. (1) Let $E = \bigcup_{i=1}^{\infty} E_i$. By monotonicity, we have

$$\lim_{i\to\infty} \operatorname{cap}_p(E_i) \leqslant \operatorname{cap}_p(E).$$

Note that monotonicity also implies that the limit on the left-hand side exists.

(2) For the reverse inequality, we may assume that

$$\lim_{k\to\infty} \operatorname{cap}_p(E_k) < \infty.$$

Let $\varepsilon > 0$ and let $v_k \in \mathscr{A}(E_k)$ be such that

$$\|v_k\|^p_{W^{1,p}(\mathbb{R}^n)} < \operatorname{cap}_p(E_k) + \frac{\varepsilon}{2^k}, \quad k = 1,2,\dots.$$

Let $u_0 = 0$ and

$$u_i = \max\{v_k : k = 1,\dots,i\}, \quad i = 1,2,\dots.$$

By Remark 2.4 (1), we have

$$\max\{u_{i-1}, v_i\} \in W^{1,p}(\mathbb{R}^n) \quad i = 1,2,\dots,$$

and

$$\min\{u_{i-1}, v_i\} \in W^{1,p}(\mathbb{R}^n), \quad i = 1,2,\dots.$$

Moreover,

$$D\min\{u_{i-1}, v_i\} = \begin{cases} Du_{i-1} & \text{a.e. in } \{u_{i-1} \leqslant v_i\}, \\ Dv_i & \text{a.e. in } \{u_{i-1} > v_i\}, \end{cases}$$

and

$$D\max\{u_{i-1}, v_i\} = \begin{cases} Du_{i-1} & \text{a.e. in } \{u_{i-1} > v_i\}, \\ Dv_i & \text{a.e. in } \{u_{i-1} \leqslant v_i\}. \end{cases}$$

Let $E_0 = \emptyset$. We note that $u_i = \max\{u_{i-1}, v_i\}$ and $\min\{u_{i-1}, v_i\} \geqslant 1$ in a neighbourhood of E_{i-1}, $i = 1,2,\dots$, and thus we have

$$\begin{aligned} &\|u_i\|^p_{W^{1,p}(\mathbb{R}^n)} + \operatorname{cap}_p(E_{i-1}) \\ &\quad\leqslant \|\max\{u_{i-1}, v_i\}\|^p_{W^{1,p}(\mathbb{R}^n)} + \|\min\{u_{i-1}, v_i\}\|^p_{W^{1,p}(\mathbb{R}^n)}, \quad i = 1,2,\dots. \end{aligned}$$

where

$$\begin{aligned} \|\max\{u_{i-1}, v_i\}\|^p_{W^{1,p}(\mathbb{R}^n)} &= \int_{\mathbb{R}^n} (|\max\{u_{i-1}, v_i\}|^p + |D\max\{u_{i-1}, v_i\}|^p)\,dx \\ &= \int_{\{u_{i-1} > v_i\}} (|u_{i-1}|^p + |Du_{i-1}|^p)\,dx + \int_{\{u_{i-1} \leqslant v_i\}} (|v_i|^p + |Dv_i|^p)\,dx \end{aligned}$$

and

$$\begin{aligned} \|\min\{u_{i-1}, v_i\}\|^p_{W^{1,p}(\mathbb{R}^n)} &= \int_{\mathbb{R}^n} (|\min\{u_{i-1}, v_i\}|^p + |D\min\{u_{i-1}, v_i\}|^p)\,dx \\ &= \int_{\{u_{i-1} \leqslant v_i\}} (|u_{i-1}|^p + |Du_{i-1}|^p)\,dx + \int_{\{u_{i-1} > v_i\}} (|v_i|^p + |Dv_i|^p)\,dx. \end{aligned}$$

This implies that

$$\begin{aligned} &\|\max\{u_{i-1}, v_i\}\|^p_{W^{1,p}(\mathbb{R}^n)} + \|\min\{u_{i-1}, v_i\}\|^p_{W^{1,p}(\mathbb{R}^n)} \\ &\quad= \int_{\mathbb{R}^n} (|u_{i-1}|^p + |Du_{i-1}|^p)\,dx + \int_{\mathbb{R}^n} (|v_i|^p + |Dv_i|^p)\,dx \\ &\quad= \|u_{i-1}\|^p_{W^{1,p}(\mathbb{R}^n)} + \|v_i\|^p_{W^{1,p}(\mathbb{R}^n)} \\ &\quad\leqslant \|u_{i-1}\|^p_{W^{1,p}(\mathbb{R}^n)} + \operatorname{cap}_p(E_i) + \frac{\varepsilon}{2^i}, \quad i = 1,2,\dots. \end{aligned}$$

Thus we have

$$\|u_i\|^p_{W^{1,p}(\mathbb{R}^n)} - \|u_{i-1}\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant \operatorname{cap}_p(E_i) - \operatorname{cap}_p(E_{i-1}) + \frac{\varepsilon}{2^i}, \quad i = 1,2,\dots,$$

and consequently

$$\begin{aligned}
\|u_i\|^p_{W^{1,p}(\mathbb{R}^n)} &= \|u_i\|^p_{W^{1,p}(\mathbb{R}^n)} - \|u_0\|^p_{W^{1,p}(\mathbb{R}^n)} \\
&= \sum_{k=1}^{i} \left(\|u_i\|^p_{W^{1,p}(\mathbb{R}^n)} - \|u_{i-1}\|^p_{W^{1,p}(\mathbb{R}^n)}\right) \\
&\leqslant \sum_{k=1}^{i} \left(\operatorname{cap}_p(E_i) - \operatorname{cap}_p(E_{i-1}) + \frac{\varepsilon}{2^i}\right) \\
&\leqslant \operatorname{cap}_p(E_i) + \varepsilon.
\end{aligned}$$

Since $u_{i-1} \leqslant u_i$, the sequence (u_i) is increasing and it converges pointwise. Let

$$u = \lim_{i\to\infty} u_i.$$

Since $u \geqslant u_i \geqslant 1$ in an open set $O_i \supset E_i$, we conclude that $u \geqslant 1$ in $\bigcup_{i=1}^{\infty} O_i$ and $\bigcup_{i=1}^{\infty} E_i \subset \bigcup_{i=1}^{\infty} O_i$. Since $E_i \subset E_{i+1}$, $i = 1,2,\dots$, by monotonicity, we have

$$\sup_i \operatorname{cap}_p(E_i) = \lim_{i\to\infty} \operatorname{cap}_p(E_i) < \infty,$$

and consequently (u_i) is a bounded sequence in $W^{1,p}(\mathbb{R}^n)$. By Theorem 2.24 we conclude that $u \in W^{1,p}(\mathbb{R}^n)$, $u_i \to u$ weakly in $L^p(\Omega)$ and $Du_i \to Du$ weakly in $L^p(\Omega;\mathbb{R}^n)$ as $i \to \infty$. It follows that $u \in \mathscr{A}(E)$. By Lemma 2.11, we obtain

$$\begin{aligned}
\operatorname{cap}_p(E) \leqslant \|u\|^p_{W^{1,p}(\mathbb{R}^n)} &\leqslant \liminf_{i\to\infty} \|u_i\|^p_{W^{1,p}(\mathbb{R}^n)} \\
&\leqslant \liminf_{i\to\infty} \operatorname{cap}_p(E_i) + \varepsilon \\
&= \lim_{i\to\infty} \operatorname{cap}_p(E_i) + \varepsilon.
\end{aligned}$$

The claim follows by letting $\varepsilon \to 0$. □

Then we discuss monotone convergence for decreasing sequence of sets.

Theorem 4.7. Let K_i, $i = 1,2,\dots$, be compact sets in $\mathbb{R}^n$ such that $K_{i+1} \subset K_i$ for every $i = 1,2,\dots$. Then

$$\operatorname{cap}_p\Big(\bigcap_{i=1}^{\infty} K_i\Big) = \lim_{i\to\infty} \operatorname{cap}_p(K_i).$$

Proof. (1) Let $K = \bigcap_{i=1}^{\infty} K_i$. By monotonicity, we have

$$\operatorname{cap}_p(K) \leqslant \lim_{i\to\infty} \operatorname{cap}_p(K_i).$$

Note that monotonicity also implies that the limit on the right-hand side exists. We note that $\operatorname{cap}_p(K) < \infty$ for every compact set $K \subset \mathbb{R}^n$, see Remark 4.12 (1).

(2) For the reverse inequality, let $\varepsilon > 0$ and let $O \supset K$ be an open set such that

$$\operatorname{cap}_p(O) < \operatorname{cap}_p(K) + \varepsilon.$$

Then $O \cup \bigcup_{i=1}^{\infty}(\mathbb{R}^n \setminus K_i)$ is an open covering of K_1. Since K_1 is compact, there exists a finite subcovering, that is, the exists an index k such that

$$K_1 \subset O \cup \bigcup_{i=1}^{k}(\mathbb{R}^n \setminus K_i) = O \cup \Big(\mathbb{R}^n \setminus \bigcap_{i=1}^{k} K_i\Big) = O \cup (\mathbb{R}^n \setminus K_k).$$

Since $K_k \subset K_1$ it follows that $K_k \subset O$. This implies that

$$\lim_{i\to\infty} \operatorname{cap}_p(K_i) \leqslant \operatorname{cap}_p(K_k) \leqslant \operatorname{cap}_p(O) < \operatorname{cap}_p(K) + \varepsilon.$$

It follows that that

$$\lim_{i\to\infty} \operatorname{cap}_p(K_i) \leqslant \operatorname{cap}_p(K).$$

□

Remark 4.8. It is essential that the sets in Theorem 4.7 are compact. For example, let $F_i = \mathbb{R}^n \setminus B(0,i)$, $i = 1,2,\dots$. Then F_i, $i = 1,2,\dots$, are closed but unbounded and by Lemma 4.9, we have

$$\operatorname{cap}_p(F_i) \geqslant |F_k| = \infty, \quad i = 1,2,\dots,$$

but

$$\operatorname{cap}_p\Big(\bigcap_{i=1}^{\infty} F_i\Big) = \operatorname{cap}_p(\emptyset) = 0.$$

A similar example can also be constructed by applying bounded sets (exercise).

4.2 Capacity and measure

We are mainly interested in sets of vanishing capacity, since they are in some sense exceptional sets in the theory of Sobolev spaces. Our first result is rather immediate.

Lemma 4.9. $|E| \leqslant \operatorname{cap}_p(E)$ for every $E \subset \mathbb{R}^n$.

THE MORAL: Sets of capacity zero are of measure zero. Thus capacity is a finer measure than Lebesgue measure.

Proof. If $\operatorname{cap}_p(E) = \infty$, there is nothing to prove. Thus we may assume that $\operatorname{cap}_p(E) < \infty$. Let $\varepsilon > 0$ and take $u \in \mathscr{A}(E)$ such that

$$\|u\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant \operatorname{cap}_p(E) + \varepsilon.$$

There is an open $O \supset E$ such that $u \geqslant 1$ in O and thus

$$|E| \leqslant |O| \leqslant \int_O |u|^p \, dx \leqslant \|u\|^p_{L^p(\mathbb{R}^n)} \leqslant \|u\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant \operatorname{cap}_p(E) + \varepsilon.$$

The claim follows by letting $\varepsilon \to 0$. □

Remark 4.10. Lemma 4.9 shows that $\operatorname{cap}_p(B(x,r)) > 0$ for every $x \in \mathbb{R}^n$, $r > 0$. This implies that capacity is nontrivial in the sense that every nonempty open set has positive capacity.

Lemma 4.11. Let $x \in \mathbb{R}^n$ and $0 < r \leqslant 1$. Then there exists $c = c(n,p)$ such that

$$\operatorname{cap}_p(B(x,r)) \leqslant cr^{n-p}$$

THE MORAL: For the Lebesgue measure of a ball we have $|B(x,r)| \leqslant cr^n$, but for the Sobolev capacity of a ball we have $\operatorname{cap}_p(B(x,r)) \leqslant cr^{n-p}$. Thus the natural scaling dimension for capacity is $n-p$. Observe, that the dimension for capacity is smaller than $n-1$.

Proof. Define a cutoff function

$$u(y) = \begin{cases} 1, & y \in B(x,r), \\ 2 - \frac{|y-x|}{r}, & y \in B(x,2r) \setminus B(x,r), \\ 0, & y \in \mathbb{R}^n \setminus B(x,2r). \end{cases}$$

Observe that $0 \leqslant u \leqslant 1$, u is $\frac{1}{r}$-Lipschitz and $|Du| \leqslant \frac{1}{r}$ almost everywhere. Thus $u \in \mathscr{A}(B(x,r))$ and

$$\begin{aligned}
\operatorname{cap}_p(B(x,r)) &\leqslant \int_{B(x,2r)} |u|^p\,dy + \int_{B(x,2r)} |Du|^p\,dy \\
&\leqslant (1 + r^{-p})|B(x,2r)| \leqslant (r^{-p} + r^{-p})|B(x,2r)| \\
&= 2r^{-p}|B(x,2r)| = cr^{n-p},
\end{aligned}$$

with $c = c(n,p)$ □

Remarks 4.12:

(1) Lemma 4.11 shows that every bounded set has finite capacity. Thus there are plenty of sets with finite capacity.

Reason. Assume that $E \subset \mathbb{R}^n$ is bounded. Then $E \subset B(0,r)$ for some r, $1 \leqslant r < \infty$, and the proof of Lemma 4.11 gives

$$\operatorname{cap}_p(E) \leqslant \operatorname{cap}_p(B(0,r)) \leqslant cr^{p+n-p} < \infty.$$ ■

(2) Lemma 4.11 implies that $\operatorname{cap}_p(\{x\}) = 0$ for every $x \in \mathbb{R}^n$ when $1 < p < n$.

Reason.

$$\operatorname{cap}_p(\{x\}) \leqslant \operatorname{cap}_p(B(x,r)) \leqslant cr^{n-p}, \quad 0 < r \leqslant 1.$$

The claim follows by letting $r \to 0$. ■

(3) When $1 < p < n$, by applying the Gagliardo–Nirenberg–Sobolev inequality (Theorem 3.3), it is possible to show (exercise) that there exists a constant $c = c(n,p) > 0$ such that

$$\operatorname{cap}_p(B(x,r)) \geqslant cr^{n-p}, \quad r > 0.$$

Remark 4.13. Let $x \in \mathbb{R}^n$ and $0 < r \leqslant \frac{1}{2}$. Then there exists $c = c(n)$ such that

$$\operatorname{cap}_n(B(x,r)) \leqslant c\left(\log\frac{1}{r}\right)^{1-n}.$$

Reason. Use the test function

$$u(y) = \begin{cases} \left(\log\frac{1}{r}\right)^{-1}\log\frac{1}{|x-y|}, & y \in B(x,1)\setminus B(x,r), \\ 1, & y \in B(x,r), \\ 0, & y \in \mathbb{R}^n \setminus B(x,1). \end{cases}$$

This implies that $\operatorname{cap}_p(\{x\}) = 0$ for every $x \in \mathbb{R}^n$ when $p = n$ (exercise). ■

We have shown that a point has zero capacity when $1 < p \leqslant n$. By countable subadditivity all countable sets have zero capacity as well. Next we show that a point has positive capacity when $p > n$.

Lemma 4.14. If $p > n$, then $\operatorname{cap}_p(\{x\}) > 0$ for every $x \in \mathbb{R}^n$.

T H E M O R A L : For $p > n$ every set containing at least one point has a positive capacity. Thus there are no nontrivial sets of capacity zero. In practice this means that capacity is a useful tool only when $p \leqslant n$.

Proof. Let $x \in \mathbb{R}^n$ and assume that $u \in \mathscr{A}(\{x\})$. Then there exists $0 < r \leqslant 1$ such that $u(y) \geqslant 1$ on $B(x,r)$. Take a cutoff function $\eta \in C_0^\infty(B(x,2))$ such that $0 \leqslant \eta \leqslant 1$, $\eta = 1$ in $B(x,r)$ and $|D\eta| \leqslant 2$. By Morrey's inequality, see Theorem 3.23, there exists $c = c(n,p) > 0$ such that

$$|(\eta u)(y) - (\eta u)(z)| \leqslant c|y-z|^{1-\frac{n}{p}} \|D(\eta u)\|_{L^p(\mathbb{R}^n)}$$

for almost every $y, z \in \mathbb{R}^n$. Choose $y \in B(x,r)$ and $z \in B(x,4)\setminus B(x,2)$ so that $(\eta u)(y) \geqslant 1$ and $(\eta u)(z) = 0$. Then $1 \leqslant |y-z| \leqslant 5$ and thus

$$\begin{aligned} \int_{B(x,2)} |D(\eta u)(y)|^p\,dy &= \|D(\eta u)\|_{L^p(\mathbb{R}^n)}^p \\ &\geqslant c\underbrace{|y-z|^{n-p}}_{\geqslant 5^{n-p}}\underbrace{|(\eta u)(y) - (\eta u)(z)|^p}_{\geqslant 1} \geqslant c > 0. \end{aligned}$$

On the other hand

$$\begin{aligned}\int_{B(x,2)} |D(\eta u)|^p\,dy &\leqslant 2^p\left(\int_{B(x,2)} |D\eta u|^p\,dy + \int_{B(x,2)} |\eta D u|^p\,dy\right)\\ &= 2^p\left(\int_{B(x,2)} \underbrace{|D\eta|^p}_{\leqslant 2^p} |u|^p\,dy + \int_{B(x,2)} \underbrace{|\eta|^p}_{\leqslant 1} |Du|^p\,dy\right)\\ &\leqslant 4^p\left(\int_{B(x,2)} |u|^p\,dy + \int_{B(x,2)} |Du|^p\,dy\right)\\ &\leqslant 4^p \|u\|^p_{W^{1,p}(\Omega)}.\end{aligned}$$

This shows that there exists $c = c(n,p) > 0$ such that $\|u\|^p_{W^{1,p}(\Omega)} \geqslant c > 0$ for every $u \in \mathscr{A}(\{x\})$ and thus $\operatorname{cap}_p(\{x\}) \geqslant c > 0$. □

In order to study the connection between capacity and measure, we need to consider lower dimensional measures than the Lebesgue measure. We recall the definition of Hausdorff measures.

Definition 4.15. Let $E \subset \mathbb{R}^n$ and $s \geqslant 0$. For $0 < \delta \leqslant \infty$ we set

$$\mathscr{H}^s_\delta(E) = \inf\left\{\sum_{i=1}^\infty r_i^s : E \subset \bigcup_{i=1}^\infty B(x_i, r_i),\, r_i \leqslant \delta\right\}.$$

The (spherical) s-Hausdorff measure of E is

$$\mathscr{H}^s(E) = \lim_{\delta\to 0} \mathscr{H}^s_\delta(E) = \sup_{\delta>0} \mathscr{H}^s_\delta(E).$$

The Hausdorff dimension of E is

$$\inf\{s : \mathscr{H}^s(E) = 0\} = \sup\{s : \mathscr{H}^s(E) = \infty\}.$$

THE MORAL: The Hausdorff measure is the natural s-dimensional measure up to scaling and the Hausdorff dimension is the measure theoretic dimension of the set. Observe that the dimension can be any nonnegative real number less or equal than the dimension of the space.

We begin by proving a useful measure theoretic lemma. In the proof we need some tools from measure and integration theory and real analysis.

Lemma 4.16. Assume that $0 < s < n$, $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and let

$$E = \left\{x \in \mathbb{R}^n : \limsup_{r\to 0} \frac{1}{r^s}\int_{B(x,r)} |f|\,dy > 0\right\}.$$

Then $\mathscr{H}^s(E) = 0$.

THE MORAL: Roughly speaking the lemma above says that the set where a locally integrable function blows up rapidly is of the corresponding Hausdorff measure zero.

Proof. (1) Assume first that $f \in L^1(\mathbb{R}^n)$.

(2) By the Lebesgue differentiation theorem

$$\lim_{r\to 0} \fint_{B(x,r)} |f|\,dy = |f(x)| < \infty,$$

for almost every $x \in \mathbb{R}^n$. If x is a Lebesgue point of $|f|$, then

$$\limsup_{r\to 0} \frac{1}{r^s} \int_{B(x,r)} |f|\,dy = c \limsup_{r\to 0} r^{n-s} \fint_{B(x,r)} |f|\,dy = 0.$$

This shows that all Lebesgue points of $|f|$ belong to the complement of E and thus $|E| = 0$.

(3) Let $\varepsilon > 0$ and

$$E_\varepsilon = \left\{ x \in \mathbb{R}^n : \limsup_{r\to 0} \frac{1}{r^s} \int_{B(x,r)} |f|\,dy > \varepsilon \right\}.$$

Since $E_\varepsilon \subset E$ and $|E| = 0$, we have $|E_\varepsilon| = 0$.

Claim: $\mathcal{H}^s(E_\varepsilon) = 0$ for every $\varepsilon > 0$.

Reason. Let $0 < \delta < 1$. For every $x \in E_\varepsilon$ there exists a r_x with $0 < r_x \leqslant \delta$ such that

$$\frac{1}{r_x^s} \int_{B(x,r_x)} |f|\,dy > \varepsilon.$$

By the Vitali covering theorem, there exists a subfamily of countably many pairwise disjoint balls $B(x_i, r_i)$, $i = 1, 2, \dots$, such that

$$E_\varepsilon \subset \bigcup_{i=1}^\infty B(x_i, 5r_i).$$

This gives

$$\mathcal{H}^s_{5\delta}(E_\varepsilon) \leqslant \sum_{i=1}^\infty (5r_i)^s \leqslant \frac{5^s}{\varepsilon} \sum_{i=1}^\infty \int_{B(x_i,r_i)} |f|\,dy = \frac{5^s}{\varepsilon} \int_{\bigcup_{i=1}^\infty B(x_i,r_i)} |f|\,dy.$$

By disjointness of the balls

$$\begin{aligned}
\left| \bigcup_{i=1}^\infty B(x, r_i) \right| &= \sum_{i=1}^\infty |B(x_i, r_i)| = c \sum_{i=1}^\infty r_i^n \\
&\leqslant c \sum_{i=1}^\infty \frac{r_i^n}{\varepsilon r_i^s} \int_{B(x_i,r_i)} |f|\,dy \\
&\leqslant c \frac{\delta^{n-s}}{\varepsilon} \int_{\mathbb{R}^n} |f|\,dy \xrightarrow{\delta\to 0} 0.
\end{aligned}$$

By absolute continuity of integral

$$\int_{\bigcup_{i=1}^\infty B(x_i,r_i)} |f|\,dy \xrightarrow{\delta\to 0} 0.$$

Thus

$$\mathcal{H}^s(E_\varepsilon) = \lim_{\delta\to 0} \mathcal{H}^s_{5\delta}(E_\varepsilon) \leqslant \frac{5^s}{\varepsilon} \lim_{\delta\to 0} \int_{\bigcup_{i=1}^\infty B(x_i,r_i)} |f|\,dy = 0.$$

This shows that $\mathcal{H}^s(E_\varepsilon) = 0$ for every $\varepsilon > 0$. ■

(4) By subadditivity of the Hausdorff measure

$$\mathcal{H}^s(E) = \mathcal{H}^s\left(\bigcup_{k=1}^{\infty} E_{\frac{1}{k}}\right) \leqslant \sum_{k=1}^{\infty} \mathcal{H}^s(E_{\frac{1}{k}}) = 0.$$

This shows that $\mathcal{H}^s(E) = 0$.

(5) Assume then that $f \in L^1_{\text{loc}}(\mathbb{R}^n)$. Then

$$\begin{aligned}
\mathcal{H}^s(E) &= \mathcal{H}^s\left(\left\{x \in \mathbb{R}^n : \limsup_{r\to 0} \frac{1}{r^s}\int_{B(x,r)} |f|\,dy > 0\right\}\right)\\
&= \mathcal{H}^s\left(\bigcup_{k=1}^{\infty}\left\{x \in \mathbb{R}^n : \limsup_{r\to 0} \frac{1}{r^s}\int_{B(x,r)} |f\chi_{B(0,k)}|\,dy > 0\right\}\right)\\
&\leqslant \sum_{k=1}^{\infty}\mathcal{H}^s\left(\left\{x \in \mathbb{R}^n : \limsup_{r\to 0} \frac{1}{r^s}\int_{B(x,r)} |f\chi_{B(0,k)}|\,dy > 0\right\}\right) = 0.
\end{aligned}$$

□

Next we compare capacity to the Hausdorff measure.

Theorem 4.17. Assume that $1 < p < n$. Then there exists $c = c(n,p)$ such that $\text{cap}_p(E) \leqslant c\mathcal{H}^{n-p}(E)$ for every $E \subset \mathbb{R}^n$.

T H E M O R A L : Capacity is smaller than $(n-p)$-dimensional Hausdorff measure. In particular, $\mathcal{H}^{n-p}(E) = 0$ implies $\text{cap}_p(E) = 0$.

Proof. Let $B(x_i, r_i)$, $i = 1,2,\dots$, be any covering of E such that the radii satisfy $r_i \leqslant \delta$. Subadditivity implies

$$\text{cap}_p(E) \leqslant \sum_{i=1}^{\infty} \text{cap}_p(B(x_i,r_i)) \leqslant c\sum_{i=1}^{\infty} r_i^{n-p}.$$

By taking the infimum over all coverings by such balls and observing that $\mathcal{H}^s_\delta(E) \leqslant \mathcal{H}^s(E)$ we obtain

$$\text{cap}_p(E) \leqslant c\mathcal{H}^{n-p}_\delta(E) \leqslant c\mathcal{H}^{n-p}(E).$$

□

We next consider the converse of the previous theorem. We prove that sets of p-capacity zero have Hausdorff dimension at most $n-p$.

Theorem 4.18. Assume that $1 < p < n$. If $E \subset \mathbb{R}^n$ with $\text{cap}_p(E) = 0$, then $\mathcal{H}^s(E) = 0$ for all $s > n-p$.

Proof. (1) Let $E \subset \mathbb{R}^n$ be such that $\text{cap}_p(E) = 0$. Then for every $i = 1,2,\dots,$ there exists $u_i \in \mathcal{A}'(E)$ such that $\|u_i\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant 2^{-i}$. Let $u = \sum_{i=1}^{\infty} u_i$.

Claim: $u \in \mathcal{A}(E)$.

Reason. Let $v_k = \sum_{i=1}^{k} u_i$, $k = 1,2,\dots$. Then $v_k \in W^{1,p}(\mathbb{R}^n)$ and

$$\begin{aligned}
\|v_k\|_{W^{1,p}(\mathbb{R}^n)} = \left\|\sum_{i=1}^{k} u_i\right\|_{W^{1,p}(\mathbb{R}^n)} &\leqslant \sum_{i=1}^{k}\|u_i\|_{W^{1,p}(\mathbb{R}^n)}\\
&\leqslant \sum_{i=1}^{\infty}\|u_i\|_{W^{1,p}(\mathbb{R}^n)} \leqslant \sum_{i=1}^{\infty} 2^{-\frac{i}{p}} < \infty.
\end{aligned}$$

Thus (v_k) is a bounded sequence in $W^{1,p}(\mathbb{R}^n)$. Since $0 \leqslant u_i \leqslant 1$, we observe that (v_k) is an increasing sequence and thus $v_k \to u$ almost everywhere. Theorem 2.24 implies $u \in W^{1,p}(\mathbb{R}^n)$. Moreover, $u \geqslant 1$ almost everywhere on a neighbourhood of E which shows that $u \in \mathscr{A}(E)$. ■

(2) **Claim:**

$$\limsup_{r\to 0} \fint_{B(x,r)} u\,dy = \infty \quad \text{for every} \quad x \in E. \tag{4.19}$$

Reason. Let $m \in \mathbb{N}$ and $x \in E$. Then for $r > 0$ small enough $B(x,r)$ is contained in an intersection of open sets O_i, $i = 1,\dots,m$, with the property that $u_i = 1$ almost everywhere on O_i. This implies that $u = \sum_{i=1}^{\infty} u_i \geqslant m$ almost everywhere in $B(x,r)$ and thus

$$\fint_{B(x,r)} u\,dy \geqslant m.$$

This proves the claim. ■

THE MORAL: This gives a method to construct a function that blows up on any set of zero capacity.

(3) **Claim:** If $s > n - p$, then

$$\limsup_{r\to 0} \frac{1}{r^s} \int_{B(x,r)} |Du|^p\,dy = \infty \quad \text{for every} \quad x \in E.$$

Reason. Let $x \in E$ and, for a contradiction, assume that

$$\limsup_{r\to 0} \frac{1}{r^s} \int_{B(x,r)} |Du|^p\,dy < \infty.$$

Then there exists $c < \infty$ such that

$$\limsup_{r\to 0} \frac{1}{r^s} \int_{B(x,r)} |Du|^p\,dy \leqslant c.$$

Then we choose $R > 0$ so small that

$$\int_{B(x,r)} |Du|^p\,dy \leqslant cr^s$$

for every $0 < r \leqslant R$. Denote $B_i = B(x, 2^{-i}R)$, $i = 1,2,\dots$. Then by Hölder's inequality and the Poincaré inequality, see Theorem 5.25, we have

$$\begin{aligned}
|u_{B_{i+1}} - u_{B_i}| &\leqslant \fint_{B_{i+1}} |u - u_{B_i}|\,dy \\
&\leqslant \frac{|B_i|}{|B_{i+1}|} \fint_{B_i} |u - u_{B_i}|\,dy \\
&\leqslant c \left(\fint_{B_i} |u - u_{B_i}|^p\,dy \right)^{\frac{1}{p}} \\
&\leqslant c2^{-i}R \left(\fint_{B_i} |Du|^p\,dy \right)^{\frac{1}{p}} \\
&\leqslant c(2^{-i}R)^{\frac{p-n+s}{p}}.
\end{aligned}$$

For $k > j$, we obtain

$$|u_{B_k} - u_{B_j}| \leqslant \sum_{i=j}^{k-1} |u_{B_{i+1}} - u_{B_i}| \leqslant c \sum_{i=j}^{k-1} (2^{-i}R)^{\frac{p-n+s}{p}}$$

and thus (u_{B_i}) is a Cauchy sequence when $s > n - p$. This contradicts (4.19) and thus the claim holds true. ■

(4) Thus

$$\begin{aligned} E &\subset \left\{ x \in \mathbb{R}^n : \limsup_{r\to 0} \frac{1}{r^s} \int_{B(x,r)} |Du|^p \, dy = \infty \right\} \\ &\subset \left\{ x \in \mathbb{R}^n : \limsup_{r\to 0} \frac{1}{r^s} \int_{B(x,r)} |Du|^p \, dy > 0 \right\}. \end{aligned}$$

Lemma 4.16 implies

$$\mathcal{H}^s(E) \leqslant \mathcal{H}^s \left(\left\{ x \in \mathbb{R}^n : \limsup_{r\to 0} \frac{1}{r^s} \int_{B(x,r)} |Du|^p \, dy > 0 \right\} \right) = 0.$$

This shows that $\mathcal{H}^s(E) = 0$ whenever $n - p < s < n$. The claim follows from this, since $\mathcal{H}^s(E) = 0$ implies $\mathcal{H}^t(E) = 0$ for every $t \geqslant s$. □

Remark 4.20. It can be shown that even $\mathcal{H}^{n-p}(E) < \infty$, $1 < p < n$, implies $\operatorname{cap}_p(E) = 0$.

4.3 Quasicontinuity

In this section we study fine properties of Sobolev functions. It turns out that Sobolev functions are defined up to a set of capacity zero.

Definition 4.21. We say that a property holds p-quasieverywhere, if it holds except for a set of p-capacity zero.

THE MORAL: Quasieverywhere is a capacitary version of almost everywhere.

Recall that by Meyers-Serrin theorem 1.21 $W^{1,p}(\mathbb{R}^n) \cap C(\mathbb{R}^n)$ is dense in $W^{1,p}(\mathbb{R}^n)$ for $1 \leqslant p < \infty$ and, by Theorem 1.15, the Sobolev space $W^{1,p}(\mathbb{R}^n)$ is complete. The next result gives a way to find a quasieverywhere converging subsequence.

Theorem 4.22. Assume that $u_i \in W^{1,p}(\mathbb{R}^n) \cap C(\mathbb{R}^n)$, $i = 1, 2, \ldots$, and that (u_i) is a Cauchy sequence in $W^{1,p}(\mathbb{R}^n)$. Then there is a subsequence of (u_i) that converges pointwise p-quasieverywhere in $\mathbb{R}^n$. Moreover, the convergence is uniform outside a set of arbitrarily small p-capacity.

THE MORAL: This is a Sobolev space version of the result that for every Cauchy sequence in $L^p(\mathbb{R}^n)$, there is a subsequence that converges pointwise almost everywhere. The claim concerning uniform convergence is a Sobolev space version of Egorov's theorem.

Proof. There exists a subsequence of (u_i), which we still denote by (u_i), such that

$$\sum_{i=1}^{\infty} 2^{ip} \|u_i - u_{i+1}\|_{W^{1,p}(\mathbb{R}^n)}^p < \infty.$$

For $i = 1, 2, \ldots$, let

$$E_i = \{x \in \mathbb{R}^n : |u_i(x) - u_{i+1}(x)| > 2^{-i}\} \quad \text{and} \quad F_j = \bigcup_{i=j}^{\infty} E_i.$$

By continuity, the set $E_i \subset \mathbb{R}^n$ is open and $2^i |u_i - u_{i+1}| \in \mathscr{A}(E_i)$ for every $i = 1, 2, \ldots$. Thus

$$\operatorname{cap}_p(E_i) \leqslant 2^{ip} \|u_i - u_{i+1}\|_{W^{1,p}(\mathbb{R}^n)}^p, \quad i = 1, 2, \ldots.$$

By subadditivity we obtain

$$\operatorname{cap}_p(F_j) \leqslant \sum_{i=j}^{\infty} \operatorname{cap}_p(E_i) \leqslant \sum_{i=j}^{\infty} 2^{ip} \|u_i - u_{i+1}\|_{W^{1,p}(\mathbb{R}^n)}^p.$$

Since $\bigcap_{j=1}^{\infty} F_j \subset F_j$ and $F_{j+1} \subset F_j$, $j = 1, 2, \ldots$, we have

$$\begin{aligned}
\operatorname{cap}_p\Big(\bigcap_{j=1}^{\infty} F_j\Big) &\leqslant \lim_{j\to\infty} \operatorname{cap}_p(F_j) \\
&\leqslant \lim_{j\to\infty} \sum_{i=j}^{\infty} 2^{ip} \|u_i - u_{i+1}\|_{W^{1,p}(\mathbb{R}^n)}^p = 0.
\end{aligned}$$

Here we used the fact that the tail of a convergent series tends to zero.

We note that

$$|u_m(x) - u_k(x)| \leqslant \sum_{i=m}^{k-1} |u_i(x) - u_{i+1}(x)| \leqslant \sum_{i=m}^{k-1} 2^{-i} \leqslant \sum_{i=m}^{\infty} 2^{-i} \leqslant 2^{1-m},$$

for every $k > m > j$ and for every

$$\begin{aligned}
x \in \mathbb{R}^n \setminus F_j = \mathbb{R}^n \setminus \bigcup_{i=j}^{\infty} E_i &= \bigcap_{i=j}^{\infty} (\mathbb{R}^n \setminus E_i) \\
&= \bigcap_{i=j}^{\infty} \{x \in \mathbb{R}^n : |u_i(x) - u_{i+1}(x)| \leqslant 2^{-i}\}.
\end{aligned}$$

This shows that $(u_i(x))$ is a Cauchy sequence if there exists $j \in \mathbb{N}$ such that $x \in \mathbb{R}^n \setminus F_j$. It follows that $(u_i(x))$ is a Cauchy sequence, if

$$x \in \bigcup_{i=1}^{\infty} (\mathbb{R}^n \setminus F_j) = \mathbb{R}^n \setminus \bigcap_{i=1}^{\infty} F_j.$$

Since $(u_i(x))$ converges if and only if it is a Cauchy sequence, we conclude that $(u_i(x))$ converges for every $\mathbb{R}^n \setminus \bigcap_{i=1}^{\infty} F_j$ with

$$\operatorname{cap}_p\Big(\bigcap_{i=1}^{\infty} F_j\Big) = 0.$$

Moreover, we have

$$\sup_{\mathbb{R}^n \setminus F_j} |u_j(x) - u_k(x)| \leqslant 2^{1-m}$$

for every $k > m > j$, which shows that the convergence is uniform in $\mathbb{R}^n \setminus F_j$ with

$$\lim_{j\to\infty} \operatorname{cap}_p(F_j) = 0.$$

□

Definition 4.23. A function $u : \mathbb{R}^n \to [-\infty, \infty]$ is p-quasicontinuous in $\mathbb{R}^n$ if for every $\varepsilon > 0$ there exists a set E such that $\operatorname{cap}_p(E) < \varepsilon$ and the restriction of u to $\mathbb{R}^n \setminus E$, denoted by $u|_{\mathbb{R}^n \setminus E}$, is continuous.

Remark 4.24. By outer regularity, see Remark 4.5, we may assume that E is open in the definition above.

The next result shows that a Sobolev function has a quasicontinous representative.

Corollary 4.25. For every $u \in W^{1,p}(\mathbb{R}^n)$ there exists a p-quasicontinuous function $v \in W^{1,p}(\mathbb{R}^n)$ such that $u = v$ almost everywhere in $\mathbb{R}^n$.

THE MORAL: Every L^p function is defined almost everywhere, but every $W^{1,p}$ function is defined quasieverywhere.

Proof. By Theorem 1.21, for every function $u \in W^{1,p}(\mathbb{R}^n)$, there are functions $u_i \in W^{1,p}(\mathbb{R}^n) \cap C(\mathbb{R}^n)$, $i = 1, 2, \ldots$, such that $u_i \to u$ in $W^{1,p}(\mathbb{R}^n)$ as $i \to \infty$. Let $\varepsilon > 0$. By Theorem 4.22 there exist a set E such that $\operatorname{cap}_p(E) < \varepsilon$ and subsequence that converges uniformly to u in $\mathbb{R}^n \setminus E$. Uniform convergence implies continuity of the limit function $v = u|_{\mathbb{R}^n \setminus E}$. By defining $v = u$ in E and applying Lemma 4.9, we have

$$|\{x \in \mathbb{R}^n : u(x) \neq v(x)\}| \leqslant |E| \leqslant \operatorname{cap}_p(E) < \varepsilon.$$

This completes the proof. □

Next we show that the quasicontinuous representative given by Corollary 4.25 is essentially unique. We begin with a useful observation.

Remark 4.26. If $G \subset \mathbb{R}^n$ is open and $E \subset \mathbb{R}^n$ with $|E| = 0$, then $\operatorname{cap}_p(G) = \operatorname{cap}_p(G \setminus E)$.

Reason. $\boxed{\geqslant}$ Monotonicity implies $\mathrm{cap}_p(G) \geqslant \mathrm{cap}_p(G \setminus E)$.

$\boxed{\leqslant}$ Let $\varepsilon > 0$ and let $u \in \mathscr{A}(G \setminus E)$ be such that

$$\|u\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant \mathrm{cap}_p(G \setminus E) + \varepsilon.$$

Then there exists an open $O \subset \mathbb{R}^n$ with $(G \setminus E) \subset O$ and $u \geqslant 1$ almost everywhere in O. Since $O \cup G$ is open $G \subset (O \cup G)$ and $u \geqslant 1$ almost everywhere in $O \cup (G \setminus E)$, and almost everywhere in $O \cup G$ since $|E| = 0$, we have $u \in \mathscr{A}(G)$.

$$\mathrm{cap}_p(G) \leqslant \|u\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant \mathrm{cap}_p(G \setminus E) + \varepsilon.$$

By letting $\varepsilon \to 0$, we obtain $\mathrm{cap}_p(G) \leqslant \mathrm{cap}_p(G \setminus E)$. ■

Theorem 4.27. Assume that u and v are p−quasicontinuous functions on $\mathbb{R}^n$. If $u = v$ almost everywhere in $\mathbb{R}^n$, then $u = v$ p-quasieverywhere in $\mathbb{R}^n$.

THE MORAL: Quasicontinuous representatives of Sobolev functions are unique.

Proof. Let $\varepsilon > 0$ and let $G \subset \mathbb{R}^n$ be an open set such that $\mathrm{cap}_p(G) < \varepsilon$ and that the restrictions of u and v to $\mathbb{R}^n \setminus G$ are continuous. Thus $\{x \in \mathbb{R}^n \setminus G : u(x) \neq v(x)\}$ is open in the relative topology on $\mathbb{R}^n \setminus G$, that is, there exists open $U \subset \mathbb{R}^n$ with

$$U \setminus G = \{x \in \mathbb{R}^n \setminus G : u(x) \neq v(x)\}$$

and

$$|U \setminus G| = |\{x \in \mathbb{R}^n \setminus G : u(x) \neq v(x)\}| = 0.$$

Moreover,

$$\{x \in \mathbb{R}^n : u(x) \neq v(x)\} \subset G \cup \{x \in \mathbb{R}^n \setminus G : u(x) \neq v(x)\} = G \cup U.$$

Remark 4.26 (1) with G and E replaced by $U \cup G$ and $U \setminus G$, respectively, implies

$$\mathrm{cap}_p(\{x \in \mathbb{R}^n : u(x) \neq v(x)\}) \leqslant \mathrm{cap}_p(G \cup U) = \mathrm{cap}_p(G) < \varepsilon.$$

This completes the proof. □

Remarks 4.28:

(1) The same proof gives the following local result: Assume that u and v are p−quasicontinuous on an open set $O \subset \mathbb{R}^n$. If $u = v$ almost everywhere in O, then $u = v$ p-quasieverywhere in O.

(2) Observe that if u and v are p−quasicontinuous and $u \leqslant v$ almost everywhere in an open set O, then $\max\{u - v, 0\} = 0$ almost everywhere in O and $\max\{u - v, 0\}$ is p−quasicontinuous. Then Theorem 4.27 implies $\max\{u - v, 0\} = 0$ p-quasieverywhere in O and consequently $u \leqslant v$ p-quasieverywhere in O.

(3) The previous theorem enables us to define the trace of a Sobolev function to an arbitrary set. If $u \in W^{1,p}(\mathbb{R}^n)$ and $E \subset \mathbb{R}^n$, then the trace of u to E is the restriction to E of any p−quasicontinuous representative of u. This definition is useful only if $\mathrm{cap}_p(E) > 0$.

4.4 Lebesgue points of Sobolev functions

We consider the Hardy-Littlewood maximal function in Definition 5.1. By the maximal function theorem with $p = 1$, see Theorem 5.3 (1), there exists $c = c(n)$ such that

$$|\{x \in \mathbb{R}^n : Mf(x) > \lambda\}| \leqslant \frac{c}{\lambda}\|f\|_{L^1(\mathbb{R}^n)}$$

for every $\lambda > 0$. By Chebyshev's inequality and the maximal function theorem with $1 < p < \infty$, see Theorem 5.3 (2), there exists $c = c(n,p)$ such that

$$|\{x \in \mathbb{R}^n : Mf(x) > \lambda\}| \leqslant \frac{1}{\lambda^p}\|Mf\|^p_{L^p(\mathbb{R}^n)} \leqslant \frac{c}{\lambda^p}\|f\|^p_{L^p(\mathbb{R}^n)}$$

for every $\lambda > 0$. Thus the Hardy-Littlewood maximal function satisfies weak type estimates with respect to Lebesgue measure for functions in $L^p(\mathbb{R}^n)$. Next we consider capacitary weak type estimates for functions in $W^{1,p}(\mathbb{R}^n)$.

Theorem 4.29. Assume that $u \in W^{1,p}(\mathbb{R}^n)$, $1 < p < \infty$. Then there exists $c = c(n,p)$ such that

$$\operatorname{cap}_p\left(\{x \in \mathbb{R}^n : Mu(x) > \lambda\}\right) \leqslant \frac{c}{\lambda^p}\|u\|^p_{W^{1,p}(\mathbb{R}^n)}.$$

for every $\lambda > 0$.

THE MORAL: This is a capacitary version of weak type estimates for the Hardy-Littlewood maximal function.

Proof. Denote $E_\lambda = \{x \in \mathbb{R}^n : Mu(x) > \lambda\}$. Then E_λ is open and by Theorem 5.4 $Mu \in W^{1,p}(\mathbb{R}^n)$. Thus

$$\frac{Mu}{\lambda} \in \mathscr{A}(E_\lambda).$$

Since the maximal operator is bounded on $W^{1,p}(\mathbb{R}^n)$, see (5.5), we obtain

$$\operatorname{cap}_p(E_\lambda) \leqslant \left\|\frac{Mu}{\lambda}\right\|^p_{W^{1,p}(\mathbb{R}^n)} = \frac{1}{\lambda^p}\|Mu\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant \frac{c}{\lambda^p}\|u\|^p_{W^{1,p}(\mathbb{R}^n)}.$$

□

This weak type inequality can be used in studying the pointwise behaviour of Sobolev functions. We recall that $x \in \mathbb{R}^n$ is a Lebesgue point for $u \in L^1_{\text{loc}}(\mathbb{R}^n)$ if the limit

$$u^*(x) = \lim_{r\to 0} \fint_{B(x,r)} u(y)\,dy$$

exists and

$$\lim_{r\to 0} \fint_{B(x,r)} |u(y) - u^*(x)|\,dy = 0.$$

The Lebesgue differentiation theorem states that almost all points are Lebesgue points for a locally integrable function. If a function belongs to $W^{1,p}(\mathbb{R}^n)$, then using the capacitary weak type estimate, see Theorem 4.29, we shall prove that it has Lebesgue points p-quasieverywhere. Moreover, we show that the p-quasicontinuous representative given by Corollary 4.25 is u^*.

We begin by proving a measure theoretic result, which is analogous to Lemma 4.16.

Lemma 4.30. Let $1 < p < \infty$, $f \in L^p(\mathbb{R}^n)$ and

$$E = \left\{ x \in \mathbb{R}^n : \limsup_{r\to 0} r^p \fint_{B(x,r)} |f(y)|^p \, dy > 0 \right\}.$$

Then $\operatorname{cap}_p(E) = 0$.

THE MORAL: Roughly speaking the lemma above says that the set where an L^p function blows up rapidly is of capacity zero. The main difference compared to Lemma 4.16 is that the size of the set is measured by capacity instead of Hausdorff measure.

Proof. The argument is similar to the proof of Lemma 4.16, but we reproduce it here.

(1) By the Lebesgue differentiation theorem

$$\lim_{r\to 0} \fint_{B(x,r)} |f(y)|^p \, dy = |f(x)|^p < \infty,$$

for almost every $x \in \mathbb{R}^n$. If x is a Lebesgue point of $|f|^p$, then

$$\limsup_{r\to 0} r^p \fint_{B(x,r)} |f(y)|^p \, dy = 0.$$

This shows that all Lebesgue points of $|f|^p$ belong to the complement of E and thus $|E| = 0$.

(2) Let $\varepsilon > 0$ and

$$E_\varepsilon = \left\{ x \in \mathbb{R}^n : \limsup_{r\to 0} r^p \fint_{B(x,r)} |f(y)|^p \, dy > \varepsilon \right\}.$$

Since $E_\varepsilon \subset E$ and $|E| = 0$, we have $|E_\varepsilon| = 0$. We show that $\operatorname{cap}_p(E_\varepsilon) = 0$ for every $\varepsilon > 0$, then the claim follows by subadditivity. Let $0 < \delta < \frac{1}{5}$. For every $x \in E_\varepsilon$ there is r_x with $0 < r_x \leqslant \delta$ such that

$$r_x^p \fint_{B(x,r_x)} |f(y)|^p \, dy > \varepsilon.$$

By the Vitali covering theorem, there exists a subfamily of countably many pairwise disjoint balls $B(x_i, r_i)$, $i = 1, 2, \dots$, such that

$$E_\varepsilon \subset \bigcup_{i=1}^{\infty} B(x_i, 5r_i).$$

By subadditivity of the capacity and Lemma 4.11 we have

$$\begin{aligned}\operatorname{cap}_p(E_\varepsilon) &\leqslant \sum_{i=1}^{\infty} \operatorname{cap}_p(B(x_i,5r_i)) \leqslant c\sum_{i=1}^{\infty} r_i^{n-p}\\ &\leqslant \frac{c}{\varepsilon}\sum_{i=1}^{\infty}\int_{B(x_i,r_i)} |f(y)|^p\,dy = \frac{c}{\varepsilon}\int_{\bigcup_{i=1}^{\infty} B(x_i,r_i)} |f(y)|^p\,dy.\end{aligned}$$

Here $c = c(n,p)$. Finally we observe that by the disjointness of the balls

$$\begin{aligned}\left|\bigcup_{i=1}^{\infty} B(x_i,r_i)\right| &= \sum_{i=1}^{\infty} |B(x_i,r_i)| \leqslant \sum_{i=1}^{\infty} \frac{r_i^p}{\varepsilon}\int_{B(x_i,r_i)} |f(y)|^p\,dy\\ &\leqslant \frac{\delta^p}{\varepsilon}\int_{\mathbb{R}^n} |f(y)|^p\,dy \xrightarrow{\delta\to 0} 0.\end{aligned}$$

By absolute continuity of integral

$$\int_{\bigcup_{i=1}^{\infty} B(x_i,r_i)} |f(y)|^p\,dy \xrightarrow{\delta\to 0} 0.$$

Thus

$$\operatorname{cap}_p(E_\varepsilon) \leqslant \frac{c}{\varepsilon}\int_{\bigcup_{i=1}^{\infty} B(x_i,r_i)} |f(y)|^p\,dy \xrightarrow{\delta\to 0} 0,$$

which implies that $\operatorname{cap}_p(E_\varepsilon) = 0$ for every $\varepsilon > 0$. □

Now we are ready for a version of the Lebesgue differentiation theorem for Sobolev functions.

Theorem 4.31. Assume that $u \in W^{1,p}(\mathbb{R}^n)$ with $1 < p < \infty$. Then there exists $E \subset \mathbb{R}^n$ such that $\operatorname{cap}_p(E) = 0$ and

$$\lim_{r\to 0} \fint_{B(x,r)} u(y)\,dy = u^*(x)$$

exists for every $x \in \mathbb{R}^n \setminus E$. Moreover

$$\lim_{r\to 0} \fint_{B(x,r)} |u(y) - u^*(x)|\,dy = 0$$

for every $x \in \mathbb{R}^n \setminus E$ and the function u^* is the p-quasicontinuous representative of u.

T H E M O R A L : A function in $W^{1,p}(\mathbb{R}^n)$ with $1 < p < \infty$ has Lebesgue points p-quasieverywhere. Moreover, the p-quasicontinuous representative is obtained as a limit of integral averages.

Proof. (1) By Theorem 1.21 there exist $u_i \in C^\infty(\mathbb{R}^n) \cap W^{1,p}(\mathbb{R}^n)$ such that

$$\|u - u_i\|_{W^{1,p}(\mathbb{R}^n)}^p \leqslant 2^{-i(p+1)}, \quad i = 1,2,\ldots.$$

Denote

$$E_i = \{x \in \mathbb{R}^n : M(u-u_i)(x) > 2^{-i}\}, \quad i = 1,2,\dots.$$

By Theorem 4.29 there exists $c = c(n,p)$ such that

$$\operatorname{cap}_p(E_i) \leqslant c2^{ip} \|u-u_i\|^p_{W^{1,p}(\mathbb{R}^n)} \leqslant c2^{-i}, \quad i = 1,2,\dots.$$

Clearly

$$\begin{aligned} |u_i(x) - u_{B(x,r)}| &\leqslant \fint_{B(x,r)} |u_i(x) - u(y)|\,dy \\ &\leqslant \fint_{B(x,r)} |u_i(x) - u_i(y)|\,dy + \fint_{B(x,r)} |u_i(y) - u(y)|\,dy, \end{aligned}$$

which implies that

$$\begin{aligned} &\limsup_{r\to 0} |u_i(x) - u_{B(x,r)}| \\ &\quad\leqslant \limsup_{r\to 0} \fint_{B(x,r)} |u_i(x) - u_i(y)|\,dy + \limsup_{r\to 0} \fint_{B(x,r)} |u_i(y) - u(y)|\,dy \\ &\quad\leqslant M(u_i - u)(x) \leqslant 2^{-i}, \end{aligned}$$

for every $x \in \mathbb{R}^n \setminus E_i$. Here we used the fact that

$$\limsup_{r\to 0} \fint_{B(x,r)} |u_i(x) - u_i(y)|\,dy = 0, \quad i = 1,2,\dots,$$

since u_i is continuous and

$$\fint_{B(x,r)} |u_i(y) - u(y)|\,dy \leqslant M(u_i - u)(x) \quad \text{for every} \quad r > 0.$$

Let $F_k = \bigcup_{i=k}^{\infty} E_i$, $k = 1,2,\dots$. Then by the subadditivity of capacity we have

$$\operatorname{cap}_p(F_k) \leqslant \sum_{i=k}^{\infty} \operatorname{cap}_p(E_i) \leqslant c \sum_{i=k}^{\infty} 2^{-i}.$$

If $x \in \mathbb{R}^n \setminus F_k$ and $i, j \geqslant k$, then

$$\begin{aligned} |u_i(x) - u_j(x)| &\leqslant \limsup_{r\to 0} |u_i(x) - u_{B(x,r)}| + \limsup_{r\to 0} |u_{B(x,r)} - u_j(x)| \\ &\leqslant 2^{-i} + 2^{-j}. \end{aligned}$$

Thus (u_i) converges uniformly in $\mathbb{R}^n \setminus F_k$ to a continuous function v in $\mathbb{R}^n \setminus F_k$. Furthermore

$$\begin{aligned} \limsup_{r\to 0} |v(x) - u_{B(x,r)}| &\leqslant |v(x) - u_i(x)| + \limsup_{r\to 0} |u_i(x) - u_{B(x,r)}| \\ &\leqslant |v(x) - u_i(x)| + 2^{-i} \end{aligned}$$

for every $x \in \mathbb{R}^n \setminus F_k$. The right-hand side of the previous inequality tends to zero as $i \to \infty$. Thus

$$\limsup_{r\to 0} |v(x) - u_{B(x,r)}| = 0$$

and consequently

$$v(x) = \lim_{r\to 0} \fint_{B(x,r)} u(y)\,dy = u^*(x)$$

for every $x \in \mathbb{R}^n \setminus F_k$. Define $F = \bigcap_{k=1}^{\infty} F_k$. Then

$$\operatorname{cap}_p(F) \leqslant \lim_{k\to\infty} \operatorname{cap}_p(F_k) \leqslant c \lim_{k\to\infty} \sum_{i=k}^{\infty} 2^{-i} = 0$$

and

$$\lim_{r\to 0} \fint_{B(x,r)} u(y)\,dy = u^*(x)$$

exists for every $x \in \mathbb{R}^n \setminus F$. This completes the proof of the first claim.

(2) To prove the second claim, consider

$$E = \left\{ x \in \mathbb{R}^n : \limsup_{r\to 0} r^p \fint_{B(x,r)} |Du(y)|^p\,dy > 0 \right\}.$$

Lemma 4.30 shows that $\operatorname{cap}_p(E) = 0$. By the Poincaré inequality, see Theorem 5.25, we have

$$\lim_{r\to 0} \fint_{B(x,r)} |u(y) - u_{B(x,r)}|^p\,dy \leqslant c \lim_{r\to 0} r^p \fint_{B(x,r)} |Du(y)|^p\,dy = 0$$

for every $x \in \mathbb{R}^n \setminus E$. We conclude that

$$\begin{aligned} &\lim_{r\to 0} \fint_{B(x,r)} |u(y) - u^*(x)|\,dy \\ &\quad \leqslant \lim_{r\to 0} \left(\fint_{B(x,r)} |u(y) - u^*(x)|^p\,dy \right)^{\frac{1}{p}} \\ &\quad \leqslant \lim_{r\to 0} \left(\fint_{B(x,r)} |u(y) - u_{B(x,r)}|^p\,dy \right)^{\frac{1}{p}} + \lim_{r\to 0} |u_{B(x,r)} - u^*(x)| = 0 \end{aligned}$$

whenever $x \in \mathbb{R}^n \setminus (E \cup F)$. Finally we observe that

$$\operatorname{cap}_p(E \cup F) \leqslant \operatorname{cap}_p(E) + \operatorname{cap}_p(F) = 0.$$

(3) Let $\varepsilon > 0$ and choose k large enough so that $\operatorname{cap}_p(F_k) < \frac{\varepsilon}{2}$. Then by the outer regularity of capacity, see Remark 4.5, there is an open set O containing F_k so that $\operatorname{cap}_p(O) < \varepsilon$. Since (u_i) converges uniformly to u^* on $\mathbb{R}^n \setminus O$ we conclude that $u^*|_{X\setminus O}$ is continuous. Thus u^* is p-quasicontinuous. □

Remarks 4.32:

(1) Since the claims in Theorem 4.31 are local, the corresponding result also holds for $u \in W^{1,p}(\Omega)$, where $\Omega \subset \mathbb{R}^n$ is open.

(2) Let $n < p < \infty$. Lemma 4.14 implies that $\operatorname{cap}_p(\{x\}) > 0$ for every $x \in \mathbb{R}^n$, and Theorem 4.31 holds for every $u \in W^{1,p}(\mathbb{R}^n)$ with $E = \emptyset$. Consequently, every function $u \in W^{1,p}(\mathbb{R}^n)$ has a continuous representative and every point $x \in \mathbb{R}^n$ is a Lebesgue point of u. Recall that this fact follows more directly from Morrey's inequality, see Theorem 3.23, which is also applied in the proof of Lemma 4.14.

4.5 Sobolev spaces with zero boundary values

In this section we return to Sobolev spaces with zero boundary values started in Section 1.9. Assume that Ω is an open subset of $\mathbb{R}^n$ and $1 \leqslant p < \infty$. Recall that $W_0^{1,p}(\Omega)$ with $1 \leqslant p < \infty$ is the closure of $C_0^\infty(\Omega)$ with respect to the Sobolev norm, see Definition 1.23. Using pointwise properties of Sobolev functions we discuss the definition of $W_0^{1,p}(\Omega)$.

The first result is a $W_0^{1,p}(\Omega)$ version of Corollary 4.25 which states that for every $u \in W^{1,p}(\mathbb{R}^n)$ there is a p-quasicontinuous function $v \in W^{1,p}(\mathbb{R}^n)$ such that $u = v$ almost everywhere in $\mathbb{R}^n$.

Theorem 4.33. If $u \in W_0^{1,p}(\Omega)$, there exists a p-quasicontinuous function $v \in W^{1,p}(\mathbb{R}^n)$ such that $u = v$ almost everywhere in Ω and $v = 0$ p-quasieverywhere in $\mathbb{R}^n \setminus \Omega$.

THE MORAL: Quasicontinuous functions in Sobolev spaces with zero boundary values are zero quasieverywhere in the complement.

Proof. Since $u \in W_0^{1,p}(\Omega)$, there exist $u_i \in C_0^\infty(\Omega)$, $i = 1, 2, \dots$, such that $u_i \to u$ in $W^{1,p}(\Omega)$ as $i \to \infty$. Since (u_i) is a Cauchy sequence in $W^{1,p}(\mathbb{R}^n)$, by Theorem 4.22 it has a subsequence of (u_i) that converges pointwise p-quasieverywhere in $\mathbb{R}^n$ to a function $v \in W^{1,p}(\mathbb{R}^n)$. Moreover, the convergence is uniform outside a set of arbitrary small p-capacity and, as in Corollary 4.25, the limit function v is p-quasicontinuous. □

Theorem 4.34. If $u \in W^{1,p}(\mathbb{R}^n)$ is p-quasicontinuous and $u = 0$ p-quasieverywhere in $\mathbb{R}^n \setminus \Omega$, then $u \in W_0^{1,p}(\Omega)$.

THE MORAL: Quasicontinuous functions in a Sobolev space on the whole space which are zero quasieverywhere in the complement belong to the Sobolev space with zero boundary values. In particular, continuous functions in a Sobolev space on the whole space which are zero everywhere in the complement belong to the Sobolev space with zero boundary values.

Proof. (1) We show that u can be approximated by $W^{1,p}(\mathbb{R}^n)$ functions with compact support in Ω. If we can construct such a sequence for $u_+ = \max\{u, 0\}$, then we can do it for $u_- = -\min\{u, 0\}$, and we obtain the result for $u = u_+ - u_-$. Thus we may assume that $u \geqslant 0$. By Theorem 1.28 we may assume that u has a compact support in $\mathbb{R}^n$ and by considering truncations $\min\{u, \lambda\}$, $\lambda > 0$, we may assume that u is bounded (exercise).

(2) Let $\delta > 0$ and let $O \subset \mathbb{R}^n$ be an open set such that $\operatorname{cap}_p(O) < \delta$ and the restriction of u to $\mathbb{R}^n \setminus O$ is continuous. Denote

$$E = \{x \in \mathbb{R}^n \setminus \Omega : u(x) \neq 0\}.$$

By assumption $\operatorname{cap}_p(E) = 0$. Let $v \in \mathscr{A}'(O \cup E)$ such that $0 \leqslant v \leqslant 1$ and

$$\|v\|^p_{W^{1,p}(\mathbb{R}^n)} < \delta,$$

see Remark 4.2. Then $v = 1$ in an open set G containing $O \cup E$. Define

$$u_\varepsilon(x) = \max\{u(x) - \varepsilon, 0\}, \quad 0 < \varepsilon < 1.$$

Let $x \in \partial\Omega \setminus G$. Since $u(x) = 0$ and the restriction of u to $\mathbb{R}^n \setminus G$ is continuous, there exists $r_x > 0$ such that $u_\varepsilon = 0$ in $B(x, r_x) \setminus G$. Thus $(1 - v)u_\varepsilon = 0$ in $B(x, r_x) \cup G$ for every $x \in \partial\Omega \setminus G$. This shows that $(1 - v)u_\varepsilon$ is zero in a neighbourhood of $\mathbb{R}^n \setminus \Omega$, which implies that $(1 - v)u_\varepsilon$ is compactly supported in Ω. Lemma 1.27 implies $(1 - v)u_\varepsilon \in W^{1,p}_0(\Omega)$. We show that this kind of functions converge to u in $W^{1,p}(\mathbb{R}^n)$.

(3) Since

$$u_\varepsilon = \begin{cases} u - \varepsilon & \text{in} \quad \{x \in \mathbb{R}^n : u(x) \geqslant \varepsilon\}, \\ 0 & \text{in} \quad \{x \in \mathbb{R}^n : u(x) \leqslant \varepsilon\}, \end{cases}$$

by Remark 2.4 we have

$$Du_\varepsilon = \begin{cases} Du & \text{a.e. in} \quad \{x \in \mathbb{R}^n : u(x) \geqslant \varepsilon\}, \\ 0 & \text{a.e. in} \quad \{x \in \mathbb{R}^n : u(x) \leqslant \varepsilon\}. \end{cases}$$

Thus

$$\|u - (1 - v)u_\varepsilon\|_{W^{1,p}(\mathbb{R}^n)} \leqslant \|u - u_\varepsilon\|_{W^{1,p}(\mathbb{R}^n)} + \|vu_\varepsilon\|_{W^{1,p}(\mathbb{R}^n)}.$$

Using the facts that $u - u_\varepsilon \leqslant \varepsilon$ and $\operatorname{supp}(u - u_\varepsilon) \subset \operatorname{supp} u$, we obtain

$$\begin{aligned} \|u - u_\varepsilon\|_{W^{1,p}(\mathbb{R}^n)} &\leqslant \|u - u_\varepsilon\|_{L^p(\mathbb{R}^n)} + \|Du - Du_\varepsilon\|_{L^p(\mathbb{R}^n)} \\ &\leqslant \varepsilon \|\chi_{\operatorname{supp} u}\|_{L^p(\mathbb{R}^n)} + \|\chi_{\{0 < u \leqslant \varepsilon\}} Du\|_{L^p(\mathbb{R}^n)} \to 0 \end{aligned}$$

as $\varepsilon \to 0$. Observe that, by the dominated convergence theorem, we have

$$\begin{aligned} \lim_{\varepsilon \to 0} \|\chi_{\{0 < u \leqslant \varepsilon\}} Du\|_{L^p(\mathbb{R}^n)} &= \left(\lim_{\varepsilon \to 0} \int_{\mathbb{R}^n} \chi_{\{0 < u \leqslant \varepsilon\}} |Du|^p \, dx \right)^{\frac{1}{p}} \\ &= \left(\int_{\mathbb{R}^n} \lim_{\varepsilon \to 0} \chi_{\{0 < u \leqslant \varepsilon\}} |Du|^p \, dx \right)^{\frac{1}{p}} = 0, \end{aligned}$$

where $\chi_{\{0 < u \leqslant \varepsilon\}} |Du|^p \leqslant |Du|^p \in L^1(\mathbb{R}^n)$ may be used as an integrable majorant. On the other hand,

$$\begin{aligned}
\|vu_\varepsilon\|_{W^{1,p}(\mathbb{R}^n)} &\leqslant \|vu_\varepsilon\|_{L^p(\mathbb{R}^n)} + \|D(vu_\varepsilon)\|_{L^p(\mathbb{R}^n)} \\
&\leqslant \|vu_\varepsilon\|_{L^p(\mathbb{R}^n)} + \|u_\varepsilon Dv\|_{L^p(\mathbb{R}^n)} + \|vDu_\varepsilon\|_{L^p(\mathbb{R}^n)} \\
&\leqslant \|uv\|_{L^p(\mathbb{R}^n)} + \|uDv\|_{L^p(\mathbb{R}^n)} + \|vDu_\varepsilon\|_{L^p(\mathbb{R}^n)} \\
&\leqslant \|u\|_{L^\infty(\mathbb{R}^n)}\|v\|_{L^p(\mathbb{R}^n)} + \|u\|_{L^\infty(\mathbb{R}^n)}\|Dv\|_{L^p(\mathbb{R}^n)} + \|vDu\|_{L^p(\mathbb{R}^n)} \\
&\leqslant 2\|u\|_{L^\infty(\mathbb{R}^n)}\|v\|_{W^{1,p}(\mathbb{R}^n)} + \|vDu\|_{L^p(\mathbb{R}^n)} \\
&\leqslant 2\delta^{\frac{1}{p}}\|u\|_{L^\infty(\mathbb{R}^n)} + \|vDu\|_{L^p(\mathbb{R}^n)}.
\end{aligned}$$

Since $v = v_\delta \to 0$ in $L^p(\mathbb{R}^n)$ as $\delta \to 0$, there is a subsequence (δ_i) for which $v_i = v_{\delta_i} \to 0$ almost everywhere as $i \to \infty$. By the dominated convergence theorem, we have

$$\begin{aligned}
\lim_{i\to\infty} \|v_i Du\|_{L^p(\mathbb{R}^n)} &= \left(\lim_{i\to\infty} \int_{\mathbb{R}^n} |v_i|^p |Du|^p \, dx\right)^{\frac{1}{p}} \\
&= \left(\int_{\mathbb{R}^n} (\lim_{i\to\infty} |v_i|^p)|Du|^p \, dx\right)^{\frac{1}{p}} = 0,
\end{aligned}$$

where $|v_i|^p|Du|^p \leqslant |Du|^p$, so that $|Du|^p \in L^1(\mathbb{R}^n)$ may be used as an integrable majorant. Thus we conclude that

$$\lim_{i\to\infty} \|v_i u_\varepsilon\|_{W^{1,p}(\mathbb{R}^n)} \leqslant \lim_{i\to\infty} \left(2\delta_i^{\frac{1}{p}} \|u\|_{L^\infty(\mathbb{R}^n)} + \|v_i Du\|_{L^p(\mathbb{R}^n)}\right) = 0.$$

Thus

$$\|u - (1 - v_i)u_\varepsilon\|_{W^{1,p}(\mathbb{R}^n)} \to 0$$

as $\varepsilon \to 0$ and $i \to \infty$. Since

$$(1 - v_i)u_\varepsilon \in W_0^{1,p}(\Omega) \quad \text{and} \quad (1 - v_i)u_\varepsilon \to u \quad \text{in} \quad W^{1,p}(\mathbb{R}^n)$$

as $\varepsilon \to 0$ and $i \to \infty$, we conclude that $u \in W_0^{1,p}(\Omega)$. □

We obtain a very useful characterization of Sobolev spaces with zero boundary values on an arbitrary open set by combining the last two theorems.

Corollary 4.35. $u \in W_0^{1,p}(\Omega)$ if and only if there exists a p-quasicontinuous function $u^* \in W^{1,p}(\mathbb{R}^n)$ such that $u^* = u$ almost everywhere in Ω and $u^* = 0$ p-quasieverywhere in $\mathbb{R}^n \setminus \Omega$.

THE MORAL: Quasicontinuous functions in Sobolev spaces with zero boundary values are precisely functions in the Sobolev space on the whole space which are zero quasieverywhere in the complement. This result can be used to show that a given function belongs to the Sobolev space with zero boundary values without constructing an approximating sequence of compactly supported smooth functions.

Remark 4.36. Let $p > n$. Lemma 4.14 implies that empty set is the only set of p-capacity zero. Thus a function is p-quasicontinuous if and only if it is continuous in $\mathbb{R}^n$. Corollary 4.35 asserts that $u \in W_0^{1,p}(\Omega)$ with $p > n$ if and only if there exists a continuous function $u^* \in W^{1,p}(\mathbb{R}^n)$ such that $u^* = u$ almost everywhere in Ω and $u^* = 0$ everywhere in $\mathbb{R}^n \setminus \Omega$.

There is also a characterization of Sobolev spaces with zero boundary values using Lebesgue points for Sobolev functions.

Theorem 4.37. Assume that $\Omega \subset \mathbb{R}^n$ is an open set and $u \in W^{1,p}(\mathbb{R}^n)$ with $1 < p < \infty$. Then $u \in W_0^{1,p}(\Omega)$ if and only if

$$\lim_{r\to 0} \fint_{B(x,r)} u(y)\,dy = 0$$

for p-quasievery $x \in \mathbb{R}^n \setminus \Omega$.

THE MORAL: A function in the Sobolev space on the whole space belongs to the Sobolev space with zero boundary values if and only if the limit of integral averages is zero quasieverywhere in the complement.

Proof. $\boxed{\Longrightarrow}$ If $u \in W_0^{1,p}(\Omega)$, then by Theorem 4.33 there exists a p-quasicontinuous function $u^* \in W^{1,p}(\mathbb{R}^n)$ such that $u^* = u$ almost everywhere in Ω and $u^* = 0$ p-quasieverywhere in $\mathbb{R}^n \setminus \Omega$. Theorem 4.31 shows that the limit

$$u^*(x) = \lim_{r\to 0} \fint_{B(x,r)} u(y)\,dy$$

exists p-quasieverywhere and that the function u^* is a p-quasicontinuous representative of u. This shows that

$$\lim_{r\to 0} \fint_{B(x,r)} u(y)\,dy = u^*(x) = 0$$

for p-quasievery $x \in \mathbb{R}^n \setminus \Omega$.

$\boxed{\Longleftarrow}$ Assume then that $u \in W^{1,p}(\mathbb{R}^n)$ and

$$\lim_{r\to 0} \fint_{B(x,r)} u(y)\,dy = 0$$

for p-quasievery $x \in \mathbb{R}^n \setminus \Omega$. Theorem 4.31 shows that the limit

$$u^*(x) = \lim_{r\to 0} \fint_{B(x,r)} u(y)\,dy$$

exists p-quasieverywhere and that the function u^* is a p-quasicontinuous representative of u. We conclude that $u^*(x) = 0$ for p-quasievery $x \in \mathbb{R}^n \setminus \Omega$. The claim follows from Corollary 4.35. $\square$

Let $E \subset \Omega$ be a relatively closed set, that is, there exists a closed $F \subset \mathbb{R}^n$ such that $E = \Omega \cap F$, with $|E| = 0$. It is clear that $W_0^{1,p}(\Omega \setminus E) \subset W_0^{1,p}(\Omega)$. By

$$W_0^{1,p}(\Omega \setminus E) = W_0^{1,p}(\Omega)$$

we mean that every $u \in W_0^{1,p}(\Omega)$ can be approximated by functions in $C_0^\infty(\Omega \setminus E)$ or in $W_0^{1,p}(\Omega \setminus E)$.

Theorem 4.38. Assume that E is a relatively closed subset of Ω. Then $W_0^{1,p}(\Omega) = W_0^{1,p}(\Omega \setminus E)$ if and only if $\operatorname{cap}_p(E) = 0$.

Proof. $\boxed{\Leftarrow}$ Assume $\operatorname{cap}_p(E) = 0$. Lemma 4.9 implies $|E| = 0$ so that it is reasonable to ask whether $W_0^{1,p}(\Omega) = W_0^{1,p}(\Omega \setminus E)$ when we consider functions defined up to a set of measure zero.

It is clear that $W_0^{1,p}(\Omega \setminus E) \subset W_0^{1,p}(\Omega)$. To see reverse inclusion, let $u_i \in C_0^\infty(\Omega)$, $i = 1,2,\dots$, be such that $u_i \to u$ in $W^{1,p}(\Omega)$ as $i \to \infty$. Since $\operatorname{cap}_p(E) = 0$ there are $v_j \in \mathscr{A}'(E)$, $j = 1,2,\dots$, be such that $\|v_j\|_{W^{1,p}(\mathbb{R}^n)} \to 0$ as $j \to \infty$. Then $(1 - v_j)u_i \in W^{1,p}(\Omega)$ and, since $v_j = 1$ in a neighbourhood of E, $\operatorname{supp}(1 - v_j)u_i$ is a compact subset of $\Omega \setminus E$ for every $i, j = 1,2,\dots$. Lemma 1.27 implies $(1 - v_j)u_i \in W_0^{1,p}(\Omega \setminus E)$, $i, j = 1,2,\dots$.

Moreover, we have

$$\|u - (1 - v_j)u_i\|_{W^{1,p}(\Omega)} \leqslant \|u - u_i\|_{W^{1,p}(\Omega)} + \|v_j u_i\|_{W^{1,p}(\Omega)},$$

where $\|u - u_i\|_{W^{1,p}(\Omega)} \to 0$ as $i \to \infty$ and

$$\begin{aligned}
\|v_j u_i\|_{W^{1,p}(\Omega)} &\leqslant \|v_j u_i\|_{L^p(\Omega)} + \|D(v_j u_i)\|_{L^p(\Omega)} \\
&\leqslant \|u_i\|_{L^\infty(\Omega)} \|v_j\|_{L^p(\Omega)} + \|v_j D u_i\|_{L^p(\Omega)} + \|u_i D v_j\|_{L^p(\Omega)} \\
&\leqslant \|u_i\|_{L^\infty(\Omega)} \|v_j\|_{L^p(\Omega)} + \|v_j D u_i\|_{L^p(\Omega)} + \|u_i\|_{L^\infty(\Omega)} \|D v_j\|_{L^p(\Omega)} \\
&\leqslant 2\|u_i\|_{L^\infty(\Omega)} \|v_j\|_{W^{1,p}(\Omega)} + \|v_j D u_i\|_{L^p(\Omega)}.
\end{aligned}$$

Since $v_j \to 0$ in $L^p(\Omega)$ as $j \to \infty$, there is a subsequence, still denoted by (v_j), for which $v_j \to 0$ almost everywhere as $j \to \infty$. By the dominated convergence theorem, we have

$$\begin{aligned}
\lim_{j \to \infty} \|v_j D u_i\|_{L^p(\Omega)} &= \left(\lim_{j \to \infty} \int_\Omega |v_j|^p |D u_i|^p \, dx \right)^{\frac{1}{p}} \\
&= \left(\int_\Omega (\lim_{j \to \infty} |v_j|^p) |D u_i|^p \, dx \right)^{\frac{1}{p}} = 0.
\end{aligned}$$

Observe that $|v_j|^p |D u_i|^p \leqslant |D u_i|^p$ for $j = 1,2,\dots$, so that $|D u_i|^p \in L^1(\Omega)$ may be used as an integrable majorant. Thus

$$\|u - (1 - v_j)u_i\|_{W^{1,p}(\Omega)} \to 0 \quad \text{as} \quad i, j \to \infty.$$

Since

$$(1-v_j)u_i \in W_0^{1,p}(\Omega\setminus E) \quad \text{and} \quad (1-v_j)u_i \to u \quad \text{in} \quad W^{1,p}(\Omega\setminus E)$$

as $i, j \to \infty$, we conclude that $u \in W_0^{1,p}(\Omega \setminus E)$.

$\boxed{\Longrightarrow}$ Let $x_0 \in \Omega$ and let $i_0 \in \mathbb{N}$ be large enough that

$$\operatorname{dist}(x_0, \mathbb{R}^n \setminus \Omega) > \tfrac{1}{i_0}.$$

Define

$$\Omega_i = \left\{x \in \Omega : \operatorname{dist}(x, \mathbb{R}^n \setminus \Omega) > \tfrac{1}{i}\right\} \cap B(x_0, i), \quad i = i_0, i_0+1, \dots.$$

Observe that $\Omega_i \Subset \Omega_{i+1} \Subset \dots \Subset \Omega$ and $\Omega = \bigcup_{i=i_0}^\infty \Omega_i$. Let $u_i : \mathbb{R}^n \to \mathbb{R}$,

$$u_i(x) = \operatorname{dist}(x, \mathbb{R}^n \setminus \Omega_{2i}).$$

Then u_i is Lipschitz continuous, $u_i \in W_0^{1,p}(\Omega)$ and $u_i(x) \geqslant \frac{1}{2i}$ for every $x \in E \cap \Omega_i$, $i = 1, 2, \dots$. Since $W_0^{1,p}(\Omega) = W_0^{1,p}(\Omega \setminus E)$ we have $u_i \in W_0^{1,p}(\Omega \setminus E)$, $i = 1, 2, \dots$.

Fix i and let $v_j \in C_0^\infty(\Omega \setminus E)$, $j = 1, 2, \dots$, such that $v_j \to u_i$ in $W^{1,p}(\Omega \setminus E)$ as $j \to \infty$. Since $3i(u_i - v_j) \geqslant 1$ in a neigbourhood of $E \cap \Omega_i$,

$$\begin{aligned}\operatorname{cap}_p(E\cap\Omega_i) &\leqslant \|3i(u_i - v_j)\|^p_{W^{1,p}(\Omega\setminus E)} \\ &= (3i)^p \|u_i - v_j\|^p_{W^{1,p}(\Omega\setminus E)} \xrightarrow{j\to\infty} 0.\end{aligned}$$

Thus $\operatorname{cap}_p(E \cap \Omega_i) = 0$, $i = 1, 2 \dots$, and by subadditivity

$$\operatorname{cap}_p(E) = \operatorname{cap}_p\Big(\bigcup_{i=1}^\infty (E\cap\Omega_i)\Big) \leqslant \sum_{i=1}^\infty \operatorname{cap}_p(E\cap\Omega_i) = 0.$$

□

Example 4.39. Let $\Omega = B(0,1) \setminus \{0\}$. By Remark 4.12 (2) and Remark 4.13 we have $\operatorname{cap}_p(\{0\}) = 0$ when $1 < p \leqslant n$. On the other hand, by Lemma 4.14, we have $\operatorname{cap}_p(\{0\}) > 0$ for every $x \in \mathbb{R}^n$ when $p > n$. Thus

$$W_0^{1,p}(B(0,1)) = W_0^{1,p}(B(0,1)\setminus\{0\})$$

if and only if $1 < p \leqslant n$. Thus

$$W_0^{1,p}(B(0,1)) = W_0^{1,p}(B(0,1)\setminus\{0\}), \quad 1 < p \leqslant n,$$

and

$$W_0^{1,p}(B(0,1)) \neq W_0^{1,p}(B(0,1)\setminus\{0\}), \quad p > n.$$

T H E M O R A L : The Sobolev space with zero boundary values $W_0^{1,p}(B(0,1)\setminus\{0\})$ does not see the boundary point $\{0\}$ when $1 < p \leqslant n$.

Example 4.40. Let $\Omega = B(0,1) \setminus \{0\}$ and $u : \Omega \to \mathbb{R}$, $u(x) = 1 - |x|$. Then $u \in W_0^{1,p}(\Omega)$ for $1 < p \leqslant n$ and $u \notin W_0^{1,p}(\Omega)$ for $p > n$.

Reason. $\boxed{1 < p \leqslant n}$ The zero extension $u_0 : \mathbb{R}^n \to \mathbb{R}$,

$$u_0(x) = \begin{cases} 1 - |x|, & x \in B(0,1), \\ 0, & x \in \mathbb{R}^n \setminus B(0,1), \end{cases}$$

is continuous in $\mathbb{R}^n$ and thus p-quasicontinuous in $\mathbb{R}^n$. Moreover, we have $u_0 \in W^{1,p}(\mathbb{R}^n)$. By Remark 4.12 (2) and Remark 4.13 we have $\operatorname{cap}_p(\{0\}) = 0$ when $1 < p \leqslant n$. It follows that $u_0 = 0$ p-quasieverywhere in $\mathbb{R}^n \setminus \Omega$. Since $u_0 = u$ in Ω, Corollary 4.35 implies that $u \in W_0^{1,p}(\Omega)$.

$\boxed{p > n}$ By Remark 4.36 we may conclude that that $u \in W_0^{1,p}(\Omega)$ with $p > n$ if and only if there exists a continuous function $u^* \in W^{1,p}(\mathbb{R}^n)$ such that $u^* = u$ almost everywhere in Ω and $u^* = 0$ everywhere in $\mathbb{R}^n \setminus \Omega$.

For a contradiction, assume that there exists a continuous function $u^* \in W^{1,p}(\mathbb{R}^n)$ such that $u^* = u$ almost everywhere in Ω and $u^* = 0$ everywhere in $\mathbb{R}^n \setminus \Omega$. In particular, $u^*(0) = 0$. This is a contradiction with the assumption that $u^* = u$ almost everywhere in Ω.

Second approach: Since $u_0 \in W^{1,p}(\mathbb{R}^n)$, by Corollary 4.37, we have $u \in W_0^{1,p}(\Omega)$ if and only if

$$\lim_{r \to 0} \fint_{B(x,r)} u_0(y)\, dy = 0$$

for every $x \in \mathbb{R}^n \setminus \Omega$. However, a direct computation shows that

$$\lim_{r \to 0} \fint_{B(0,r)} u_0(y)\, dy = 1.$$

It follows that $u \notin W_0^{1,p}(\Omega)$. ∎

T H E M O R A L : A function that belongs to the Sobolev space with zero boundary values does not have to be zero at every point of the boundary.

5 Maximal function approach to Sobolev spaces

5.1 Maximal operator on Sobolev spaces

We recall the definition of the maximal function.

Definition 5.1. The centered Hardy-Littlewood maximal function $Mf : \mathbb{R}^n \to [0,\infty]$ of $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ is

$$Mf(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)|\,dy,$$

where $B(x,r) = \{y \in \mathbb{R}^n : |y-x| < r\}$ is the open ball with the radius $r > 0$ and the center $x \in \mathbb{R}^n$.

THE MORAL: The maximal function gives the maximal integral average of the absolute value of the function on balls centered at a point.

Note that the Lebesgue differentiation theorem implies

$$\begin{aligned}|f(x)| &= \lim_{r\to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)|\,dy \\ &\leqslant \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)|\,dy = Mf(x)\end{aligned}$$

for almost every $x \in \mathbb{R}^n$.

We are interested in behaviour of the maximal operator in L^p-spaces and begin with a relatively obvious result.

Lemma 5.2. If $f \in L^\infty(\mathbb{R}^n)$, then $Mf \in L^\infty(\mathbb{R}^n)$ and $\|Mf\|_{L^\infty(\mathbb{R}^n)} \leqslant \|f\|_{L^\infty(\mathbb{R}^n)}$.

THE MORAL: If the original function is essentially bounded, then the maximal function is essentially bounded and thus finite almost everywhere. Intuitively this is clear, since the integral averages cannot be larger than the essential supremum of the function. Another way to state this is that $M : L^\infty(\mathbb{R}^n) \to L^\infty(\mathbb{R}^n)$ is a bounded operator.

Proof. For every $x \in \mathbb{R}^n$ and $r > 0$ we have

$$\frac{1}{|B(x,r)|}\int_{B(x,r)} |f(y)|\,dy \leqslant \frac{1}{|B(x,r)|}\|f\|_{L^\infty(\mathbb{R}^n)}|B(x,r)| = \|f\|_{L^\infty(\mathbb{R}^n)}.$$

By taking supremum over $r > 0$, we have $Mf(x) \leqslant \|f\|_{L^\infty(\mathbb{R}^n)}$ for every $x \in \mathbb{R}^n$ and thus $\|Mf\|_{L^\infty(\mathbb{R}^n)} \leqslant \|f\|_{L^\infty(\mathbb{R}^n)}$. □

The following maximal function theorem was first proved by Hardy and Littlewood in the one-dimensional case and by Wiener in higher dimensions.

Theorem 5.3 (Hardy-Littlewood-Wiener).

(1) If $f \in L^1(\mathbb{R}^n)$, there exists a constant $c = c(n)$ such that

$$|\{x \in \mathbb{R}^n : Mf(x) > \lambda\}| \leqslant \frac{c}{\lambda}\|f\|_{L^1(\mathbb{R}^n)} \quad \text{for every} \quad \lambda > 0.$$

(2) If $f \in L^p(\mathbb{R}^n)$, $1 < p \leqslant \infty$, then $Mf \in L^p(\mathbb{R}^n)$ and there exists a constant $c = c(n,p)$ such that

$$\|Mf\|_{L^p(\mathbb{R}^n)} \leqslant c\|f\|_{L^p(\mathbb{R}^n)}.$$

THE MORAL: The first assertion states that the Hardy-Littlewood maximal operator maps $L^1(\mathbb{R}^n)$ to weak $L^1(\mathbb{R}^n)$ and the second claim shows that $M : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$ is a bounded operator for $p > 1$.

WARNING: $f \in L^1(\mathbb{R}^n)$ does not imply that $Mf \in L^1(\mathbb{R}^n)$ and thus the Hardy-Littlewood maximal operator is not bounded in $L^1(\mathbb{R}^n)$. In this case we only have the weak type estimate.

Assume that u is Lipschitz continuous with constant L, that is

$$|u_h(y) - u(y)| = |u(y+h) - u(y)| \leqslant L|h|$$

for every $y, h \in \mathbb{R}^n$, where we denote $u_h(y) = u(y+h)$. Since the maximal function commutes with translations and the maximal operator is sublinear, we have

$$\begin{aligned}
|(Mu)_h(x) - Mu(x)| &= |M(u_h)(x) - Mu(x)| \\
&\leqslant M(u_h - u)(x) \\
&= \sup_{r>0} \frac{1}{|B(x,r)|}\int_{B(x,r)} |u_h(y) - u(y)|\,dy \\
&\leqslant L|h|.
\end{aligned}$$

This means that the maximal function is Lipschitz continuous with the same constant as the original function if Mu is not identically infinity. Observe, that this proof applies to Hölder continuous functions as well.

Next we show that the Hardy-Littlewood maximal operator is bounded in Sobolev spaces.

Theorem 5.4. Let $1 < p < \infty$. If $u \in W^{1,p}(\mathbb{R}^n)$, then $Mu \in W^{1,p}(\mathbb{R}^n)$. Moreover, there exists $c = c(n,p)$ such that

$$\|Mu\|_{W^{1,p}(\mathbb{R}^n)} \leqslant c\|u\|_{W^{1,p}(\mathbb{R}^n)}. \tag{5.5}$$

T H E M O R A L : $M : W^{1,p}(\mathbb{R}^n) \to W^{1,p}(\mathbb{R}^n)$, $p > 1$, is a bounded operator. Thus the maximal operator is not only bounded on $L^p(\mathbb{R}^n)$ but also on $W^{1,p}(\mathbb{R}^n)$ for $p > 1$.

Proof. The proof is based on the characterization of $W^{1,p}(\mathbb{R}^n)$ by integrated difference quotients, see Theorem 2.32. By the maximal function theorem with $1 < p < \infty$, see Theorem 5.3 (2) , we have $Mu \in L^p(\mathbb{R}^n)$ and

$$\begin{aligned}
\|(Mu)_h - Mu\|_{L^p(\mathbb{R}^n)} &= \|M(u_h) - Mu\|_{L^p(\mathbb{R}^n)} \\
&\leqslant \|M(u_h - u)\|_{L^p(\mathbb{R}^n)} \\
&\leqslant c\|u_h - u\|_{L^p(\mathbb{R}^n)} \\
&\leqslant c\|Du\|_{L^p(\mathbb{R}^n)}|h|
\end{aligned}$$

for every $h \in \mathbb{R}^n$. Theorem 2.32 gives $Mu \in W^{1,p}(\mathbb{R}^n)$ with

$$\|DMu\|_{L^p(\mathbb{R}^n)} \leqslant c\|Du\|_{L^p(\mathbb{R}^n)}.$$

Thus by the maximal function theorem

$$\begin{aligned}
\|Mu\|_{W^{1,p}(\mathbb{R}^n)} &= \left(\|Mu\|^p_{L^p(\mathbb{R}^n)} + \|DMu\|^p_{L^p(\mathbb{R}^n)}\right)^{\frac{1}{p}} \\
&\leqslant \|Mu\|_{L^p(\mathbb{R}^n)} + \|DMu\|_{L^p(\mathbb{R}^n)} \\
&\leqslant c\left(\|u\|_{L^p(\mathbb{R}^n)} + \|Du\|_{L^p(\mathbb{R}^n)}\right) \\
&\leqslant c\|u\|_{W^{1,p}(\mathbb{R}^n)}.
\end{aligned}$$

□

A more careful analysis gives a pointwise estimate for the partial derivatives.

Theorem 5.6. Let $1 < p < \infty$. If $u \in W^{1,p}(\mathbb{R}^n)$, then $Mu \in W^{1,p}(\mathbb{R}^n)$ and

$$|D_j Mu| \leqslant M(D_j u), \quad j = 1, 2, \ldots, n, \tag{5.7}$$

almost everywhere in $\mathbb{R}^n$.

T H E M O R A L : Differentiation commutes with a linear operator. The sublinear maximal operator semicommutes with differentiation.

Proof. If $\chi_{B(0,r)}$ is the characteristic function of $B(0,r)$ and

$$\chi_r = \frac{\chi_{B(0,r)}}{|B(0,r)|},$$

then

$$\begin{aligned}\frac{1}{|B(x,r)|}\int_{B(x,r)} |u(y)|\,dy &= \frac{1}{|B(0,r)|}\int_{B(0,r)} |u(x-y)|\,dy \\ &= \frac{1}{|B(0,r)|}\int_{\mathbb{R}^n} \chi_{B(0,r)}|u(x-y)|\,dy \\ &= (|u| * \chi_r)(x),\end{aligned}$$

where $*$ denotes the convolution. We observe that $|u| * \chi_r \in W^{1,p}(\mathbb{R}^n)$ and

$$D_j(|u| * \chi_r) = \chi_r * D_j|u|, \quad j = 1,2,\dots,n,$$

almost everywhere in $\mathbb{R}^n$, see the proof of Theorem 1.19 (exercise).

Let r_m, $m = 1,2,\dots$, be an enumeration of positive rationals. Since u is locally integrable, we may restrict ourselves to the positive rational radii in the definition of the maximal function. Hence

$$Mu(x) = \sup_m (|u| * \chi_{r_m})(x).$$

We define functions $v_k : \mathbb{R}^n \to \mathbb{R}$, $k = 1,2,\dots$, by

$$v_k(x) = \max_{1\leqslant m\leqslant k} (|u| * \chi_{r_m})(x).$$

Now (v_k) is an increasing sequence of functions in $W^{1,p}(\mathbb{R}^n)$, which converges to Mu pointwise and

$$\begin{aligned}|D_j v_k| &\leqslant \max_{1\leqslant m\leqslant k} \left|D_j(|u| * \chi_{r_m})\right| \\ &= \max_{1\leqslant m\leqslant k} \left|\chi_{r_m} * D_j|u|\right| \\ &\leqslant M(D_j|u|) = M(D_j u), \quad j = 1,2,\dots,n,\end{aligned}$$

almost everywhere in $\mathbb{R}^n$. Here we also used Remark 2.4 and the fact that by Theorem 2.3

$$\left|D_j|u|\right| = |D_j u|, \quad j = 1,2,\dots,n,$$

almost everywhere. Thus

$$\|Dv_k\|_{L^p(\mathbb{R}^n)} \leqslant \sum_{j=1}^n \|D_j v_k\|_{L^p(\mathbb{R}^n)} \leqslant \sum_{j=1}^n \|M(D_j u)\|_{L^p(\mathbb{R}^n)}$$

and the maximal function theorem implies

$$\begin{aligned}\|v_k\|_{W^{1,p}(\mathbb{R}^n)} &\leqslant \|Mu\|_{L^p(\mathbb{R}^n)} + \sum_{j=1}^n \|M(D_j u)\|_{L^p(\mathbb{R}^n)} \\ &\leqslant c\|u\|_{L^p(\mathbb{R}^n)} + c\sum_{j=1}^n \|D_j u\|_{L^p(\mathbb{R}^n)} \leqslant c < \infty\end{aligned}$$

for every $k=1,2,\dots$ Hence (v_k) is a bounded sequence in $W^{1,p}(\mathbb{R}^n)$ which converges to Mu pointwise. Theorem 2.24 implies $Mu\in W^{1,p}(\mathbb{R}^n)$, $v_k\to Mu$ weakly in $L^p(\mathbb{R}^n)$ and $D_jv_k\to D_jMu$ weakly in $L^p(\mathbb{R}^n)$ as $k\to\infty$.

Next we prove the pointwise estimate for the gradient. By Mazur's lemma, see Theorem 2.17, there is a sequence of convex combinations such that

$$w_k=\sum_{l=k}^{m_k}a_{k,l}D_jv_l\to D_jMu,\quad j=1,\dots,n,$$

in $L^p(\mathbb{R}^n)$ as $k\to\infty$. There is a subsequence of (w_k) which converges almost everywhere to D_jMu. Thus we have

$$|w_k|\leqslant\sum_{l=k}^{m_k}a_{k,l}\left|D_jv_l\right|\leqslant\sum_{l=k}^{m_k}a_{k,l}M(D_ju)=M(D_ju)$$

for every $k=1,2,\dots$ and finally

$$|D_jMu|=\lim_{k\to\infty}|w_k|\leqslant M(D_ju),\quad j=1,\dots,n,$$

almost everywhere in $\mathbb{R}^n$. This completes the proof. □

Remarks 5.8:

(1) Estimate (5.5) also follows from (5.7). To see this, we may use the maximal function theorem, see Theorem 5.3 (2) , and (5.7) to obtain

$$\begin{aligned}\|Mu\|_{W^{1,p}(\mathbb{R}^n)}&\leqslant\|Mu\|_{L^p(\mathbb{R}^n)}+\|DMu\|_{L^p(\mathbb{R}^n)}\\&\leqslant c\|u\|_{L^p(\mathbb{R}^n)}+\|M|Du|\|_{L^p(\mathbb{R}^n)}\\&\leqslant c\|u\|_{W^{1,p}(\mathbb{R}^n)},\end{aligned}$$

where c is the constant in Theorem 5.3 (2) .

(2) If $u\in W^{1,\infty}(\mathbb{R}^n)$, then a slight modification of our proof shows that Mu belongs to $W^{1,\infty}(\mathbb{R}^n)$. Moreover,

$$\begin{aligned}\|Mu\|_{W^{1,\infty}(\mathbb{R}^n)}&=\|Mu\|_{L^\infty(\mathbb{R}^n)}+\|DMu\|_{L^\infty(\mathbb{R}^n)}\\&\leqslant\|u\|_{L^\infty(\mathbb{R}^n)}+\|M|Du|\|_{L^\infty(\mathbb{R}^n)}\\&\leqslant\|u\|_{W^{1,\infty}(\mathbb{R}^n)}.\end{aligned}$$

Hence in this case the maximal operator is bounded with constant one. Recall, that after a redefinition on a set of measure zero $u\in W^{1,\infty}(\mathbb{R}^n)$ is a bounded and Lipschitz continuous function, see Theorem 3.31.

5.2 Representation formulas and Riesz potentials

We begin with considering the one-dimensional case. If $u\in C_0^1(\mathbb{R})$, there exists an interval $[a,b]\subset\mathbb{R}$ such that $u(x)=0$ for every $x\in\mathbb{R}\setminus[a,b]$. By the fundamental theorem of calculus,

$$u(x) = u(a) + \int_a^x u'(y)\,dy = \int_{-\infty}^x u'(y)\,dy, \tag{5.9}$$

since $u(a) = 0$. On the other hand,

$$0 = u(b) = u(x) + \int_x^b u'(y)\,dy = u(x) + \int_x^\infty u'(y)\,dy,$$

so that

$$u(x) = -\int_x^\infty u'(y)\,dy. \tag{5.10}$$

Equalities (5.9) and (5.10) imply

$$\begin{aligned}2u(x) &= \int_{-\infty}^x u'(y)\,dy - \int_x^\infty u'(y)\,dy\\ &= \int_{-\infty}^x \frac{u'(y)(x-y)}{|x-y|}\,dy + \int_x^\infty \frac{u'(y)(x-y)}{|x-y|}\,dy\\ &= \int_{\mathbb{R}} \frac{u'(y)(x-y)}{|x-y|}\,dy,\end{aligned}$$

from which it follows that

$$u(x) = \frac{1}{2}\int_{\mathbb{R}} \frac{u'(y)(x-y)}{|x-y|}\,dy \quad \text{for every} \quad x \in \mathbb{R}.$$

Next we extend the fundamental theorem of calculus to $\mathbb{R}^n$.

Lemma 5.11 (Representation formula). If $u \in C_0^1(\mathbb{R}^n)$, then

$$u(x) = \frac{1}{\omega_{n-1}}\int_{\mathbb{R}^n} \frac{Du(y)\cdot(x-y)}{|x-y|^n}\,dy \quad \text{for every} \quad x \in \mathbb{R}^n,$$

where $\omega_{n-1} = n\Omega_n$ is the $(n-1)$-dimensional measure of $\partial B(0,1)$.

THE MORAL: This is a representation formula for a compactly supported continuously differentiable function in terms of its gradient. A function can be integrated back from its derivative using this formula.

Proof. If $x \in \mathbb{R}^n$ and $e \in \partial B(0,1)$, by the fundamental theorem of calculus

$$u(x) = -\int_0^\infty \frac{\partial}{\partial t}(u(x+te))\,dt = -\int_0^\infty Du(x+te)\cdot e\,dt.$$

By the Fubini theorem

$$\begin{aligned}
\omega_{n-1}u(x) &= u(x)\int_{\partial B(0,1)} 1\,dS(e) \\
&= -\int_{\partial B(0,1)}\int_0^\infty Du(x+te)\cdot e\,dt\,dS(e) \\
&= -\int_0^\infty \int_{\partial B(0,1)} Du(x+te)\cdot e\,dS(e)\,dt \qquad \text{(Fubini)} \\
&= -\int_0^\infty \int_{\partial B(0,t)} Du(x+y)\cdot \frac{y}{t}\frac{1}{t^{n-1}}\,dS(y)\,dt \\
&\quad (y = te,\ dS(e) = t^{1-n}\,dS(y)) \\
&= -\int_0^\infty \int_{\partial B(0,t)} Du(x+y)\cdot \frac{y}{|y|^n}\,dS(y)\,dt \\
&= -\int_{\mathbb{R}^n} \frac{Du(x+y)\cdot y}{|y|^n}\,dy \\
&= -\int_{\mathbb{R}^n} \frac{Du(z)\cdot (z-x)}{|z-x|^n}\,dz \qquad (z = x+y,\ dy = dz) \\
&= \int_{\mathbb{R}^n} \frac{Du(y)\cdot (x-y)}{|x-y|^n}\,dy.
\end{aligned}$$

□

Remark 5.12. By the Cauchy-Schwarz inequality and Lemma 5.11, we have

$$\begin{aligned}
|u(x)| &= \left|\frac{1}{\omega_{n-1}}\int_{\mathbb{R}^n} \frac{Du(y)\cdot(x-y)}{|x-y|^n}\,dy\right| \\
&\leqslant \frac{1}{\omega_{n-1}}\int_{\mathbb{R}^n} \frac{|Du(y)||x-y|}{|x-y|^n}\,dy \\
&= \frac{1}{\omega_{n-1}}\int_{\mathbb{R}^n} \frac{|Du(y)|}{|x-y|^{n-1}}\,dy \\
&= \frac{1}{\omega_{n-1}} I_1(|Du|)(x),
\end{aligned}$$

where $I_\alpha f$, $0<\alpha<n$, is the Riesz potential

$$I_\alpha f(x) = \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-\alpha}}\,dy.$$

THE MORAL: This gives a useful pointwise bound for a compactly supported smooth function in terms of the Riesz potential of the gradient.

Remark 5.13. A similar estimate holds almost everywhere if $u \in W^{1,p}(\mathbb{R}^n)$ or $u \in W_0^{1,p}(\Omega)$ (exercise).

We begin with a technical lemma for the Riesz potential for $\alpha = 1$.

Lemma 5.14. If $E \subset \mathbb{R}^n$ is a measurable set with $|E| < \infty$, then

$$\int_E \frac{1}{|x-y|^{n-1}}\,dy \leqslant c(n)|E|^{\frac{1}{n}}.$$

Proof. Let $B = B(x,r)$ be a ball with $|B| = |E|$. Then $|E| = |E \cap B| + |E \setminus B|$ and $|B| = |B \cap E| + |B \setminus E|$ which implies that $|E \setminus B| = |B \setminus E|$. Since $|x - y| \geqslant r$ for every $y \in E \setminus B$ and $|x - y| < r$ for every $y \in B \setminus E$, we have

$$\int_{E\setminus B} \frac{1}{|x-y|^{n-1}}\,dy \leqslant |E\setminus B|\frac{1}{r^{n-1}} = |B\setminus E|\frac{1}{r^{n-1}} \leqslant \int_{B\setminus E} \frac{1}{|x-y|^{n-1}}\,dy.$$

It follows that

$$\begin{aligned}\int_E \frac{1}{|x-y|^{n-1}}\,dy &= \int_{E\setminus B} \frac{1}{|x-y|^{n-1}}\,dy + \int_{E\cap B} \frac{1}{|x-y|^{n-1}}\,dy \\ &\leqslant \int_{B\setminus E} \frac{1}{|x-y|^{n-1}}\,dy + \int_{E\cap B} \frac{1}{|x-y|^{n-1}}\,dy \\ &= \int_B \frac{1}{|x-y|^{n-1}}\,dy \\ &= c(n)r = c(n)|B|^{\frac{1}{n}} = c(n)|E|^{\frac{1}{n}}.\end{aligned}$$

□

Lemma 5.15. Assume that $|\Omega| < \infty$ and $1 \leqslant p < \infty$. Then

$$\|I_1(|f|\chi_\Omega)\|_{L^p(\Omega)} \leqslant c(n,p)|\Omega|^{\frac{1}{n}}\|f\|_{L^p(\Omega)}.$$

T H E M O R A L : If $|\Omega| < \infty$, then $I_1 : L^p(\Omega) \to L^p(\Omega)$ is a bounded operator for $1 \leqslant p < \infty$.

Proof. If $p > 1$, Hölder's inequality and Lemma 5.14 give

$$\begin{aligned}\int_\Omega \frac{|f(y)|}{|x-y|^{n-1}}\,dy &= \int_\Omega \frac{|f(y)|}{|x-y|^{\frac{1}{p}(n-1)}} \frac{1}{|x-y|^{\frac{1}{p'}(n-1)}}\,dy \\ &\leqslant \left(\int_\Omega \frac{|f(y)|^p}{|x-y|^{n-1}}\,dy\right)^{\frac{1}{p}} \left(\int_\Omega \frac{1}{|x-y|^{n-1}}\,dy\right)^{\frac{1}{p'}} \\ &\leqslant c|\Omega|^{\frac{1}{np'}} \left(\int_\Omega \frac{|f(y)|^p}{|x-y|^{n-1}}\,dy\right)^{\frac{1}{p}} \\ &= c|\Omega|^{\frac{p-1}{np}} \left(\int_\Omega \frac{|f(y)|^p}{|x-y|^{n-1}}\,dy\right)^{\frac{1}{p}}.\end{aligned}$$

For $p = 1$, the inequality above is clear. Thus by Fubini's theorem and Lemma 5.14, we have

$$\begin{aligned}\int_\Omega |I_1(|f|\chi_\Omega)(x)|^p\,dx &\leqslant c|\Omega|^{\frac{p-1}{n}} \int_\Omega\int_\Omega \frac{|f(y)|^p}{|x-y|^{n-1}}\,dy\,dx \\ &\leqslant c|\Omega|^{\frac{p-1}{n}}|\Omega|^{\frac{1}{n}} \int_\Omega |f(y)|^p\,dy.\end{aligned}$$

□

Next we show that the Riesz potential can be bounded by the Hardy-Littlewood maximal function. We shall do this for the general α although $\alpha = 1$ will be most important for us.

Lemma 5.16. Let $0 < \alpha < n$. Then there exists a constant $c = c(n,\alpha)$ such that

$$\int_{B(x,r)} \frac{|f(y)|}{|x-y|^{n-\alpha}}\,dy \leqslant cr^{\alpha} Mf(x)$$

for every $x \in \mathbb{R}^n$ and $r > 0$.

Proof. Let $x \in \mathbb{R}^n$ and denote $A_i = B(x, r2^{-i})$, $i = 0,1,2,\dots$. Then

$$\begin{aligned}
\int_{B(x,r)} \frac{|f(y)|}{|x-y|^{n-\alpha}}\,dy &= \sum_{i=0}^{\infty} \int_{A_i \setminus A_{i+1}} \frac{|f(y)|}{|x-y|^{n-\alpha}}\,dy \\
&\leqslant \sum_{i=0}^{\infty} \left(\frac{r}{2^{i+1}}\right)^{\alpha-n} \int_{A_i} |f(y)|\,dy \\
&= \Omega_n \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^{\alpha-n} \left(\frac{r}{2^i}\right)^{\alpha} \frac{1}{\Omega_n} \left(\frac{r}{2^i}\right)^{-n} \int_{A_i} |f(y)|\,dy \\
&= \Omega_n \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^{\alpha-n} \left(\frac{r}{2^i}\right)^{\alpha} \frac{1}{|A_i|} \int_{A_i} |f(y)|\,dy \\
&\leqslant cMf(x) r^{\alpha} \sum_{i=0}^{\infty} \left(\frac{1}{2^{\alpha}}\right)^{i} \\
&= cr^{\alpha} Mf(x).
\end{aligned}$$

□

Theorem 5.17 (Sobolev inequality for Riesz potentials). Assume that $\alpha > 0$, $p > 1$ and $\alpha p < n$. Then there exists a constant $c = c(n,p,\alpha)$ such tha

$$\|I_\alpha f\|_{L^{p^*}(\mathbb{R}^n)} \leqslant c\|f\|_{L^p(\mathbb{R}^n)}, \quad p^* = \frac{pn}{n-\alpha p}$$

for every $f \in L^p(\mathbb{R}^n)$ we have

T H E M O R A L : This is the Sobolev inequality for the Riesz potentials. Observe that p^* is the Sobolev conjugate of p if $\alpha = 1$.

Proof. If $f = 0$ almost everywhere, the claim is clear. Thus we may assume that $f \neq 0$ on a set of positive measure and thus $Mf > 0$ everywhere. By Hölder's inequality

$$\int_{\mathbb{R}^n \setminus B(x,r)} \frac{|f(y)|}{|x-y|^{n-\alpha}}\,dy \leqslant \left(\int_{\mathbb{R}^n \setminus B(x,r)} |f(y)|^p\,dy\right)^{\frac{1}{p}} \left(\int_{\mathbb{R}^n \setminus B(x,r)} |x-y|^{(\alpha-n)p'}\,dy\right)^{\frac{1}{p'}},$$

where

$$\begin{aligned}
\int_{\mathbb{R}^n \setminus B(x,r)} |x-y|^{(\alpha-n)p'}\,dy &= \int_r^{\infty} \int_{\partial B(x,\rho)} |x-y|^{(\alpha-n)p'}\,dS(y)\,d\rho \\
&= \int_r^{\infty} \rho^{(\alpha-n)p'} \underbrace{\int_{\partial B(x,\rho)} 1\,dS(y)}_{=\omega_{n-1}\rho^{n-1}}\,d\rho \\
&= \omega_{n-1} \int_r^{\infty} \rho^{(\alpha-n)p'+n-1}\,d\rho \\
&= \frac{\omega_{n-1}}{(n-\alpha)p'-n} r^{n-(n-\alpha)p'}.
\end{aligned}$$

The exponent can be written in the form

$$n-(n-\alpha)p' = n-(n-\alpha)\frac{p}{p-1} = \frac{\alpha p - n}{p-1},$$

and thus

$$\int_{\mathbb{R}^n\setminus B(x,r)} \frac{|f(y)|}{|x-y|^{n-\alpha}}\,dy \leqslant c r^{\alpha-\frac{n}{p}} \|f\|_{L^p(\mathbb{R}^n)}.$$

Lemma 5.16 implies

$$\begin{aligned}
|I_\alpha f(x)| &\leqslant \int_{\mathbb{R}^n} \frac{|f(y)|}{|x-y|^{n-\alpha}}\,dy \\
&= \int_{B(x,r)} \frac{|f(y)|}{|x-y|^{n-\alpha}}\,dy + \int_{\mathbb{R}^n\setminus B(x,r)} \frac{|f(y)|}{|x-y|^{n-\alpha}}\,dy \\
&\leqslant c\left(r^\alpha Mf(x) + r^{\alpha-\frac{n}{p}}\|f\|_{L^p(\mathbb{R}^n)}\right).
\end{aligned}$$

By choosing

$$r = \left(\frac{Mf(x)}{\|f\|_{L^p(\mathbb{R}^n)}}\right)^{-\frac{p}{n}},$$

we obtain

$$|I_\alpha f(x)| \leqslant c Mf(x)^{1-\frac{\alpha p}{n}} \|f\|_{L^p(\mathbb{R}^n)}^{\frac{\alpha p}{n}}. \tag{5.18}$$

By raising both sides to the power $p^* = \frac{np}{n-\alpha p}$, we have

$$|I_\alpha f(x)|^{p^*} \leqslant c Mf(x)^p \|f\|_{L^p(\mathbb{R}^n)}^{\frac{\alpha p}{n}p^*}$$

The maximal function theorem, see Theorem 5.3 (2) , implies

$$\begin{aligned}
\int_{\mathbb{R}^n} |I_\alpha f(x)|^{p^*}\,dy &\leqslant c\|f\|_{L^p(\mathbb{R}^n)}^{\frac{\alpha p}{n}p^*} \int_{\mathbb{R}^n} (Mf(x))^p\,dx \\
&= c\|f\|_{L^p(\mathbb{R}^n)}^{\frac{\alpha p}{n}p^*} \|Mf\|_{L^p(\mathbb{R}^n)}^p \\
&\leqslant c\|f\|_{L^p(\mathbb{R}^n)}^{\frac{\alpha p}{n}p^*} \|f\|_{L^p(\mathbb{R}^n)}^p
\end{aligned}$$

and thus

$$\|I_\alpha f\|_{L^{p^*}(\mathbb{R}^n)} \leqslant c\|f\|_{L^p(\mathbb{R}^n)}^{\frac{\alpha p}{n}+\frac{p}{p^*}} = c\|f\|_{L^p(\mathbb{R}^n)}.$$ □

Remark 5.19. From the proof of the previous theorem we also obtain a weak type estimate when $p=1$. Indeed, by (5.18) with $p=1$, there exists $c=c(n,\alpha)$ such that

$$|I_\alpha f(x)| \leqslant c Mf(x)^{1-\frac{\alpha}{n}} \|f\|_{L^1(\mathbb{R}^n)}^{\frac{\alpha}{n}}$$

and thus the maximal function theorem with $p=1$, see Theorem 5.3 (1), implies

$$\begin{aligned}
|\{x\in\mathbb{R}^n : |I_\alpha f(x)| > t\}| &\leqslant \left|\left\{x\in\mathbb{R}^n : cMf(x)^{\frac{n-\alpha}{n}} \|f\|_{L^1(\mathbb{R}^n)}^{\frac{\alpha}{n}} > t\right\}\right| \\
&\leqslant \left|\left\{x\in\mathbb{R}^n : Mf(x) > ct^{\frac{n}{n-\alpha}} \|f\|_{L^1(\mathbb{R}^n)}^{-\frac{\alpha}{n}\cdot\frac{n}{n-\alpha}}\right\}\right| \\
&\leqslant ct^{-\frac{n}{n-\alpha}} \|f\|_{L^1(\mathbb{R}^n)}^{\frac{\alpha}{n-\alpha}} \|f\|_{L^1(\mathbb{R}^n)} \\
&= ct^{-\frac{n}{n-\alpha}} \|f\|_{L^1(\mathbb{R}^n)}^{\frac{n}{n-\alpha}}
\end{aligned}$$

for every $t>0$. This also implies that

$$|\{x\in\mathbb{R}^n : |I_\alpha f(x)|\geqslant t\}|\leqslant ct^{-\frac{n}{n-\alpha}}\|f\|_{L^1(\mathbb{R}^n)}^{\frac{n}{n-\alpha}}$$

for every $t>0$.

This gives a second proof for the Sobolev-Gagliardo-Nirenberg inequality, see Theorem 3.3.

Corollary 5.20 (Sobolev-Gagliardo-Nirenberg inequality). If $1\leqslant p<n$, there exists a constant $c=c(n,p)$ such that

$$\|u\|_{L^{p^*}(\mathbb{R}^n)}\leqslant c\|Du\|_{L^p(\mathbb{R}^n)},\quad p^*=\frac{np}{n-p},$$

for every $u\in C_0^1(\mathbb{R}^n)$.

THE MORAL: The Sobolev-Gagliardo-Nirenberg inequality is a consequence of pointwise estimates for the function in terms of the Riesz potential of the gradient and the Sobolev inequality for the Riesz potentials.

Proof. $\boxed{1<p<\infty}$ By Remark 5.12

$$|u(x)|\leqslant\frac{1}{\omega_{n-1}}I_1(|Du|)(x)\quad\text{for every}\quad x\in\mathbb{R}^n,$$

Thus Theorem 5.17 with $\alpha=1$ gives

$$\|u\|_{L^{p^*}(\mathbb{R}^n)}\leqslant c\|I_1(|Du|)\|_{L^{p^*}(\mathbb{R}^n)}\leqslant c\|Du\|_{L^p(\mathbb{R}^n)}.$$

$\boxed{p=1}$ Let

$$A_j=\{x\in\mathbb{R}^n : 2^j<|u(x)|\leqslant 2^{j+1}\},\quad j\in\mathbb{Z},$$

and let $\varphi:\mathbb{R}\to\mathbb{R}$, $\varphi(t)=\max\{0,\min\{t,1\}\}$, be an auxiliary function. For $j\in\mathbb{Z}$ define $u_j:\mathbb{R}^n\to[0,1]$,

$$u_j(x)=\varphi(2^{1-j}|u(x)|-1)=\begin{cases}0, & |u(x)|\leqslant 2^{j-1},\\ 2^{1-j}|u(x)|-1, & 2^{j-1}<|u(x)|\leqslant 2^j,\\ 1, & |u(x)|>2^j.\end{cases}$$

A version of Lemma 2.1 for Lipschitz functions (exercise) implies that $u_j\in W^{1,1}(\mathbb{R}^n)$, $j\in\mathbb{Z}$. Observe that $Du_j=0$ almost everywhere in $\mathbb{R}^n\setminus A_{j-1}$, $j\in\mathbb{Z}$.

Then

$$\begin{aligned}
|A_j| &\leqslant |\{x \in \mathbb{R}^n : |u(x)| > 2^j\}| \\
&= |\{x \in \mathbb{R}^n : u_j(x) = 1\}| \qquad (|u(x)| > 2^j \Longrightarrow 2^{1-j}|u(x)| - 1 > 1) \\
&\leqslant \left|\{x \in \mathbb{R}^n : I_1(|Du_j|)(x) \geqslant \omega_{n-1}\}\right| \qquad \text{(Remark 5.12)} \\
&\leqslant c\left(\int_{\mathbb{R}^n} |Du_j(x)|\,dx\right)^{\frac{n}{n-1}} \qquad \text{(Remark 5.19)} \\
&= c\left(\int_{A_{j-1}} |Du_j(x)|\,dx\right)^{\frac{n}{n-1}} \\
&\leqslant c\left(\int_{A_{j-1}} \varphi'(2^{1-j}|u(x)|-1)2^{1-j}|Du(x)|\,dx\right)^{\frac{n}{n-1}} \\
&= c2^{-j\frac{n}{n-1}}\left(\int_{A_{j-1}} |Du(x)|\,dx\right)^{\frac{n}{n-1}}.
\end{aligned}$$

By summing over $j \in \mathbb{Z}$, we obtain

$$\begin{aligned}
\int_{\mathbb{R}^n} |u(x)|^{\frac{n}{n-1}}\,dx &= \sum_{j\in\mathbb{Z}} \int_{A_j} |u(x)|^{\frac{n}{n-1}}\,dx \\
&\leqslant \sum_{j\in\mathbb{Z}} 2^{(j+1)\frac{n}{n-1}}|A_j| \\
&\leqslant c\sum_{j\in\mathbb{Z}}\left(\int_{A_{j-1}} |Du(x)|\,dx\right)^{\frac{n}{n-1}} \\
&\leqslant c\left(\sum_{j\in\mathbb{Z}}\int_{A_{j-1}} |Du(x)|\,dx\right)^{\frac{n}{n-1}} \\
&= c\left(\int_{\mathbb{R}^n} |Du(x)|\,dx\right)^{\frac{n}{n-1}}.
\end{aligned}$$

In the last equality we used the fact that the sets A_j, $j \in \mathbb{Z}$, are pairwise disjoint.□

Remark 5.21. The Sobolev-Gagliardo-Nirenberg inequality for $u \in W^{1,p}(\mathbb{R}^n)$ follows from Corollary 5.20 by using the fact that $C_0^1(\mathbb{R}^n)$ is dense in $W^{1,p}(\mathbb{R}^n)$, $1 \leqslant p < n$.

5.3 Sobolev-Poincaré inequalities

Next we consider Sobolev-Poincaré inequalities in balls, compare with Theorem 3.12 and Theorem 3.13 for the corresponding estimates over cubes.

First we study the one-dimensional case. Assume that $u \in C^1(\mathbb{R})$ and let $y,z \in B(x,r) = (x-r,x+r)$. By the fundamental theorem of calculus

$$u(z) - u(y) = \int_z^y u'(t)\,dt.$$

Thus

$$|u(z)-u(y)| \leqslant \int_z^y |u'(t)|\,dt \leqslant \int_{x-r}^{x+r} |u'(t)|\,dt = \int_{B(x,r)} |u'(t)|\,dt$$

and

$$\begin{aligned}|u(z)-u_{B(x,r)}| &= \left|u(z) - \fint_{B(x,r)} u(y)\,dy\right| \\ &= \left|\fint_{B(x,r)} u(z)\,dy - \fint_{B(x,r)} u(y)\,dy\right| \\ &\leqslant \fint_{B(x,r)} |u(z)-u(y)|\,dy \leqslant \int_{B(x,r)} |u'(y)|\,dy.\end{aligned}$$

This is a pointwise estimate of the oscillation of the function. Next we generalize this to $\mathbb{R}^n$.

Lemma 5.22. Let $u \in C^1(\mathbb{R}^n)$ and $B(x,r) \subset \mathbb{R}^n$. There exists $c = c(n)$ such that

$$\left|u(z) - u_{B(x,r)}\right| \leqslant c \int_{B(x,r)} \frac{|Du(y)|}{|z-y|^{n-1}}\,dy$$

for every $z \in B(x,r)$.

T H E M O R A L : This is a pointwise estimate for the oscillation of the function in terms of the Riesz potential of the gradient.

Proof. For any $y, z \in B(x,r)$, we have

$$u(y)-u(z) = \int_0^1 \frac{\partial}{\partial t}(u(ty+(1-t)z))\,dt = \int_0^1 Du(ty+(1-t)z)\cdot(y-z)\,dt.$$

By the Cauchy-Schwarz inequality

$$|u(y)-u(z)| \leqslant |y-z| \int_0^1 |Du(ty+(1-t)z)|\,dt.$$

Let $\rho > 0$. In the next display, we make a change of variables

$$w = ty+(1-t)z \iff y = \frac{1}{t}(w-(1-t)z), \quad dS(y) = t^{1-n}\,dS(w).$$

Then we have $|w-z| = t|y-z|$ and $t^{n-1} = \left(\frac{|z-w|}{\rho}\right)^{n-1}$, where $\rho = |y-z|$. We arrive at

$$\begin{aligned}
&\int_{B(x,r)\cap\partial B(z,\rho)} |u(y)-u(z)|\,dS(y)\\
&\leqslant \rho\int_0^1\int_{B(x,r)\cap\partial B(z,\rho)} |Du(ty+(1-t)z)|\,dS(y)\,dt\\
&= \rho\int_0^1 \frac{1}{t^{n-1}}\int_{B(x,r)\cap\partial B(z,t\rho)} |Du(w)|\,dS(w)\,dt\\
&= \rho^n\int_0^1\int_{B(x,r)\cap\partial B(z,t\rho)} \frac{|Du(w)|}{|z-w|^{n-1}}\,dS(w)\,dt\\
&= \rho^{n-1}\int_0^\rho\int_{B(x,r)\cap\partial B(z,s)} \frac{|Du(w)|}{|z-w|^{n-1}}\,dS(w)\,ds \qquad (s=t\rho,\ dt=\tfrac{1}{\rho}\,ds)\\
&= \rho^{n-1}\int_{B(x,r)\cap B(z,\rho)} \frac{|Du(w)|}{|z-w|^{n-1}}\,dw. \qquad \text{(polar coordinates)}
\end{aligned}$$

Since $B(x,r)\subset B(z,2r)$, an integration in polar coordinates gives

$$\begin{aligned}
\left|u(z)-u_{B(x,r)}\right| &\leqslant \fint_{B(x,r)} |u(z)-u(y)|\,dy\\
&= \frac{1}{|B(x,r)|}\int_0^{2r}\int_{B(x,r)\cap\partial B(z,\rho)} |u(y)-u(z)|\,dS(y)\,d\rho\\
&\leqslant \frac{1}{|B(x,r)|}\int_0^{2r}\rho^{n-1}\int_{B(x,r)\cap B(z,\rho)} \frac{|Du(y)|}{|z-y|^{n-1}}\,dy\,d\rho\\
&\leqslant \frac{1}{|B(x,r)|}\int_0^{2r}\rho^{n-1}\,d\rho\int_{B(x,r)} \frac{|Du(y)|}{|z-y|^{n-1}}\,dy\\
&= c(n)\int_{B(x,r)} \frac{|Du(y)|}{|z-y|^{n-1}}\,dy.
\end{aligned}$$

□

Remarks 5.23:

(1) Assume that $u\in C^1(\mathbb{R}^n)$. By Lemma 5.22 and Lemma 5.16, we have

$$\begin{aligned}
|u(z)-u_{B(x,r)}| &\leqslant c\int_{B(x,r)} \frac{|Du(y)|}{|z-y|^{n-1}}\,dy\\
&= cI_1(|Du|\chi_{B(x,r)})(z)\\
&\leqslant c\int_{B(z,2r)} \frac{|Du(y)|\chi_{B(x,r)}(y)}{|z-y|^{n-1}}\,dy\\
&\leqslant crM(|Du|\chi_{B(x,r)})(z),
\end{aligned}$$

for every $z\in B(x,r)$.

Next we show that the corresponding inequalities hold true almost everywhere if $u\in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n)$, $1\leqslant p<\infty$. Since $C^\infty(B(x,r))$ is dense in $W^{1,p}(B(x,r))$, there exists a sequence $u_i\in C^\infty(B(x,r))$, $i=1,2,\dots$, such that $u_i\to u$ in $W^{1,p}(B(x,r))$ as $i\to\infty$. By passing to a subsequence, if necessary, we obtain an exceptional set $N_1 \subset \mathbb{R}^n$ with $|N_1| = 0$ such that

$$\lim_{i\to\infty} u_i(z) = u(z) < \infty$$

for every $z \in B(x,r)\setminus N_1$. By linearity of the Riesz potential and by Lemma 5.15, we have

$$\begin{aligned}
\big\| I_1(|Du_i|\chi_{B(x,r)}) - I_1(|Du|\chi_{B(x,r)})\big\|_{L^p(B(x,r))} \\
&= \big\| I_1((|Du_i| - |Du|)\chi_{B(x,r)})\big\|_{L^p(B(x,r))} \\
&\leqslant c|B(x,r)|^{\frac{1}{n}} \big\| |Du_i| - |Du| \big\|_{L^p(B(x,r))},
\end{aligned}$$

which implies that

$$I_1(|Du_i|\chi_{B(x,r)}) \to I_1(|Du|\chi_{B(x,r)}) \quad \text{in} \quad L^p(B(x,r)) \quad \text{as} \quad i\to\infty.$$

By passing to a subsequence, if necessary, we obtain an exceptional set $N_2 \subset B(x,r)$ with $|N_2| = 0$ such that

$$\lim_{i\to\infty} I_1(|Du_i|\chi_{B(x,r)})(z) = I_1(|Du|\chi_{B(x,r)})(z) < \infty$$

for every $z \in B(x,r)\setminus N_2$. Thus

$$\begin{aligned}
|u(z) - u_{B(x,r)}| &= \lim_{i\to\infty} |u_i(z) - (u_i)_{B(x,r)}| \\
&\leqslant c \lim_{i\to\infty} I_1(|Du_i|\chi_{B(x,r)})(z) \\
&= cI_1(|Du|\chi_{B(x,r)})(z) \\
&\leqslant crM(|Du|\chi_{B(x,r)})(z),
\end{aligned}$$

for every $z \in B(x,r)\setminus (N_1 \cup N_2)$.

(2) By Lemma 5.22 and (5.18), we have

$$\begin{aligned}
|u(z) - u_{B(x,r)}| &\leqslant c\int_{B(x,r)} \frac{|Du(y)|}{|z-y|^{n-1}}\,dy \\
&= cI_1(|Du|\,\chi_{B(x,r)})(z) \\
&\leqslant cM(|Du|\,\chi_{B(x,r)})(z)^{1-\frac{p}{n}} \big\| |Du|\,\chi_{B(x,r)}\big\|_{L^p(\mathbb{R}^n)}^{\frac{p}{n}}
\end{aligned}$$

for every $z \in B(x,r)$. The corresponding inequalities hold true almost everywhere if $u \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n)$, $1 \leqslant p < \infty$.

This gives a proof for the Sobolev-Poincaré inequality on balls, see Theorem 3.13 for the correspoding statement for cubes. Maximal function arguments can be used for cubes as well.

Theorem 5.24 (Sobolev-Poincaré inequality on balls). Assume that $u \in W^{1,p}(\mathbb{R}^n)$ and let $1 < p < n$. There exists $c = c(n,p)$ such that

$$\left(\fint_{B(x,r)} |u - u_{B(x,r)}|^{p^*}\,dy\right)^{\frac{1}{p^*}} \leqslant cr\left(\fint_{B(x,r)} |Du|^p\,dy\right)^{\frac{1}{p}}$$

for every $B(x,r) \subset \mathbb{R}^n$.

THE MORAL: The Sobolev-Poincaré inequality is a consequence of pointwise estimates for the oscillation of the function in terms of the Riesz potential of the gradient and the Sobolev inequality for the Riesz potentials.

Proof. By Remark 5.23, we have

$$|u(y) - u_{B(x,r)}| \leqslant c I_1(|Du|\chi_{B(x,r)})(y)$$

for almost every $y \in B(x,r)$. Thus Theorem 5.17 implies

$$\begin{aligned}\left(\int_{B(x,r)} |u - u_{B(x,r)}|^{p^*}\,dy\right)^{\frac{1}{p^*}} &\leqslant c\left(\int_{\mathbb{R}^n} I_1(|Du|\chi_{B(x,r)})^{p^*}\,dy\right)^{\frac{1}{p^*}}\\ &\leqslant c\left(\int_{\mathbb{R}^n} (|Du|\chi_{B(x,r)})^{p}\,dy\right)^{\frac{1}{p}}\\ &= c\left(\int_{B(x,r)} |Du|^{p}\,dy\right)^{\frac{1}{p}}. \qquad \square\end{aligned}$$

A similar argument can be used to prove a counterpart of Theorem 3.12 as well.

Theorem 5.25 (Poincaré inequality on balls). Assume that $u \in W^{1,p}(\mathbb{R}^n)$and let $1 < p < \infty$. There exists $c = c(n,p)$ such that

$$\left(\fint_{B(x,r)} |u - u_{B(x,r)}|^{p}\,dy\right)^{\frac{1}{p}} \leqslant cr\left(\fint_{B(x,r)} |Du|^{p}\,dy\right)^{\frac{1}{p}}$$

for every $B(x,r) \subset \mathbb{R}^n$.

Proof. By Remark 5.23, we have

$$|u(y) - u_{B(x,r)}| \leqslant crM(|Du|\chi_{B(x,r)})(y)$$

for almost every $y \in B(x,r)$. The maximal function theorem with $p > 1$, see Theorem 5.3 (2) , implies

$$\begin{aligned}\int_{B(x,r)} |u - u_{B(x,r)}|^{p}\,dy &\leqslant cr^p \int_{\mathbb{R}^n} M(|Du|\chi_{B(x,r)})^{p}\,dy\\ &\leqslant cr^p \int_{\mathbb{R}^n} (|Du|\chi_{B(x,r)})^{p}\,dy\\ &= cr^p \int_{B(x,r)} |Du|^{p}\,dy. \qquad \square\end{aligned}$$

The maximal function approach to Sobolev-Poincaré inequalities is more involved in the case $p = 1$, since then we only have a weak type estimate. However, it is possible to consider that case as well, but this requires a different proof. We begin with two rather technical lemmas.

Lemma 5.26. Assume that $E \subset \mathbb{R}^n$ is a measurable set and that $f : E \to [0,\infty]$ is a measurable function for which

$$\left|\{x \in E : f(x) = 0\}\right| \geqslant \tfrac{1}{2}|E|.$$

Then for every $a \in \mathbb{R}$ and $\lambda > 0$, we have

$$\left|\{x \in E : f(x) > \lambda\}\right| \leqslant \left|\left\{x \in E : |f(x) - a| > \tfrac{\lambda}{2}\right\}\right|.$$

Proof. Assume first that $|a| \leqslant \frac{\lambda}{2}$. If $x \in E$ with $f(x) > \lambda$, then

$$|f(x) - a| \geqslant f(x) - |a| > \frac{\lambda}{2}.$$

Thus $\{x \in E : f(x) > \lambda\} \subset \left\{x \in E : |f(x) - a| > \frac{\lambda}{2}\right\}$ and

$$\left|\{x \in E : f(x) > \lambda\}\right| \leqslant \left|\left\{x \in E : |f(x) - a| > \tfrac{\lambda}{2}\right\}\right|.$$

Assume then that $|a| > \frac{\lambda}{2}$. If $x \in E$ with $f(x) = 0$, then

$$|f(x) - a| = |a| > \frac{\lambda}{2}.$$

Thus

$$\{x \in E : f(x) = 0\} \subset \left\{x \in E : f(x) > \tfrac{\lambda}{2}\right\}.$$

If $|E| = \infty$, then by assumption

$$\left|\{x \in E : f(x) = 0\}\right| \geqslant \tfrac{1}{2}|E| = \infty$$

and thus $\left|\left\{x \in E : f(x) \geqslant \frac{\lambda}{2}\right\}\right| = \infty$. On the other hand, if $|E| < \infty$, then

$$\begin{aligned}
\left|\{x \in E : f(x) > \lambda\}\right| &\leqslant |E| - \left|\{x \in E : f(x) = 0\}\right| \\
&\leqslant |\{x \in E : f(x) = 0\}| \\
&\leqslant \left|\left\{x \in E : |f(x) - a| > \tfrac{\lambda}{2}\right\}\right|.
\end{aligned}$$

This completes the proof. □

Lemma 5.27. Assume that $u \in C^{0,1}(\mathbb{R}^n)$, that is, u is a bounded Lipschitz continuous function in $\mathbb{R}^n$, and let $B(x,r)$ be a ball in $\mathbb{R}^n$. Then there exists $\lambda_0 \in \mathbb{R}$ for which

$$\left|\{y \in B(x,r) : u(y) \geqslant \lambda_0\}\right| \geqslant \tfrac{1}{2}|B(x,r)| \quad \text{and} \quad \left|\{y \in B(x,r) : u(y) \leqslant \lambda_0\}\right| \geqslant \tfrac{1}{2}|B(x,r)|.$$

Proof. Denote $E_\lambda = \{y \in B(x,r) : u(y) \geqslant \lambda\}$, $\lambda \in \mathbb{R}$, and set

$$\lambda_0 = \sup\left\{\lambda \in \mathbb{R} : |E_\lambda| \geqslant \tfrac{1}{2}|B(x,r)|\right\}.$$

Note that $|\lambda_0| \leqslant \|u\|_{L^\infty(\mathbb{R}^n)} < \infty$. Thus there exists an increasing sequence of real numbers (λ_i) such that $\lambda_i \to \lambda_0$ and

$$|E_{\lambda_i}| \geqslant \tfrac{1}{2}|B(x,r)| \quad \text{for every} \quad i = 1,2,\dots.$$

Since $E_{\lambda_0} = \bigcap_{i=1}^{\infty} E_{\lambda_i}$, $E_{\lambda_1} \supset E_{\lambda_2} \supset \dots$ and $|E_{\lambda_i}| \leqslant |B(x,r)| < \infty$, we conclude that

$$|E_{\lambda_0}| = \lim_{i\to\infty} |E_{\lambda_i}| \geqslant \tfrac{1}{2}|B(x,r)|.$$

This shows that

$$\big|\{y \in B(x,r) : u(y) \geqslant \lambda_0\}\big| \geqslant \tfrac{1}{2}|B(x,r)|.$$

A similar argument shows the other claim (exercise). □

The next result is Theorem 5.24 with $p = 1$.

Theorem 5.28. Assume that $u \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^n)$. There exists $c = c(n)$ such that

$$\left(\fint_{B(x,r)} |u - u_{B(x,r)}|^{\frac{n}{n-1}}\,dy\right)^{\frac{n-1}{n}} \leqslant cr \fint_{B(x,r)} |Du|\,dy$$

for every $B(x,r) \subset \mathbb{R}^n$.

Proof. By Lemma 5.27 there is a number $\lambda_0 \in \mathbb{R}$ for which

$$\big|\{y \in B(x,r) : u(y) \geqslant \lambda_0\}\big| \geqslant \tfrac{1}{2}|B(x,r)| \quad \text{and} \quad \big|\{y \in B(x,r) : u(y) \leqslant \lambda_0\}\big| \geqslant \tfrac{1}{2}|B(x,r)|.$$

Denote

$$v_+ = \max\{u - \lambda_0, 0\} \quad \text{and} \quad v_- = -\min\{u - \lambda_0, 0\}.$$

Both of these functions belong to $W^{1,1}_{\mathrm{loc}}(\mathbb{R}^n)$. For the rest of the proof $v \geqslant 0$ denotes either v_+ or v_-. All statements are valid in both cases.

Let

$$A_j = \{y \in B(x,r) : 2^j < v(y) \leqslant 2^{j+1}\}, \quad j \in \mathbb{Z},$$

and let $\varphi : \mathbb{R} \to \mathbb{R}$, $\varphi(t) = \max\{0, \min\{t,1\}\}$, be an auxiliary function. We define $v_j : B(x,r) \to [0,1]$,

$$v_j(y) = \varphi(2^{1-j}v(y) - 1), \quad j \in \mathbb{Z}.$$

Lemma 2.1 implies $v_j \in W^{1,1}(B(x,r))$, $j \in \mathbb{Z}$. By Remark 5.23 (2) with $p = 1$, we have

$$|v_j(y) - (v_j)_{B(x,r)}|^{\frac{n}{n-1}} \leqslant cM(|Dv_j|\chi_{B(x,r)})(y)\big\||Dv_j|\chi_{B(x,r)}\big\|_{L^1(\mathbb{R}^n)}^{\frac{1}{n-1}}.$$

Lemma 5.26 with $\lambda = \frac{1}{2}$ and $a = (v_j)_{B(x,r)}$ gives

$$\begin{aligned}
|A_j| &\leqslant |\{y \in B(x,r) : v(y) > 2^j\}| \\
&\leqslant \left|\left\{y \in B(x,r) : v_j(y) > \tfrac{1}{2}\right\}\right| \\
&\leqslant \left|\left\{y \in B(x,r) : |v_j(y) - (v_j)_{B(x,r)}| > \tfrac{1}{4}\right\}\right| \\
&\leqslant \left|\left\{y \in \mathbb{R}^n : M(|Dv_j|\chi_{B(x,r)})(y) \geqslant c\big\||Dv_j|\chi_{B(x,r)}\big\|_{L^1(\mathbb{R}^n)}^{\frac{1}{1-n}}\right\}\right|.
\end{aligned}$$

The last term can be estimated using the weak type estimate for the maximal function, see Theorem 5.3 (1), and the fact that

$$|Dv_j| = 2^{1-j}|Dv|\chi_{A_{j-1}}$$

almost everywhere in $B(x,r)$. Thus we arrive at

$$
\begin{aligned}
\Big|\Big\{y\in\mathbb{R}^n : M(|Dv_j|\chi_{B(x,r)})(y) \geqslant c\big\||Dv_j|\chi_{B(x,r)}\big\|_{L^1(\mathbb{R}^n)}^{\frac{1}{1-n}}\Big\}\Big|
&\leqslant c\||Dv_j|\chi_{B(x,r)}\|_{L^1(\mathbb{R}^n)}^{\frac{1}{n-1}}\int_{\mathbb{R}^n}|Dv_j(y)|\chi_{B(x,r)}(y)\,dy\\
&= c\big\||Dv_j|\chi_{B(x,r)}\big\|_{L^1(\mathbb{R}^n)}^{\frac{n}{n-1}}\\
&\leqslant c2^{-\frac{jn}{n-1}}\big\||Dv|\chi_{A_{j-1}\cap B(x,r)}\big\|_{L^1(\mathbb{R}^n)}^{\frac{n}{n-1}}.
\end{aligned}
$$

Combining the above estimates for $|A_j|$, we obtain

$$
\begin{aligned}
\int_{B(x,r)} v(y)^{\frac{n}{n-1}}\,dy = \sum_{j\in\mathbb{Z}}\int_{A_j} v(y)^{\frac{n}{n-1}}\,dy &= \sum_{j\in\mathbb{Z}} 2^{\frac{(j+1)n}{n-1}}|A_j|\\
&\leqslant c\sum_{j\in\mathbb{Z}} 2^{\frac{(j+1)n}{n-1}}2^{-\frac{jn}{n-1}}\big\||Dv|\chi_{A_{j-1}\cap B(x,r)}\big\|_{L^1(\mathbb{R}^n)}^{\frac{n}{n-1}}\\
&\leqslant c\Big\|\sum_{j\in\mathbb{Z}}|Dv|\chi_{A_{j-1}\cap B(x,r)}\Big\|_{L^1(\mathbb{R}^n)}^{\frac{n}{n-1}}\\
&\leqslant c\big\||Du|\chi_{B(x,r)}\big\|_{L^1(\mathbb{R}^n)}^{\frac{n}{n-1}}.
\end{aligned}
$$

Since $|u-\lambda_0| = v_+ + v_-$, we obtain

$$
\begin{aligned}
\Big(\fint_{B(x,r)}|u-u_{B(x,r)}|^{\frac{n}{n-1}}\,dy\Big)^{\frac{n-1}{n}} &\leqslant 2\Big(\fint_{B(x,r)}|u-\lambda_0|^{\frac{n}{n-1}}\,dy\Big)^{\frac{n-1}{n}}\\
&\leqslant 2\Big(\fint_{B(x,r)}v_+(y)^{\frac{n}{n-1}}\,dy\Big)^{\frac{n-1}{n}} + 2\Big(\fint_{B(x,r)}v_-(y)^{\frac{n}{n-1}}\,dy\Big)^{\frac{n-1}{n}}\\
&\leqslant c\big\||Du|\chi_{B(x,r)}\big\|_{L^1(\mathbb{R}^n)}\\
&= c\int_{B(x,r)}|Du(y)|\,dy. \qquad \square
\end{aligned}
$$

THE MORAL: The proof shows that in this case a weak type estimate implies a strong type estimate. Observe carefully, that this does not hold in general. The reason why this works here is that we consider gradients, which have the property that they vanish on the set where the function itself is constant.

Next we give a maximal function proof for Morrey's inequality, see Theorem 3.23 and Remark 3.25 (3).

Theorem 5.29 (Morrey's inequality). Assume that $u \in C^1(\mathbb{R}^n)$ and let $n < p < \infty$. There exists $c = c(n,p)$ such that

$$
|u(y)-u(z)| \leqslant cr\Big(\fint_{B(x,r)}|Du|^p\,dw\Big)^{\frac{1}{p}}
$$

for every $B(x,r)\subset\mathbb{R}^n$ and $y,z\in B(x,r)$.

Proof. By Lemma 5.22

$$\begin{aligned}|u(y)-u(z)| &\leqslant |u(y)-u_{B(x,r)}|+|u_{B(x,r)}-u(z)|\\ &\leqslant c\int_{B(x,r)}\frac{|Du(w)|}{|y-w|^{n-1}}\,dw + c\int_{B(x,r)}\frac{|Du(w)|}{|z-w|^{n-1}}\,dw\end{aligned}$$

for every $y,z\in B(x,r)$. Hölder's inequality gives

$$\int_{B(x,r)}\frac{|Du(w)|}{|y-w|^{n-1}}\,dw \leqslant \left(\int_{B(x,r)}|Du|^p\,dw\right)^{\frac{1}{p}}\left(\int_{B(x,r)}|y-w|^{(1-n)p'}\,dw\right)^{\frac{1}{p'}},$$

where

$$\begin{aligned}\int_{B(x,r)}|y-w|^{(1-n)p'}\,dw &\leqslant \int_{B(y,2r)}|y-w|^{(1-n)p'}\,dw\\ &= \int_0^{2r}\int_{\partial B(y,\rho)}\rho^{(1-n)p'}\,dS(w)\,d\rho\\ &= \omega_{n-1}\int_0^{2r}\rho^{(1-n)p'+n-1}\,d\rho = cr^{n-(n-1)p'}.\end{aligned}$$

Since

$$(n-(n-1)p')\tfrac{1}{p'} = 1-\tfrac{n}{p},$$

we have

$$\int_{B(x,r)}\frac{|Du(w)|}{|y-w|^{n-1}}\,dw \leqslant cr^{1-\frac{n}{p}}\left(\int_{B(x,r)}|Du|^p\,dw\right)^{\frac{1}{p}}.$$

The same argument applies to the other integral as well, so that

$$|u(y)-u(z)| \leqslant cr^{1-\frac{n}{p}}\left(\int_{B(x,r)}|Du|^p\,dw\right)^{\frac{1}{p}}. \qquad \square$$

5.4 Sobolev inequalities on domains

In this section we study open sets $\Omega\subset\mathbb{R}^n$ for which the Sobolev-Poincaré inequality

$$\left(\int_\Omega |u-u_\Omega|^{p^*}\,dy\right)^{\frac{1}{p^*}} \leqslant c(p,n,\Omega)\left(\int_\Omega |Du|^p\,dy\right)^{\frac{1}{p}},\quad 1\leqslant p<n,\quad p^*=\frac{np}{n-p},$$

holds true for every $u\in W^{1,p}(\Omega)$. We already know that this inequality holds if Ω is a ball, but are there other sets for which it holds true as well? The following example shows that, in general, a function $u\in W^{1,p}(\Omega)$ is not integrable to any power $q>p$.

Example 5.30. Let

$$\Omega = \left\{(x_1,x_2)\in\mathbb{R}^2 : 0<x_1<1, |x_2|<\exp(-x_1^{-2})\right\}$$

and $u:\Omega\to\mathbb{R}$, $u(x_1,x_2)=x_1^3\exp(x_1^{-2})$. Then $u\in W^{1,1}(\Omega)$, but $u\notin L^q(\Omega)$ for every $q>1$ (exercise).

We begin by introducing an appropriate class of domains.

Definition 5.31. A bounded open set $\Omega \subset \mathbb{R}^n$ is a John domain, if there is $c_J \geqslant 1$ and a point $x_0 \in \Omega$ so that every point $x \in \Omega$ can be joined to x_0 by a path $\gamma : [0,1] \to \Omega$ such that $\gamma(0) = x$, $\gamma(1) = x_0$ and

$$\operatorname{dist}(\gamma(t), \partial\Omega) \geqslant c_J^{-1} |x - \gamma(t)|$$

for every $t \in [0,1]$.

THE MORAL: In a John domain every point can be connected to the distinguished point with a curve that is relatively far from the boundary.

Remarks 5.32:

(1) A bounded and connected open set $\Omega \subset \mathbb{R}^n$ satisfies the interior cone condition, if there exists a bounded cone

$$C = \{x \in \mathbb{R}^n : x_1^2 + \cdots + x_{n-1}^2 \leqslant a x_n^2, 0 \leqslant x_n \leqslant b\}$$

such that every point of Ω is a vertex of a cone congruent to C and entirely contained in Ω. Every domain with interior cone condition is a John domain (exercise). Roughly speaking the main difference between the interior cone condition and a John domain is that rigid cones are replaced by twisted cones.

(2) The collection of John domains is relatively large. For example, a domain whose boundary is von Koch snowflake is a John domain.

Theorem 5.33. If $\Omega \subset \mathbb{R}^n$ is a John domain and $1 \leqslant p < n$, then

$$\left(\int_\Omega |u - u_\Omega|^{p^*} \, dx \right)^{\frac{1}{p^*}} \leqslant c(p,n,c_J) \left(\int_\Omega |Du|^p \, dx \right)^{\frac{1}{p}}, \quad 1 < p < n,$$

for every $u \in C^\infty(\Omega) \cap W^{1,p}(\Omega)$.

THE MORAL: The Sobolev-Poincaré inequality holds for many other sets than balls as well.

WARNING: A rooms and passages example shows that the Sobolev-Poincaré inequality does not hold for all sets.

Proof. Let $x_0 \in \Omega$ be the distinguished point in the John domain. Denote $B_0 = B(x_0, r_0)$, $r_0 = \frac{1}{4} \operatorname{dist}(x_0, \partial\Omega)$. We show that there is a constant $M = M(c_J, n)$ such that for every $x \in \Omega$ there is a chain of balls $B_i = B(x_i, r_i) \subset \Omega$, $i = 1, 2, \ldots$, with the properties

(1) $|B_i \cup B_{i+1}| \leqslant M |B_i \cap B_{i+1}|$, $i = 1, 2, \ldots$,

(2) $\operatorname{dist}(x,B_i) \leqslant Mr_i$, $r_i \to 0$, $x_i \to x$ as $i \to \infty$ and

(3) no point of Ω belongs to more than M balls B_i.

To construct the chain, first assume that x is far from x_0, say $x \in \Omega \setminus B(x_0,2r_0)$. Let γ be a John path that connects x to x_0. All balls on the chain are centered on γ. We construct the balls recursively staring with B_0. Assume that $B_0,\dots,B_i$ have been constructed. Starting from the center x_i of B_i we move along γ towards x until we leave B_i for the last time. Let x_{i+1} be the point on γ where this happens and define

$$B_{i+1} = B(x_{i+1},r_{i+1}), \quad r_{i+1} = \frac{1}{4c_J}|x-x_{i+1}|.$$

By construction $B_i \subset \Omega$. Property (1) and $\operatorname{dist}(x,B_i) \leqslant Mr_i$ in (2) follow from the fact that the consecutive balls have comparable radii and that the radii are comparable to the distances of the centers of the balls to x.

To prove (3) assume that $y \in B_{i_1} \cap \dots \cap B_{i_k}$. Observe that the radii of B_{i_j}, $j = 1,\dots,k$, are comparable to $|x-y|$. By construction, if $i_j < i_m$, the center of B_{i_m} does not belong to B_{i_j}. This implies that the distances between the centers of B_{i_j} are comparable to $|x-y|$. The number of points in $\mathbb{R}^n$ with pairwise comparable distances is bounded, that is, if $z_1,\dots,z_m \in \mathbb{R}^n$ satisfy

$$\frac{r}{c} < \operatorname{dist}(z_i,z_j) < cr \quad \text{for} \quad i \neq j,$$

then $m \leqslant N = N(c,n)$. Thus k is bounded by a constant depending only on n and c_J. This implies (3). Property (3) implies , $r_i \to 0$, $x_i \to x$ as $i \to \infty$.

The case $x \in B(x_0,2r_0)$ is left as an exercise.

Since

$$u_{B_i} = \fint_{B_i} u(y)\,dy \to u(x)$$

for every $x \in \Omega$ as $i \to \infty$, we obtain

$$\begin{aligned}
|u(x) - u_{B_0}| &\leqslant \sum_{i=0}^{\infty} |u_{B_i} - u_{B_{i+1}}| \\
&\leqslant \sum_{i=0}^{\infty} \left(|u_{B_i} - u_{B_i \cap B_{i+1}}| + |u_{B_i \cap B_{i+1}} - u_{B_{i+1}}| \right) \\
&\leqslant \sum_{i=0}^{\infty} \left(\frac{|B_i|}{|B_i \cap B_{i+1}|} \fint_{B_i} |u - u_{B_i}|\,dy + \frac{|B_{i+1}|}{|B_i \cap B_{i+1}|} \fint_{B_{i+1}} |u - u_{B_{i+1}}|\,dy \right) \\
&\leqslant c \sum_{i=0}^{\infty} \fint_{B_i} |u - u_{B_i}|\,dy \quad \text{(property (1))} \\
&\leqslant c \sum_{i=0}^{\infty} r_i \fint_{B_i} |Du|\,dy \quad \text{(Poincaré inequality, see Theorem 5.25)} \\
&= c \sum_{i=0}^{\infty} \int_{B_i} \frac{|Du|}{r_i^{n-1}}\,dy.
\end{aligned}$$

Property (2) implies $|x-y| \leqslant cr_i$ for every $y \in B_i$ and

$$\frac{1}{r_i^{n-1}} \leqslant \frac{c}{|x-y|^{n-1}} \quad \text{for every} \quad y \in B_i.$$

Thus

$$|u(x)-u_{B_0}| \leqslant c\sum_{i=0}^{\infty}\int_{B_i}\frac{|Du(y)|}{|x-y|^{n-1}}\,dy \leqslant c\int_{\Omega}\frac{|Du(y)|}{|x-y|^{n-1}}\,dy.$$

The last inequality follows from property (3). We observe that

$$|u(x)-u_{\Omega}| \leqslant |u(x)-u_{B_0}|+|u_{B_0}-u_{\Omega}|,$$

where by Lemma 5.14 we have

$$\begin{aligned}
|u_{B_0}-u_{\Omega}| &\leqslant \frac{1}{|\Omega|}\int_{\Omega}|u(x)-u_{B_0}|\,dx\\
&\leqslant c\frac{1}{|\Omega|}\int_{\Omega}\int_{\Omega}\frac{|Du(y)|}{|x-y|^{n-1}}\,dx\,dy\\
&= c\frac{1}{|\Omega|}\int_{\Omega}|Du(y)|\left(\int_{\Omega}\frac{1}{|x-y|^{n-1}}\,dx\right)dy\\
&\leqslant c|\Omega|^{-1+\frac{1}{n}}\int_{\Omega}|Du(y)|\,dy.
\end{aligned}$$

By the John condition we have

$$c|\Omega|^{\frac{1}{n}} \geqslant \operatorname{dist}(x_0,\partial\Omega) \geqslant c_J^{-1}|x-x_0|$$

and by taking supremum over $x\in\Omega$ we obtain

$$\operatorname{diam}\Omega \leqslant c(n,c_J)|\Omega|^{\frac{1}{n}}$$

and thus

$$|\Omega|^{-\frac{n-1}{n}} \leqslant \frac{c}{|x-y|^{n-1}} \quad \text{for every} \quad y\in\Omega.$$

This implies

$$|u_{B_0}-u_{\Omega}| \leqslant c\int_{\Omega}\frac{|Du(y)|}{|x-y|^{n-1}}\,dy$$

and thus

$$|u(x)-u_{\Omega}| \leqslant c\int_{\Omega}\frac{|Du(y)|}{|x-y|^{n-1}}\,dy = cI_1(|Du|\chi_{\Omega})(x)$$

for almost every $x\in\Omega$. Theorem 5.17 implies

$$\begin{aligned}
\left(\int_{\Omega}|u(x)-u_{\Omega}|^{p^*}\,dx\right)^{\frac{1}{p^*}} &\leqslant c\left(\int_{\mathbb{R}^n}|I_1(|Du(x)|\chi_{\Omega}(x))|^{p^*}\,dx\right)^{\frac{1}{p^*}}\\
&\leqslant c\left(\int_{\mathbb{R}^n}(|Du(x)|\chi_{\Omega}(x))^{p}\,dx\right)^{\frac{1}{p}}\\
&= c\left(\int_{\Omega}|Du(x)|^{p}\,dx\right)^{\frac{1}{p}}.
\end{aligned}$$

□

Example 5.34. Since an annulus $B(x,2r)\setminus B(x,r)\subset\mathbb{R}^n$, $r>0$, is a John domain with $c_J=c(n)$ if $n\geqslant 2$, we have

$$\left(\int_{B(x,2r)\setminus B(x,r)}|u-u_{B(x,2r)\setminus B(x,r)}|^{p^*}\,dy\right)^{\frac{1}{p^*}} \leqslant c(p,n)\left(\int_{B(x,2r)\setminus B(x,r)}|Du|^{p}\,dy\right)^{\frac{1}{p}}$$

for every $u \in C^\infty(\mathbb{R}^n) \cap W^{1,p}(\Omega)$. This version of the Sobolev-Poincaré inequality is sometimes useful in PDEs. Note that the assumption $n \geqslant 2$ guarantees that an annulus is connected.

Similar argument as in the proof of the Sobolev-Poincaré inequality gives the following pointwise estimate.

Theorem 5.35. Assume that $u \in C^1(\mathbb{R}^n)$. There exists a constant $c = c(n)$ such that

$$|u(x) - u(y)| \leqslant c|x - y|(M|Du|(x) + M|Du|(y))$$

for every $x, y \in \mathbb{R}^n$.

Proof. Let $x, y \in \mathbb{R}^n$. Then $x, y \in B(x, 2|x - y|)$ and $B(x, 2|x - y|) \subset B(y, 4|x - y|)$. By Remark 5.23 we obtain

$$\begin{aligned} |u(x) - u(y)| &\leqslant |u(x) - u_{B(x,2|x-y|)}| + |u_{B(x,2|x-y|)} - u(y)| \\ &\leqslant c|x - y|(M|Du|(x) + M|Du|(y)). \end{aligned}$$

□

Remarks 5.36:

(1) If $|Du| \in L^p(\mathbb{R}^n)$, $1 < p \leqslant \infty$, then by Theorem 5.3 (2) we have $M|Du| \in L^p(\mathbb{R}^n)$.

(2) If $|Du| \in L^1(\mathbb{R}^n)$, then by Theorem 5.3 (1) we have $M|Du| < \infty$ almost everywhere.

(3) If $|Du| \in L^\infty(\mathbb{R}^n)$, then $M|Du| \leqslant \|M|Du|\|_{L^\infty(\mathbb{R}^n)} \leqslant \|Du\|_{L^\infty(\mathbb{R}^n)}$ everywhere. Thus

$$|u(x) - u(y)| \leqslant c\|Du\|_{L^\infty(\mathbb{R}^n)}|x - y|$$

for every $x, y \in \mathbb{R}^n$. In other words, u is Lipschitz continuous.

Theorem 5.37. Assume that $u \in W^{1,p}(\mathbb{R}^n)$, $1 < p < \infty$. There exists a constant $c = c(n)$ and a set $N \subset \mathbb{R}^n$ with $|N| = 0$ such that

$$|u(x) - u(y)| \leqslant c|x - y|(M|Du|(x) + M|Du|(y))$$

for every $x, y \in \mathbb{R}^n \setminus N$.

Proof. $C_0^\infty(\mathbb{R}^n)$ is dense in $W^{1,p}(\mathbb{R}^n)$ by Lemma 1.28. Thus there exists a sequence $u_i \in C_0^\infty(\mathbb{R}^n)$, $i = 1, 2, \dots$, such that $u_i \to u$ in $W^{1,p}(\mathbb{R}^n)$ as $i \to \infty$. By passing to a subsequence, if necessary, we obtain an exceptional set $N_1 \subset \mathbb{R}^n$ with $|N_1| = 0$ such that

$$\lim_{i \to \infty} u_i(x) = u(x) < \infty$$

for every $x \in \mathbb{R}^n \setminus N_1$. By the sublinearity of the maximal operator and the maximal function theorem

$$\begin{aligned}\big\|M|Du_i| - M|Du|\big\|_{L^p(\mathbb{R}^n)} &\leqslant \big\|M(|Du_i| - |Du|)\big\|_{L^p(\mathbb{R}^n)}\\ &\leqslant c\big\||Du_i| - |Du|\big\|_{L^p(\mathbb{R}^n)}\\ &\leqslant c\|Du_i - Du\|_{L^p(\mathbb{R}^n)}\end{aligned}$$

which implies that $M|Du_i| \to M|Du|$ in $L^p(\mathbb{R}^n)$ as $i \to \infty$. By passing to a subsequence, if necessary, we obtain an exceptional set $N_2 \subset \mathbb{R}^n$ with $|N_2| = 0$ such that

$$\lim_{i\to\infty} M|Du_i|(x) = M|Du|(x) < \infty$$

for every $x \in \mathbb{R}^n \setminus N_2$. By Theorem 5.35

$$\begin{aligned}|u(x) - u(y)| &= \lim_{i\to\infty} |u_i(x) - u_i(y)|\\ &\leqslant c|x-y| \lim_{i\to\infty}(M|Du_i|(x) + M|Du_i|(y))\\ &\leqslant c|x-y|(M|Du|(x) + M|Du|(y))\end{aligned}$$

for every $x \in \mathbb{R}^n \setminus (N_1 \cup N_2)$. □

Remark 5.38. Compare the proof above to Remark 5.23, which shows that the result holds for $u \in W^{1,p}(\mathbb{R}^n)$, $1 \leqslant p \leqslant \infty$.

The following definition motivated by Theorem 5.35.

Definition 5.39. Assume that $1 < p < \infty$ and let $u \in L^p(\mathbb{R}^n)$. For a measurable function $g : \mathbb{R}^n \to [0,\infty]$ we denote $g \in \mathscr{D}(u)$ if there exists a set $N \subset \mathbb{R}^n$ such that $|N| = 0$ and

$$|u(x) - u(y)| \leqslant |x-y|(g(x) + g(y)) \tag{5.40}$$

for every $x, y \in \mathbb{R}^n \setminus N$. We say that $u \in L^p(\mathbb{R}^n)$ belongs to the Hajłasz-Sobolev space $M^{1,p}(\mathbb{R}^n)$, if there exists $g \in L^p(\mathbb{R}^n)$ with $g \in \mathscr{D}(u)$. This space is endowed with the norm

$$\|u\|_{M^{1,p}(\mathbb{R}^n)} = \|u\|_{L^p(\mathbb{R}^n)} + \inf_{g\in\mathscr{D}(u)} \|g\|_{L^p(\mathbb{R}^n)}.$$

T H E M O R A L : The space $M^{1,p}(\mathbb{R}^n)$ is defined through the pointwise inequality (5.40).

Theorem 5.41. Assume that $1 < p < \infty$. Then $M^{1,p}(\mathbb{R}^n) = W^{1,p}(\mathbb{R}^n)$ and the associate norms are equivalent, that is, there exists c such that

$$\frac{1}{c}\|u\|_{W^{1,p}(\mathbb{R}^n)} \leqslant \|u\|_{M^{1,p}(\mathbb{R}^n)} \leqslant c\|u\|_{W^{1,p}(\mathbb{R}^n)}$$

for every measurable function u that belongs to $M^{1,p}(\mathbb{R}^n) = W^{1,p}(\mathbb{R}^n)$.

THE MORAL: This is a pointwise characterization of Sobolev spaces. This can be used as a definition of the first order Sobolev spaces on metric measure spaces.

Proof. $\boxed{\supset}$ Assume that $u \in W^{1,p}(\mathbb{R}^n)$. By Theorem 5.37 there exists $c = c(n)$ and a set $N \subset \mathbb{R}^n$ with $|N| = 0$ such that

$$|u(x) - u(y)| \leqslant c|x-y|(M|Du|(x) + M|Du|(y))$$

for every $x, y \in \mathbb{R}^n \setminus N$. Thus $g = cM|Du| \in \mathscr{D}(u) \cap L^p(\mathbb{R}^n)$ and by the maximal function theorem

$$\begin{aligned}
\|u\|_{M^{1,p}(\mathbb{R}^n)} &= \|u\|_{L^p(\mathbb{R}^n)} + \inf_{g \in \mathscr{D}(u)} \|g\|_{L^p(\mathbb{R}^n)} \\
&\leqslant \|u\|_{L^p(\mathbb{R}^n)} + \|cM|Du|\|_{L^p(\mathbb{R}^n)} \\
&\leqslant \|u\|_{L^p(\mathbb{R}^n)} + c\|Du\|_{L^p(\mathbb{R}^n)} \\
&\leqslant c\|u\|_{W^{1,p}(\mathbb{R}^n)},
\end{aligned}$$

where $c = c(n,p)$.

$\boxed{\subset}$ Assume then that $u \in M^{1,p}(\mathbb{R}^n)$. Then $u \in L^p(\mathbb{R}^n)$ and there exists $g \in L^p(\mathbb{R}^n)$ with $g \in \mathscr{D}(u)$. Then

$$|u(x+h) - u(x)| \leqslant |h|(g(x+h) + g(x))$$

for almost every $x, h \in \mathbb{R}^n$ and thus

$$\begin{aligned}
\int_{\mathbb{R}^n} |u(x+h) - u(x)|^p \, dx &\leqslant |h|^p \int_{\mathbb{R}^n} (g(x+h) + g(x))^p \, dx \\
&\leqslant 2^p |h|^p \int_{\mathbb{R}^n} (g(x+h)^p + g(x)^p) \, dx \\
&\leqslant 2^{p+1} \|g\|^p_{L^p(\mathbb{R}^n)} |h|^p.
\end{aligned}$$

By the characterization of the Sobolev space with the integrated difference quotients, see Theorem 2.32, we conclude $u \in W^{1,p}(\mathbb{R}^n)$ and

$$\|u\|_{W^{1,p}(\mathbb{R}^n)} \leqslant c\|u\|_{L^p(\mathbb{R}^n)} + c\|g\|_{L^p(\mathbb{R}^n)}.$$

The inequality $\|u\|_{W^{1,p}(\mathbb{R}^n)} \leqslant c\|u\|_{M^{1,p}(\mathbb{R}^n)}$ follows by taking infimum over all $g \in \mathscr{D}(u) \cap L^p(\mathbb{R}^n)$. □

WARNING: The characterization of $W^{1,p}(\mathbb{R}^n)$, $1 < p < \infty$, in Theorem 5.41 does not hold for $W^{1,1}(\mathbb{R}^n)$.

Example 5.42. Let $\Omega = (-\frac{1}{2}, \frac{1}{2}) \subset \mathbb{R}$ and $u : \Omega \to \mathbb{R}$,

$$u(x) = -\frac{x}{|x| \log |x|}.$$

Then $u'(x) = |x|^{-1}(\log|x|)^{-2}$ is the weak derivative of u in Ω, $u' \in L^1(\Omega)$ and $u \in L^1(\Omega)$ (exercise). Thus $u \in W^{1,1}(\Omega)$. If $g \in \mathscr{D}(u)$, then

$$-\frac{2}{\log x} = |u(x) - u(-x)| \leqslant 2x(g(x) + g(-x))$$

for almost every $0 < x < \frac{1}{2}$. It follows that

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} g(x)\,dx = \int_0^{\frac{1}{2}} (g(x) + g(-x))\,dx \geqslant -\int_0^{\frac{1}{2}} \frac{1}{\log x}\,dx = \infty.$$

Thus $u \notin M^{1,1}(\Omega)$. We may extend u to $\mathbb{R}$ in an appropriate way in oder to obtain the corresponding example in $W^{1,1}(\mathbb{R})$.

WARNING: The characterization of $W^{1,p}(\mathbb{R}^n)$, $1 < p < \infty$, in Theorem 5.41 does not hold for for $W^{1,p}(\Omega)$, where Ω is an open subset of $\mathbb{R}^n$.

Example 5.43. Let

$$\Omega = \{x = (x_1, x_2) \in \mathbb{R}^2 : |x| < 1\} \setminus \{(x_1, 0) \in \mathbb{R}^2 : -1 < x_1 \leqslant 0\} \subset \mathbb{R}^2$$

and $u : \Omega \to \mathbb{R}$,

$$u(x) = (2|x| - 1)_+ \arg x,$$

where $\arg(x) = \theta$, where $-\pi < \theta \leqslant \pi$ is the argument of x in polar coordinates. Then $u \in W^{1,p}(\Omega)$ for every $p > 1$ (exercise). Let

$$A_\pm = \left\{x = (x_1, x_2) \in \Omega : x_1 < -\tfrac{3}{4}, \pm x_2 > 0\right\} \quad \text{and} \quad A = A_- \cup A_+.$$

Then

$$u(x) > (2 \cdot \tfrac{3}{4} - 1)\tfrac{\pi}{2} = \tfrac{1}{2} \cdot \tfrac{\pi}{2} > \tfrac{1}{2}$$

for every $x \in A_+$ and, similarly, $u(x) < -\frac{1}{2}$ for every $x \in A_-$. If $g \in \mathscr{D}(u)$, then

$$\begin{aligned} g(x_1, x_2) + g(x_1, -x_2) &\geqslant \frac{|u(x_1, x_2) - u(x_1, -x_2)|}{|(x_1, x_2) - (x_1, -x_2)|} \\ &\geqslant \frac{u(x_1, x_2) - u(x_1, -x_2)}{2x_2} > \frac{1}{2x_2} \end{aligned}$$

for almost every $x = (x_1, x_2) \in A_+$. Thus

$$\int_A g(x)^p\,dx \geqslant \int_{A_+} \left(\frac{1}{2x_2}\right)^p dx_1\,dx_2 = \infty$$

for every $p > 1$. Thus $u \notin M^{1,p}(\Omega)$.

Remark 5.44. The pointwise characterization of Sobolev spaces in Theorem 5.41 is very useful in studying properties of Sobolev spaces. For example, if $u \in M^{1,p}(\mathbb{R}^n)$ and $g \in \mathscr{D}(u) \cap L^p(\mathbb{R}^n)$, then by the triangle inequality

$$\big||u(x)| - |u(y)|\big| \leqslant |u(x) - u(y)| \leqslant |x - y|(g(x) + g(y))$$

Thus $g \in \mathscr{D}(|u|) \cap L^p(\mathbb{R}^n)$ and consequently $|u| \in M^{1,p}(\mathbb{R}^n)$.

The pointwise characterization of Sobolev spaces in Theorem 5.41 can be used to show a similar result as Theorem 2.24.

Lemma 5.45. The function u belongs to $W^{1,p}(\mathbb{R}^n)$ if and only if $u \in L^p(\mathbb{R}^n)$ and there are functions $u_i \in L^p(\mathbb{R}^n)$, $i = 1,2,\dots$, such that $u_i \to u$ almost everywhere and $g_i \in \mathscr{D}(u_i) \cap L^p(\mathbb{R}^n)$ such that $g_i \to g$ almost everywhere for some $g \in L^p(\mathbb{R}^n)$.

Proof. If $u \in W^{1,p}(\mathbb{R}^n)$, then the claim of the lemma is clear. To see the converse, suppose that $u, g \in L^p(\mathbb{R}^n)$, $g_i \in \mathscr{D}(u_i) \cap L^p(\mathbb{R}^n)$ and $u_i \to u$ almost everywhere and $g_i \to g$ almost everywhere. Then

$$|u_i(x) - u_i(y)| \leqslant |x-y|\big(g_i(x) + g_i(y)\big) \tag{5.46}$$

for all $x, y \in \mathbb{R}^n \setminus F_i$ with $|F_i| = 0$, $i = 1,2,\dots$ Let $A \subset \mathbb{R}^n$ be such that $u_i(x) \to u(x)$ and $g_i(x) \to g(x)$ for all $x \in \mathbb{R}^n \setminus A$ and $|A| = 0$. Write $F = A \cup \bigcup_{i=1}^{\infty} F_i$. Then $|F| = 0$. Let $x, y \in \mathbb{R}^n \setminus F$, $x \neq y$. From (5.46) we obtain

$$|u(x) - u(y)| \leqslant |x-y|\big(g(x) + g(y)\big)$$

and thus $g \in \mathscr{D}(u) \cap L^p(\mathbb{R}^n)$. This completes the proof. □

5.5 Pointwise estimates

In this section we revisit pointwise inequalities for Sobolev functions.

Definition 5.47. Let $0 < \beta < \infty$ and $R > 0$. The fractional sharp maximal function of $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ is defined by

$$f^{\#}_{\beta,R}(x) = \sup_{0<r<R} r^{-\beta} \fint_{B(x,r)} |f - f_{B(x,r)}|\,dy,$$

If $R = \infty$, we write $f^{\#}_{\beta}(x)$.

THE MORAL: The fractional sharp maximal function controls the mean oscillation of the function instead of the average of the function as in the Hardy-Littlewood maximal function.

Next we prove a more general pointwise inequality than in Theorem 5.37.

Lemma 5.48. Assume that $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ and let $0 < \beta < \infty$. Then there is $c = c(\beta, n)$ and a set E with $|E| = 0$ such that

$$|f(x) - f(y)| \leqslant c|x-y|^{\beta}\big(f^{\#}_{\beta,4|x-y|}(x) + f^{\#}_{\beta,4|x-y|}(y)\big) \tag{5.49}$$

for every $x, y \in \mathbb{R}^n \setminus E$.

THE MORAL: This is a pointwise inequality for a function without the gradient.

Proof. Let E be the complement of the set of Lebesgue points of f. By Lebesgue's theorem $|E| = 0$. Fix $x \in \mathbb{R}^n \setminus E$, $0 < r < \infty$ and denote $B_i = B(x, 2^{-i} r)$, $i = 0, 1, \dots$ Then

$$\begin{aligned}
|f(x) - f_{B(x,r)}| &\leqslant \sum_{i=0}^{\infty} |f_{B_{i+1}} - f_{B_i}| \\
&\leqslant \sum_{i=0}^{\infty} \frac{|B_i|}{|B_{i+1}|} \fint_{B_i} |f - f_{B_i}|\,dy \\
&\leqslant c \sum_{i=0}^{\infty} (2^{-i} r)^{\beta} (2^{-i} r)^{-\beta} \fint_{B_i} |f - f_{B_i}|\,dy \\
&\leqslant c r^{\beta} f^{\#}_{\beta,r}(x).
\end{aligned}$$

Let $y \in B(x,r) \setminus E$. Then $B(x,r) \subset B(y,2r)$ and we obtain

$$\begin{aligned}
|f(y) - f_{B(x,r)}| &\leqslant |f(y) - f_{B(y,2r)}| + |f_{B(y,2r)} - f_{B(x,r)}| \\
&\leqslant c r^{\beta} f^{\#}_{\beta,2r}(y) + \fint_{B(x,r)} |f - f_{B(y,2r)}|\,dz \\
&\leqslant c r^{\beta} f^{\#}_{\beta,2r}(y) + c \fint_{B(y,2r)} |f - f_{B(y,2r)}|\,dz \\
&\leqslant c r^{\beta} f^{\#}_{\beta,2r}(y).
\end{aligned}$$

Let $x, y \in \mathbb{R}^n \setminus E$, $x \neq y$ and $r = 2|x - y|$. Then $x, y \in B(x,r)$ and hence

$$\begin{aligned}
|f(x) - f(y)| &\leqslant |f(x) - f_{B(x,r)}| + |f(y) - f_{B(x,r)}| \\
&\leqslant c|x - y|^{\beta} \big(f^{\#}_{\beta,4|x-y|}(x) + f^{\#}_{\beta,4|x-y|}(y)\big).
\end{aligned}$$

This completes the proof. □

Remark 5.50. Lemma 5.48 gives a Campanato type characterization for Hölder continuity. Let $0 < \beta \leqslant 1$ and assume that $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ with $f^{\#}_{\beta} \in L^{\infty}(\mathbb{R}^n)$. In other words, there exists a constant $c < \infty$ such that

$$r^{-\beta} \fint_{B(x,r)} |f - f_{B(x,r)}|\,dy \leqslant c$$

for every ball $B(x,r) \subset \mathbb{R}^n$. By Lemma 5.48, there exists a set $E \subset \mathbb{R}^n$ with $|E| = 0$ such that

$$|f(x) - f(y)| \leqslant c(n,\beta)|x - y|^{\beta} \big(f^{\#}_{\beta}(x) + f^{\#}_{\beta}(y)\big)$$

for every $x, y \in \mathbb{R}^n \setminus E$. This implies that

$$|u(x) - u(y)| \leqslant c(n,\beta) \|f^{\#}_{\beta}\|_{L^{\infty}(\mathbb{R}^n)} |x - y|^{\beta},$$

for every $x,y \in \mathbb{R}^n \setminus E$ with $|E| = 0$. In other words, if $f^\#_\beta \in L^\infty(\mathbb{R}^n)$, then f can be redefined on a set of measure zero so that the function is Hölder continuous in $\mathbb{R}^n$ with exponent β. On the other hand, if $f \in C^{0,\beta}(\mathbb{R}^n)$, then

$$\begin{aligned}|f(y) - f_{B(x,r)}| &= \left| f(y) - \fint_{B(x,r)} f(z)\,dz \right| \\ &\leqslant \fint_{B(x,r)} |f(y) - f(z)|\,dz \leqslant cr^\beta\end{aligned}$$

for every $y \in B(x,r)$. Thus

$$f^\#_{\beta,R}(x) = \sup_{0<r<R} r^{-\beta} \fint_{B(x,r)} |f(y) - f_{B(x,r)}|\,dy \leqslant c$$

for every $x \in \mathbb{R}^n$ and this implies that $f^\#_\beta \in L^\infty(\mathbb{R}^n)$. Thus f can be redefined on a set of measure zero so that the function is Hölder continuous with exponent β if and only if $f^\#_\beta \in L^\infty(\mathbb{R}^n)$. In the limiting case $\beta = 0$ we obtain the space of bounded mean oscillation $BMO(\mathbb{R}^n)$, which consists of functions $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ satisfying $f^\#_0 \in L^\infty(\mathbb{R}^n)$.

Definition 5.51. Let $0 \leqslant \alpha < n$ and $R > 0$. The fractional maximal function of $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ is

$$M_{\alpha,R} f(x) = \sup_{0<r<R} r^\alpha \fint_{B(x,r)} |f|\,dy,$$

For $R = \infty$, we write $M_{\alpha,\infty} = M_\alpha$. If $\alpha = 0$, we obtain the Hardy–Littlewood maximal function and we write $M_0 = M$.

If $u \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^n)$, then by the Poincaré inequality with $p = 1$, see Theorem 5.28, there is $c = c(n)$ such that

$$\fint_{B(x,r)} |u - u_{B(x,r)}|\,dy \leqslant cr \fint_{B(x,r)} |Du|\,dy$$

for every ball $B(x,r) \subset \mathbb{R}^n$. It follows that

$$r^{\alpha-1} \fint_{B(x,r)} |u - u_{B(x,r)}|\,dy \leqslant cr^\alpha \fint_{B(x,r)} |Du|\,dy$$

and consequently

$$u^\#_{1-\alpha,R}(x) \leqslant cM_{\alpha,R}|Du|(x) \tag{5.52}$$

for every $x \in \mathbb{R}^n$ and $R > 0$. Thus we have proved the following useful inequality.

Corollary 5.53. Let $u \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^n)$ and $0 \leqslant \alpha < 1$. Then there exist a constant $c = c(n,\alpha)$ and a set $E \subset \mathbb{R}^n$ with $|E| = 0$ such that

$$|u(x) - u(y)| \leqslant c|x-y|^{1-\alpha}\big(M_{\alpha,4|x-y|}|Du|(x) + M_{\alpha,4|x-y|}|Du|(y)\big)$$

for every $x,y \in \mathbb{R}^n \setminus E$.

Remark 5.54. Corollary 5.53 gives a Morrey type condition for Hölder continuity. Compare to Remark 5.50, where Hölder continuity was characterized by a Campanato approach. Let $0 \leqslant \alpha < 1$ and assume that $u \in W^{1,1}_{\text{loc}}(\mathbb{R}^n)$ with $M_\alpha|\nabla u| \in L^\infty(\mathbb{R}^n)$. In other words, there exists a constant $c < \infty$ such that

$$r^\alpha \fint_{B(x,r)} |Du|\,dy \leqslant c$$

for every ball $B(x,r) \subset \mathbb{R}^n$. By Corollary 5.53, there exists $E \subset \mathbb{R}^n$ with $|E| = 0$ such that

$$|u(x) - u(y)| \leqslant c(n,\alpha)|x-y|^{1-\alpha}\big(M_\alpha|Du|(x) + M_\alpha|Du|(y)\big) \tag{5.55}$$

for every $x, y \in \mathbb{R}^n \setminus E$. This implies that

$$|u(x) - u(y)| \leqslant c(n,\alpha)\|M_\alpha|Du|\|_{L^\infty(\mathbb{R}^n)}|x-y|^{1-\alpha},$$

for every $x, y \in \mathbb{R}^n \setminus E$ with $|E| = 0$. In other words, if $M_\alpha|\nabla u| \in L^\infty(\mathbb{R}^n)$ then u can be redefined on a set of measure zero so that the function is Hölder continuous in $\mathbb{R}^n$ with exponent $1-\alpha$. This shows that u is Hölder continuous with the exponent $1-\alpha$, after a possible redefinition on a set of measure zero.

Remark 5.56. From (5.55) we recover Morrey's inequality in Theorem 3.23. To see this, assume that $u \in W^{1,p}(\mathbb{R}^n)$ with $n < p < \infty$. By Hölder's inequality we have

$$M_{\frac{n}{p}}|Du|(x) \leqslant c(n)\big(M_n|Du|^p(x)\big)^{\frac{1}{p}} \leqslant c(n)\|Du\|_{L^p(\mathbb{R}^n)} < \infty,$$

for every $x \in \mathbb{R}^n$. Thus (5.55), with $\alpha = \frac{n}{p}$, implies

$$|u(x) - u(y)| \leqslant c(n,p)\|Du\|_{L^p(\mathbb{R}^n)}|x-y|^{1-\frac{n}{p}}$$

for every $x, y \in \mathbb{R}^n \setminus E$ with $|E| = 0$. This shows that u is Hölder continuous with the exponent $1 - \frac{n}{p}$ after a possible redefinition on a set of measure zero.

The next result shows that this gives a characterization of $W^{1,p}(\mathbb{R}^n)$ for $1 < p \leqslant \infty$.

Theorem 5.57. Let $1 < p < \infty$. Then the following four conditions are equivalent.

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

(2) $u \in L^p(\mathbb{R}^n)$ and there is $g \in L^p(\mathbb{R}^n)$, $g \geqslant 0$, such that

$$|u(x) - u(y)| \leqslant |x-y|(g(x) + g(y))$$

for every $x, y \in \mathbb{R}^n \setminus E$ with $|E| = 0$.

(3) $u \in L^p(\mathbb{R}^n)$ and there is $g \in L^p(\mathbb{R}^n)$, $g \geqslant 0$, such that the Poincaré inequality

$$\fint_{B(x,r)} |u - u_{B(x,r)}|\,dy \leqslant cr \fint_{B(x,r)} g\,dy$$

holds for every $x \in \mathbb{R}^n$ and $r > 0$.

(4) $u \in L^p(\mathbb{R}^n)$ and $u_1^\# \in L^p(\mathbb{R}^n)$.

Proof. (1) We have already seen that (1) implies (2).

(2) To prove that (2) implies (3), we integrate the pointwise inequality twice over the ball $B(x,r)$. After the first integration we obtain

$$\begin{aligned}|u(y)-u_{B(x,r)}| &= \left|u(y)-\fint_{B(x,r)} u(z)\,dz\right| \\ &\leqslant \fint_{B(x,r)} |u(y)-u(z)|\,dz \\ &\leqslant 2r\left(g(y)+\fint_{B(x,r)} g(z)\,dz\right)\end{aligned}$$

from which we have

$$\begin{aligned}\fint_{B(x,r)} |u(y)-u_{B(x,r)}|\,dy &\leqslant 2r\left(\fint_{B(x,r)} g(y)\,dy+\fint_{B(x,r)} g(z)\,dz\right) \\ &\leqslant 4r\fint_{B(x,r)} g(y)\,dy.\end{aligned}$$

(3) To show that (3) implies (4) we observe that

$$u_1^\#(x) = \sup_{r>0}\frac{1}{r}\fint_{B(x,r)} |u-u_{B(x,r)}|\,dy \leqslant c\sup_{r>0}\fint_{B(x,r)} g\,dy = cMg(x).$$

(4) Then we show that (4) implies (1). By Lemma 5.48

$$|u(x)-u(y)| \leqslant c|x-y|(u_1^\#(x)+u_1^\#(y))$$

for every $x,y \in \mathbb{R}^n \setminus E$ with $|E|=0$. If we denote $g = cu_1^\#$, then $g \in L^p(\mathbb{R}^n)$ and

$$|u(x)-u(y)| \leqslant |x-y|(g(x)+g(y))$$

for every $x,y \in \mathbb{R}^n \setminus E$ with $|E|=0$. Then we use the characterization of Sobolev spaces $W^{1,p}(\mathbb{R}^n)$, $1<p<\infty$, with integrated difference quotients, see Theorem 2.32. Let $h \in \mathbb{R}^n$. Then

$$|u_h(x)-u(x)| = |u(x+h)-u(x)| \leqslant |h|(g_h(x)+g(x)),$$

from which we conclude that

$$\|u_h-u\|_{L^p(\mathbb{R}^n)} \leqslant |h|(\|g_h\|_{L^p(\mathbb{R}^n)}+\|g\|_{L^p(\mathbb{R}^n)}) = 2|h|\|g\|_{L^p(\mathbb{R}^n)}.$$

The claim follows from this. □

Remark 5.58. It can be shown that $u \in W^{1,1}(\mathbb{R}^n)$ if and only if $u \in L^1(\mathbb{R}^n)$ and there exists a nonnegative function $g \in L^1(\mathbb{R}^n)$ and $\sigma \geqslant 1$ such that

$$|u(x)-u(y)| \leqslant |x-y|\big(M_{\sigma|x-y|}g(x)+M_{\sigma|x-y|}g(y)\big)$$

for every $x,y \in \mathbb{R}^n \setminus E$ with $|E|=0$. Moreover, if this inequality holds, then $|Du| \leqslant c(n,\sigma)g$ almost everywhere, see Theorem 4 in P. Hajłasz: A new characterization of the Sobolev space, Studia Math. 159 (2003), no. 2, 263–275.

5.6 Lipschitz truncation

Smooth functions in $C^\infty(\Omega)$ and $C_0^\infty(\Omega)$ are often used as canonical test functions in mathematical analysis. However, in many occasions smooth functions can be replaced by a more flexible class of Lipschitz functions. One highly useful property of Lipschitz functions, not shared by the smooth functions, is that the pointwise minimum and maximum over L-Lipschitz functions are still L-Lipschitz. The same is in fact true also for pointwise infimum and supremum of L-Lipschitz functions, if these are finite at a single point. In particular, it follows that if $u : A \to \mathbb{R}$ is an L-Lipschitz function, then the truncations $\max\{u,c\}$ and $\min\{u,c\}$ with $c \in \mathbb{R}$ are L-Lipschitz.

Theorem 5.59 (McShane). Assume that $A \subset \mathbb{R}^n$, $0 \leqslant L < \infty$ and that $f : A \to \mathbb{R}$ is an L-Lipschitz function. There exists an L-Lipschitz function $f^* : \mathbb{R}^n \to \mathbb{R}$ such that $f^*(x) = f(x)$ for every $x \in A$.

THE MORAL: Every Lipschitz continuous function defined on a subset A of $\mathbb{R}^n$ can be extended as a Lipschitz continuous function to the whole $\mathbb{R}^n$.

Proof. Let $f^* : \mathbb{R}^n \to \mathbb{R}$,

$$f^*(x) = \inf\{f(a) + L|x-a| : a \in A\}.$$

We claim that $f^*(b) = f(b)$ for every $b \in A$. To see this we observe that

$$f(b) - f(a) \leqslant |f(b) - f(a)| \leqslant L|b-a|,$$

which implies $f(b) \leqslant f(a) + L|b-a|$ for every $a \in A$. By taking infimum over $a \in A$ we obtain $f(b) \leqslant f^*(b)$. On the other hand, by the definition $f^*(b) \leqslant f(b)$ for every $b \in A$. Thus $f^*(b) = f(b)$ for every $b \in A$.

Then we claim that f^* is L-Lipschitz in $\mathbb{R}^n$. Let $x, y \in \mathbb{R}^n$. Then

$$\begin{aligned} f^*(x) &= \inf\{f(a) + L|x-a| : a \in A\} \\ &\leqslant \inf\{f(a) + L(|y-a| + |x-y|) : a \in A\} \\ &\leqslant \inf\{f(a) + L|y-a| : a \in A\} + L|x-y| \\ &= f^*(y) + L|x-y|. \end{aligned}$$

By switching the roles of x and y, we arrive at $f^*(y) \leqslant f^*(x) + L|x-y|$. This implies that $-L|x-y| \leqslant f^*(x) - f^*(y) \leqslant L|x-y|$. ☐

Remark 5.60. The function $f_* : \mathbb{R}^n \to \mathbb{R}$,

$$f_*(x) = \sup\{f(a) - L|x-a| : a \in A\}.$$

is an L-Lipschitz extension of f as well. We can see, that f^* is the largest L-Lipschitz extension of f in the sense that if $g: \mathbb{R}^n \to \mathbb{R}$ is L-Lipschitz and $g|_A = f$, then $g \leqslant f^*$. Correspondingly, the function f_* is the smallest L-Lipschitz extension of f.

Since $C_0^\infty(\mathbb{R}^n)$ is dense in $W^{1,p}(\mathbb{R}^n)$, also compactly supported Lipschitz functions are dense in $W^{1,p}(\mathbb{R}^n)$. By Theorem 5.37, we give a density result for Lipschitz functions in $W^{1,p}(\mathbb{R}^n)$. The main difference of the following Lipschitz truncation result to the standard mollification approximation $u_\varepsilon \to u$ as $\varepsilon \to 0$ is that the function is not changed in a good set $\{x \in \mathbb{R}^n : u_\varepsilon(x) = u(x)\}$ and there is an estimate for the measure of the bad set $\{x \in \mathbb{R}^n : u_\varepsilon(x) \neq u(x)\}$.

Theorem 5.61. Assume that $u \in W^{1,p}(\mathbb{R}^n)$, $1 < p < \infty$. Then for every $\varepsilon > 0$ there exists a Lipschitz continuous function $u_\varepsilon : \mathbb{R}^n \to \mathbb{R}$ such that

(1) $|\{x \in \mathbb{R}^n : u_\varepsilon(x) \neq u(x)\}| < \varepsilon$ and

(2) $\|u - u_\varepsilon\|_{W^{1,p}(\mathbb{R}^n)} < \varepsilon$.

Proof. **Step 1:** Let

$$E_\lambda = \{x \in \mathbb{R}^n : M|Du|(x) \leqslant \lambda\}, \quad \lambda > 0.$$

By Theorem 5.37, there exists a constant $c = c(n)$ such that

$$|u(x) - u(y)| \leqslant c|x-y|(M|Du|(x) + M|Du|(y)) \leqslant c\lambda|x-y|$$

for almost every $x, y \in E_\lambda$. This shows that u is $c\lambda$-Lipschitz continuous in E_λ. The McShane extension theorem, see Theorem 5.59, gives a $c\lambda$-Lipschitz continuous extension $v_\lambda : \mathbb{R}^n \to \mathbb{R}$. We truncate v_λ and obtain a $2c\lambda$-Lipschitz function

$$u_\lambda = \max\{-\lambda, \min\{v_\lambda, \lambda\}\}.$$

Observe that $|u_\lambda| \leqslant \lambda$ in $\mathbb{R}^n$ and $u_\lambda = u$ almost everywhere in $E_\lambda \cap \{|u| \leqslant \lambda\}$. By Theorem 3.31, we have $u_\lambda \in W^{1,\infty}(\mathbb{R}^n)$. In particular, this implies that the weak gradient Du_λ exists and $u_\lambda \in W^{1,p}_{\text{loc}}(\mathbb{R}^n)$.

Step 2: Consider measure of the set

$$\mathbb{R}^n \setminus E_\lambda = \{x \in \mathbb{R}^n : M|Du|(x) > \lambda\}.$$

We show that there exits $c = c(n,p)$ such that

$$\begin{aligned}\lambda^p|\mathbb{R}^n \setminus E_\lambda| &\leqslant \lambda^p|\{x \in \mathbb{R}^n : M|Du|(x)| > \lambda\}| \\ &\leqslant c\int_{\{x\in\mathbb{R}^n:|Du(x)|>\frac{\lambda}{2}\}} |Du(x)|^p\,dx \to 0\end{aligned}$$

as $\lambda \to \infty$, since $Du \in L^p(\mathbb{R}^n)$. Consequently $\lambda^p|\mathbb{R}^n \setminus E_\lambda| \to 0$ and $|\mathbb{R}^n \setminus E_\lambda| \to 0$ as $\lambda \to \infty$. This follows by choosing $f = |Du|$ in the following general fact for the Hardy-Littlewood maximal function.

Claim: If $f \in L^p(\mathbb{R}^n)$, there exists $c = c(n,p)$ such that

$$|\{x \in \mathbb{R}^n : Mf(x) > \lambda\}| \leqslant \frac{c}{\lambda} \int_{\{x\in\mathbb{R}^n : |f(x)| > \frac{\lambda}{2}\}} |f(x)|^p \, dx, \quad \lambda > 0.$$

Reason. Let $f = f_1 + f_2$, where $f_1 = f\chi_{\{|f|>\frac{\lambda}{2}\}}$ and $f_2 = f\chi_{\{|f|\leqslant\frac{\lambda}{2}\}}$. Then

$$\int_{\mathbb{R}^n} |f_1(x)|\,dx = \int_{\{x\in\mathbb{R}^n : |f(x)|>\frac{\lambda}{2}\}} |f_1(x)|^p |f_1(x)|^{1-p} \, dx \leqslant \left(\frac{\lambda}{2}\right)^{1-p} \|f\|^p_{L^p(\mathbb{R}^n)} < \infty.$$

This shows that $f_1 \in L^1(\mathbb{R}^n)$. On the other hand, $|f_2(x)| \leqslant \frac{\lambda}{2}$ for every $x \in \mathbb{R}^n$, which implies $\|f_2\|_{L^\infty(\mathbb{R}^n)} \leqslant \frac{\lambda}{2}$ and $f_2 \in L^\infty(\mathbb{R}^n)$. Thus every L^p function can be represented as a sum of an L^1 function and an L^∞ function. By Lemma 5.2, we have

$$\|Mf_2\|_{L^\infty(\mathbb{R}^n)} \leqslant \|f_2\|_{L^\infty(\mathbb{R}^n)} \leqslant \frac{\lambda}{2}$$

From this we conclude using sublinearity of the maximal operator that

$$Mf(x) = M(f_1 + f_2)(x) \leqslant Mf_1(x) + Mf_2(x) \leqslant Mf_1(x) + \frac{\lambda}{2}$$

for every $x \in \mathbb{R}^n$ and thus $Mf(x) > \lambda$ implies $Mf_1(x) > \frac{\lambda}{2}$. It follows that

$$|\{x \in \mathbb{R}^n : Mf(x) > \lambda\}| \leqslant \left|\{x \in \mathbb{R}^n : Mf_1(x) > \tfrac{\lambda}{2}\}\right|$$

for every $\lambda > 0$.

$\boxed{p = 1}$ By the maximal function theorem on $L^1(\mathbb{R}^n)$, see Theorem 5.3 (1) , we have

$$\left|\{x \in \mathbb{R}^n : Mf_1(x) > \tfrac{\lambda}{2}\}\right| \leqslant \frac{c}{\lambda}\|f_1\|_{L^1(\mathbb{R}^n)} = \frac{c}{\lambda}\int_{\{x\in\mathbb{R}^n : |f(x)|>\frac{\lambda}{2}\}} |f(x)|\,dx$$

for every $\lambda > 0$.

$\boxed{1 < p < \infty}$ By Chebyshev's inequality and by the maximal function theorem $L^p(\mathbb{R}^n)$, $p > 1$, see Theorem 5.3 (2) , we have

$$\begin{aligned}\left|\{x \in \mathbb{R}^n : Mf_1(x) > \tfrac{\lambda}{2}\}\right| &\leqslant \left(\frac{2}{\lambda}\right)^p \int_{\mathbb{R}^n} (Mf_1(x))^p \, dx \leqslant \frac{c}{\lambda^p}\int_{\mathbb{R}^n} |f_1(x)|^p \, dx \\ &= \frac{c}{\lambda^p}\int_{\{x\in\mathbb{R}^n : |f(x)|>\frac{\lambda}{2}\}} |f(x)|\,dx\end{aligned}$$

for every $\lambda > 0$. ■

Step 3: Next we prove an estimate for $\|u - u_\varepsilon\|_{W^{1,p}(\mathbb{R}^n)}$. Since $u_\lambda = u$ in E_λ and $|u_\lambda| \leqslant \lambda$ in $\mathbb{R}^n$, we have

$$\begin{aligned}\|u_\lambda - u\|^p_{L^p(\mathbb{R}^n)} &= \int_{\mathbb{R}^n\setminus E_\lambda} |u_\lambda - u|^p \, dx \\ &\leqslant 2^p \left(\int_{\mathbb{R}^n\setminus E_\lambda} |u_\lambda|^p \, dx + \int_{\mathbb{R}^n\setminus E_\lambda} |u|^p \, dx\right) \\ &\leqslant 2^p \left(\lambda^p |\mathbb{R}^n \setminus E_\lambda| + \int_{\mathbb{R}^n\setminus E_\lambda} |u|^p \, dx\right) \xrightarrow{\lambda\to\infty} 0.\end{aligned}$$

To prove the corresponding estimate for the gradients, we note that

$$D(u_\lambda - u) = \chi_{\mathbb{R}^n\setminus E_\lambda} D(u_\lambda - u) = \chi_{\mathbb{R}^n\setminus E_\lambda} Du_\lambda - \chi_{\mathbb{R}^n\setminus E_\lambda} Du$$

almost everywhere. Since u_λ is $c\lambda$-Lipschitz continuous and thus $|Du_\lambda| \leqslant c\lambda$ almost everywhere, we have

$$\begin{aligned}
\|D(u_\lambda - u)\|^p_{L^p(\mathbb{R}^n)} &= \int_{\mathbb{R}^n\setminus E_\lambda} |D(u_\lambda - u)|^p \, dx \\
&\leqslant 2^p \left(\int_{\mathbb{R}^n\setminus E_\lambda} |Du_\lambda|^p \, dx + \int_{\mathbb{R}^n\setminus E_\lambda} |Du|^p \, dx \right) \\
&\leqslant 2^p \left((2c\lambda)^p |\mathbb{R}^n \setminus E_\lambda| + \int_{\mathbb{R}^n\setminus E_\lambda} |Du|^p \, dx \right) \xrightarrow{\lambda\to\infty} 0.
\end{aligned}$$

Thus $\|u - u_\lambda\|_{W^{1,p}(\mathbb{R}^n)} \to 0$ as $\lambda \to \infty$. Observe that

$$\{x \in \mathbb{R}^n : u(x) \neq u_\lambda(x)\} \subset (\mathbb{R}^n \setminus E_\lambda) \cup \{x \in \mathbb{R}^n : |u(x)| > \lambda\},$$

with $|\mathbb{R}^n \setminus E_\lambda| \to 0$ as $\lambda \to \infty$ and

$$|\{x \in \mathbb{R}^n : |u(x)| > \lambda\}| \leqslant \frac{1}{\lambda^p} \int_{\mathbb{R}^n} |u|^p \, dy \xrightarrow{\lambda\to\infty} 0.$$

This proves the claims. □

Next we discuss another approach to prove Theorem 5.61. Assume that $u \in W^{1,p}(\mathbb{R}^n)$, $1 < p < \infty$. Let

$$E_\lambda = \{x \in \mathbb{R}^n : M|Du|(x) \leqslant \lambda\}, \quad \lambda > 0.$$

Let Q_i, $i = 1,2,\dots$ be a Whitney decomposition of an open set $\mathbb{R}^n \setminus E_\lambda$ with the following properties:

- each Q_i is open and cubes Q_i, $i = 1,2,\dots$, are pairwise disjoint,
- $\mathbb{R}^n \setminus E_\lambda = \cup_{i=1}^\infty \overline{Q}_i$,
- $4Q_i \subset \mathbb{R}^n \setminus E_\lambda$, $i = 1,2,\dots$,
- $\sum_{i=1}^\infty \chi_{2Q_i} \leqslant N < \infty$ and
- $c_1 \operatorname{dist}(Q_i, E_\lambda) \leqslant \operatorname{diam}(Q_i) \leqslant c_2 \operatorname{dist}(Q_i, E_\lambda)$ for some constants c_1 and c_2.

For the Whitney decomposition, see [17], pages 167–170.

Then we construct a partition of unity associated with the covering $2Q_i$, $i = 1,2,\dots$ This can be done in two steps. First, let $\varphi_i \in C_0^\infty(2Q_i)$ be a cutoff function with $0 \leqslant \varphi_i \leqslant 1$, $\varphi_i = 1$ in Q_i and

$$|D\varphi_i| \leqslant \frac{c}{\operatorname{diam}(Q_i)},$$

for $i = 1,2,\dots$ Then let

$$\phi_i(x) = \frac{\varphi_i(x)}{\sum_{j=1}^{\infty} \varphi_j(x)}$$

for every $i = 1,2,\dots$. Observe that the sum is over finitely many terms only since $\varphi_i \in C_0^\infty(2Q_i)$ and the cubes $2Q_i$, $i = 1,2,\dots$, are of bounded overlap. The functions ϕ_i have the property

$$\sum_{i=1}^{\infty} \phi_i(x) = \chi_{\mathbb{R}^n \setminus E_\lambda}(x)$$

for every $x \in \mathbb{R}^n$.

We define the function u_λ by

$$u_\lambda(x) = \begin{cases} u(x), & x \in E_\lambda, \\ \sum_{i=1}^{\infty} \phi_i(x) u_{2Q_i}, & x \in \mathbb{R}^n \setminus E_\lambda. \end{cases}$$

The function u_λ is a Whitney type extension of $u|_{E_\lambda}$ to the set $\mathbb{R}^n \setminus E_\lambda$.

THE MORAL: The extension is defined by integral averages in Whitney cubes.

Since the cubes $2Q_i$, $i = 1,2,\dots$, are of bounded overlap, the function u_λ is locally a finite linear combination of smooth functions in $\mathbb{R}^n \setminus E_\lambda$ and thus $u_\lambda \in C^\infty(\mathbb{R}^n \setminus E_\lambda)$.

Claim: There exists a constant $c = c(n)$ such that u_λ is $c\lambda$-Lipschitz continuous on $\mathbb{R}^n$, that is,

$$|u_\lambda(x) - u_\lambda(y)| \leqslant c\lambda|x-y|$$

for every $x, y \in \mathbb{R}^n$.

Reason. **Step 1:** For every $x \in \mathbb{R}^n \setminus E_\lambda$ there exists $\overline{x} \in E_\lambda$ such that $|\overline{x} - x| \leqslant 2\operatorname{dist}(x, E_\lambda)$. Then

$$\begin{aligned} |u_\lambda(\overline{x}) - u_\lambda(x)| &= \left| u(\overline{x}) - \sum_{i=1}^{\infty} \phi_i(x) u_{2Q_i} \right| = \left| u(\overline{x}) \sum_{i=1}^{\infty} \phi_i(x) - \sum_{i=1}^{\infty} \phi_i(x) u_{2Q_i} \right| \\ &= \left| \sum_{i=1}^{\infty} \phi_i(x)(u(\overline{x}) - u_{2Q_i}) \right| \leqslant \sum_{i \in I_x} |\phi_i(x)(u(\overline{x}) - u_{2Q_i})| \\ &\leqslant \sum_{i \in I_x} |u(\overline{x}) - u_{2Q_i}|, \end{aligned}$$

where $i \in I_x$ if and only if $x \in 2Q_i$. By the properties of the Whitney decomposition, there exists a constant $\alpha > 0$ such that $2Q_i \subset B(\overline{x}, \alpha l(Q_i)) = B(\overline{x}, r_i)$ for every $i \in I_x$.

As in the proof of Theorem 5.48, we obtain

$$\begin{aligned}
|u(\overline{x}) - u_{2Q_i}| &\leqslant |u(\overline{x}) - u_{B(\overline{x},r_i)}| + |u_{B(\overline{x},r_i)} - u_{2Q_i}| \\
&\leqslant c r_i u_1^{\#}(\overline{x}) + \left| \fint_{2Q_i} (u(x) - u_{B(\overline{x},r_i)})\,dx \right| \\
&\leqslant c r_i u_1^{\#}(\overline{x}) + \fint_{2Q_i} |u(x) - u_{B(\overline{x},r_i)}|\,dx \\
&\leqslant c r_i u_1^{\#}(\overline{x}) + \frac{|B(\overline{x},r_i)|}{|2Q_i|} \fint_{B(\overline{x},r_i)} |u(x) - u_{B(\overline{x},r_i)}|\,dx \\
&\leqslant c r_i u_1^{\#}(\overline{x}) + c r_i \frac{1}{r_i} \fint_{B(\overline{x},r_i)} |u(x) - u_{B(\overline{x},r_i)}|\,dx \\
&\leqslant c r_i u_1^{\#}(\overline{x}).
\end{aligned} \tag{5.62}$$

Since $\overline{x} \in E_\lambda$, by (5.52) we have

$$u_1^{\#}(\overline{x}) \leqslant cM|Du|(\overline{x}) \leqslant c\lambda.$$

This implies that

$$|u(\overline{x}) - u_{2Q_i}| \leqslant c r_i \lambda = c\lambda l(Q_i) \leqslant c\lambda \operatorname{dist}(x, E_\lambda) \leqslant c\lambda|\overline{x} - x|$$

for every $i \in I_x$. Since the cubes $2Q_i$, $i = 1, 2, \dots$, have bounded overlap, the cardinality of I_x is uniformly bounded and we obtain

$$|u_\lambda(\overline{x}) - u_\lambda(x)| \leqslant c\lambda|\overline{x} - x| \tag{5.63}$$

for every $x \in \mathbb{R}^n \setminus E_\lambda$.

Step 2: Let $x, y \in \mathbb{R}^n \setminus E_\lambda$ and let

$$\gamma = \max\{\operatorname{dist}(x, E_\lambda), \operatorname{dist}(y, E_\lambda)\}. \tag{5.64}$$

If $|x - y| \geqslant \gamma$, by (5.63) we have

$$\begin{aligned}
|u_\lambda(x) - u_\lambda(y)| &\leqslant |u_\lambda(x) - u_\lambda(\overline{x})| + |u_\lambda(\overline{x}) - u_\lambda(\overline{y})| + |u_\lambda(\overline{y}) - u_\lambda(y)| \\
&\leqslant c\lambda|x - \overline{x}| + |u(\overline{x}) - u(\overline{y})| + c\lambda|y - \overline{y}|.
\end{aligned}$$

Since $x, y \in \mathbb{R}^n \setminus E_\lambda$, Theorem 5.37 implies

$$|u(\overline{x}) - u(\overline{y})| \leqslant c|\overline{x} - \overline{y}|(M|Du|(\overline{x}) + M|Du|(\overline{y})) \leqslant c\lambda|\overline{x} - \overline{y}|$$

for almost every $\overline{x}, \overline{y}$ and thus

$$|u_\lambda(x) - u_\lambda(y)| \leqslant c\lambda\big(|x - \overline{x}| + |\overline{x} - \overline{y}| + |y - \overline{y}|\big),$$

where $|\overline{x} - \overline{y}| \leqslant |\overline{x} - x| + |x - y| + |y - \overline{y}|$. It follows that

$$\begin{aligned}
|u_\lambda(x) - u_\lambda(y)| &\leqslant c\lambda\big(|x - \overline{x}| + |x - y| + |y - \overline{y}|\big) \\
&\leqslant c\lambda|x - y|
\end{aligned}$$

for every $x, y \in \mathbb{R}^n \setminus E_\lambda$ with $|x-y| \geqslant \gamma$.

Step 3: Let $x, y \in \mathbb{R}^n \setminus E_\lambda$ with $|x-y| < \gamma$. Since

$$\sum_{i=1}^{\infty}(\phi_i(x)-\phi_i(y)) = 0 \quad \text{and} \quad |\phi_i(x)-\phi_i(y)| \leqslant \frac{c}{l(Q_i)}|x-y|$$

for every $x, y \in \mathbb{R}^n \setminus E_\lambda$, we have

$$\begin{aligned} |u_\lambda(x)-u_\lambda(y)| &= \left|\sum_{i=1}^{\infty}\phi_i(x)u_{2Q_i} - \sum_{i=1}^{\infty}\phi_i(y)u_{2Q_i}\right| \\ &= \left|\sum_{i=1}^{\infty}(\phi_i(x)-\phi_i(y))u_{2Q_i}\right| \\ &= \left|\sum_{i=1}^{\infty}(\phi_i(x)-\phi_i(y))(u(\overline{x})-u_{2Q_i})\right| \\ &\leqslant c|x-y|\sum_{i\in I_x\cup I_y}\frac{1}{l(Q_i)}|u(\overline{x})-u_{2Q_i}|. \end{aligned} \tag{5.65}$$

By the properties of the Whitney decomposition and since $|x-y| < \gamma$, there exists a constant $\alpha > 0$ such that $2Q_i \subset B(\overline{x}, \alpha l(Q_i)) = B(\overline{x}, r_i)$ for every $i \in I_x \cup I_y$. As in (5.62), we obtain

$$|u(\overline{x})-u_{2Q_i}| \leqslant cr_i u_1^{\#}(\overline{x}) \leqslant cr_i M|Du|(\overline{x}) \leqslant c\lambda l(Q_i) \tag{5.66}$$

for every $i \in I_x \cup I_y$. Since the cardinality of $I_x \cup I_y$ is uniformly bounded, we obtain

$$|u_\lambda(x)-u_\lambda(y)| \leqslant c\lambda|x-y|$$

for every $x, y \in \mathbb{R}^n \setminus E_\lambda$ with $|x-y| < \gamma$.

Step 4: Step 2 and Step 3 imply

$$|u_\lambda(x)-u_\lambda(y)| \leqslant c\lambda|x-y|$$

for every $x, y \in \mathbb{R}^n \setminus E_\lambda$. This shows that u_λ is $c\lambda$-Lipschitz continuous in $\mathbb{R}^n \setminus E_\lambda$. For almost every $x, y \in E_\lambda$, Theorem 5.37 implies

$$|u_\lambda(x)-u_\lambda(y)| = |u(x)-u(y)| \leqslant c|x-y|(M|Du|(x)+M|Du|(y)) \leqslant c\lambda|x-y|.$$

Since the points for which this holds true are dense in E_λ, we conclude that u_λ is Lipschitz on E_λ.

If $x \in \mathbb{R}^n \setminus E_\lambda$ and $y \in E_\lambda$, we have

$$|u_\lambda(x)-u_\lambda(y)| \leqslant |u_\lambda(x)-u_\lambda(\overline{x})| + |u_\lambda(\overline{x})-u_\lambda(y)|.$$

By (5.63) we have

$$|u_\lambda(x)-u_\lambda(\overline{x})| \leqslant c\lambda|x-\overline{x}| \leqslant c\lambda \operatorname{dist}(x, E_\lambda) \leqslant c\lambda|x-y|$$

and by Theorem 5.37 we have

$$\begin{aligned}|u_\lambda(\overline{x})-u_\lambda(y)| = |u(\overline{x})-u(y)| &\leqslant c|\overline{x}-y|(M|Du|(\overline{x})+M|Du|(y))\\ &\leqslant c\lambda|\overline{x}-y| \leqslant c\lambda(|\overline{x}-x|+|x-y|)\\ &\leqslant c\lambda(\operatorname{dist}(x,E_\lambda)+|x-y|) \leqslant c\lambda|x-y|.\end{aligned}$$

It follows that

$$|u_\lambda(x)-u_\lambda(y)| \leqslant c\lambda|x-y|$$

for every $x \in \mathbb{R}^n \setminus E_\lambda$ and $y \in E_\lambda$. This shows that u_λ is $c\lambda$-Lipschitz continuous in $\mathbb{R}^n$. ■

By Theorem 3.31, we have $u_\lambda \in W^{1,\infty}_{\mathrm{loc}}(\mathbb{R}^n)$. In particular, this implies that the weak gradient Du_λ exists and $u_\lambda \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n)$ and we can proceed as in the proof of Theorem 5.61. However, we discuss a maximal function argument to show that $u_\lambda \in W^{1,p}(\mathbb{R}^n)$.

Claim: $u_\lambda \in W^{1,p}(\mathbb{R}^n)$.

Reason. **Step 1:** Let $x,y \in \mathbb{R}^n \setminus E_\lambda$ with $|x-y|<\gamma$, where γ is as in (5.64). By the properties of the Whitney decomposition and since $|x-y|<\gamma$, there exists a constant $\alpha>0$ such that $2Q_i \subset B(x,\alpha l(Q_i)) = B(x,r_i)$ for every $i \in I_x \cup I_y$. As in (5.65) and (5.66), we obtain

$$\begin{aligned}|u_\lambda(x)-u_\lambda(y)| &\leqslant c|x-y| \sum_{i\in I_x\cup I_y} \frac{1}{l(Q_i)}|u(x)-u_{2Q_i}|\\ &\leqslant c|x-y|M|Du|(x)\end{aligned}$$

for every $x,y \in \mathbb{R}^n \setminus E_\lambda$ with $|x-y|<\gamma$.

Step 2: Let $x,y \in \mathbb{R}^n \setminus E_\lambda$ with $|x-y| \geqslant \gamma$. Then and applying

$$\begin{aligned}|u_\lambda(x)-u_\lambda(y)| &= \left|\sum_{i=1}^\infty \phi_i(x)u_{2Q_i} - \sum_{i=1}^\infty \phi_i(y)u_{2Q_i}\right|\\ &= \left|\sum_{i=1}^\infty \big(\phi_i(x)(u_{2Q_i}-u(x)) - \phi_i(y)(u_{2Q_i}-u(y))\big) + (u(x)-u(y))\right|\\ &\leqslant \sum_{i\in I_x}|u(x)-u_{2Q_i}| + \sum_{i\in I_y}|u(y)-u_{2Q_i}| + |u(x)-u(y)|\end{aligned}$$

As in (5.66), we obtain

$$\begin{aligned}\sum_{i\in I_x}|u(x)-u_{2Q_i}| &\leqslant c\sum_{i\in I_x} l(Q_i)M|Du|(x) \leqslant c\operatorname{dist}(x,E_\lambda)M|Du|(x)\\ &\leqslant c\gamma M|Du|(x) \leqslant c|x-y|M|Du|(x)\end{aligned}$$

and similarly

$$\sum_{i\in I_y}|u(y)-u_{2Q_i}| \leqslant c|x-y|M|Du|(y).$$

By Theorem 5.37 we have

$$|u(x)-u(y)| \leqslant c|x-y|(M|Du|(x)+M|Du|(y))$$

and we conclude that

$$|u_\lambda(x)-u_\lambda(y)| \leqslant c|x-y|(M|Du|(x)+M|Du|(y))$$

for every $x, y \in \mathbb{R}^n \setminus E_\lambda$ with $|x-y| \geqslant \gamma$. Step 2 and Step 3 show that

$$|u_\lambda(x)-u_\lambda(y)| \leqslant c|x-y|(M|Du|(x)+M|Du|(y))$$

for every $x, y \in \mathbb{R}^n \setminus E_\lambda$.

Step 3: Let $x \in E_\lambda$ and $y \in \mathbb{R}^n \setminus E_\lambda$. Then

$$|u_\lambda(x)-u_\lambda(y)| = |u(x)-u_\lambda(y)| = \left|\sum_{i=1}^{\infty} \phi_i(y)(u(x)-u_{2Q_i})\right| \leqslant \sum_{i\in I_y} |u(x)-u_{2Q_i}|.$$

As above, we obtain

$$\begin{aligned}|u(x)-u_{2Q_i}| &\leqslant |u(x)-u(y)|+|u(y)-u_{2Q_i}| \\ &\leqslant c|x-y|(M|Du|(x)+M|Du|(y))+cl(Q_i)M|Du|(y)\end{aligned}$$

for every $i \in I_y$. Since $l(Q_i) \leqslant c\,\mathrm{dist}(y,E_\lambda)$ and $\mathrm{dist}(y,E_\lambda) \leqslant |x-y|$, we obtain

$$|u_\lambda(x)-u_\lambda(y)| \leqslant c|x-y|(M|Du|(x)+M|Du|(y))$$

for every $x \in E_\lambda$ and $y \in \mathbb{R}^n \setminus E_\lambda$.

Step 4: By Step 1, Step 2 and Step 3, we have

$$|u_\lambda(x)-u_\lambda(y)| \leqslant c|x-y|(M|Du|(x)+M|Du|(y))$$

for almost every $x, y \in \mathbb{R}^n$. Since $|Du| \in L^p(\mathbb{R}^n)$ with $1 < p < \infty$, the Hardy–Littlewood–Wiener theorem, see Theorem 5.3 (2) , implies that $M|Du| \in L^p(\mathbb{R}^n)$. It follows from Theorem 5.57 that $u \in W^{1,p}(\mathbb{R}^n)$. ■

Claim: There exists a constant $c = c(n,p)$ such that

$$\|u_\lambda\|_{W^{1,p}(\mathbb{R}^n\setminus E_\lambda)} \leqslant c\|u\|_{W^{1,p}(\mathbb{R}^n\setminus E_\lambda)}.$$

Reason. Since the cubes $2Q_i$, $i = 1,2,\dots$, are of bounded overlap, we have

$$\begin{aligned}\int_{\mathbb{R}^n\setminus E_\lambda} |u_\lambda|^p\,dx &= \int_{\mathbb{R}^n\setminus E_\lambda} \left|\sum_{i=1}^{\infty}\phi_i(x)u_{2Q_i}\right|^p dx \leqslant c\sum_{i=1}^{\infty}\int_{2Q_i}|u_{2Q_i}|^p\,dx \\ &\leqslant c\sum_{i=1}^{\infty}|2Q_i| \fint_{2Q_i}|u|^p\,dx \leqslant c\int_{\mathbb{R}^n\setminus E_\lambda}|u|^p\,dx.\end{aligned}$$

Then we consider an estimate for the gradient. We recall that

$$\Phi(x) = \sum_{i=1}^{\infty}\phi_i(x) = 1$$

for every $x \in \mathbb{R}^n \setminus E_\lambda$. Since the cubes $2Q_i$, $i = 1,2,\dots$, are of bounded overlap, we see that $\Phi \in C^\infty(\mathbb{R}^n \setminus E_\lambda)$ and

$$D_j\Phi(x) = \sum_{i=1}^{\infty} D_j\phi_i(x) = 0, \quad j = 1,2,\dots,n,$$

for every $x \in \mathbb{R}^n \setminus E_\lambda$. Hence we obtain

$$\begin{aligned}|D_j u_\lambda(x)| &= \Big|\sum_{i=1}^{\infty} D_j\phi_i(x) u_{2Q_i}\Big| = \Big|\sum_{i=1}^{\infty} D_j\phi_i(x)(u(x) - u_{2Q_i})\Big| \\ &\leqslant c\sum_{i=1}^{\infty} \operatorname{diam}(Q_i)^{-1}|u(x) - u_{2Q_i}|\chi_{2Q_i}(x)\end{aligned}$$

and consequently

$$|D_j u_\lambda(x)|^p \leqslant c\sum_{i=1}^{\infty} \operatorname{diam}(Q_i)^{-p}|u(x) - u_{2Q_i}|^p \chi_{2Q_i}(x)$$

for every $x \in \mathbb{R}^n \setminus E_\lambda$. Here we again used the fact that the cubes $2Q_i$, $i = 1,2,\dots$, are of bounded overlap. By applying the Poincaré inequality, we have

$$\begin{aligned}\int_{\mathbb{R}^n\setminus E_\lambda} |D_j u_\lambda|^p\,dx &\leqslant c\int_{\mathbb{R}^n\setminus E_\lambda} \Big(\sum_{i=1}^{\infty} \operatorname{diam}(Q_i)^{-p}|u - u_{2Q_i}|^p\chi_{2Q_i}\Big)\,dx \\ &\leqslant \sum_{i=1}^{\infty}\int_{2Q_i} \operatorname{diam}(Q_i)^{-p}|u - u_{2Q_i}|^p\,dx \\ &\leqslant c\sum_{i=1}^{\infty}\int_{2Q_i} |Du|^p\,dx \leqslant c\int_{\mathbb{R}^n\setminus E_\lambda} |Du|^p\,dx\end{aligned}$$

for every $j = 1,2,\dots,n$. ■

In particular, the previous claim implies that

$$\begin{aligned}\|u - u_\lambda\|_{W^{1,p}(\mathbb{R}^n)} &= \|u - u_\lambda\|_{W^{1,p}(\mathbb{R}^n\setminus E_\lambda)} \\ &\leqslant \|u\|_{W^{1,p}(\mathbb{R}^n\setminus E_\lambda)} + \|u_\lambda\|_{W^{1,p}(\mathbb{R}^n\setminus E_\lambda)} \\ &\leqslant c\|u\|_{W^{1,p}(\mathbb{R}^n\setminus E_\lambda)}.\end{aligned}$$

As in the proof of Theorem 5.61, we have $|\mathbb{R}^n \setminus E_\lambda| \to 0$ as $\lambda \to \infty$. This implies that $u_\lambda \to u$ in $W^{1,p}(\mathbb{R}^n)$ as $\lambda \to \infty$.

Remark 5.67. We know that $u_\lambda \in W^{1,p}(\mathbb{R}^n \setminus E_\lambda)$ and that it is Lipschitz continuous in $\mathbb{R}^n$. Moreover $u \in W^{1,p}(\mathbb{R}^n)$ and $u = u_\lambda$ in E_λ. This implies that $w = u - u_\lambda \in W^{1,p}(\mathbb{R}^n \setminus E_\lambda)$ and that $w = 0$ in E_λ. By the ACL property, u is absolutely continuous on almost every line segment parallel to the coordinate axes. Take any such line. Now w is absolutely continuous on the part of the line segment which intersects $\mathbb{R}^n \setminus E_\lambda$. On the other hand $w = 0$ in the complement of E_λ. Hence the continuity of w in the line segment implies that w is absolutely continuous on the whole line segment. By the ACL characterization of Sobolev spaces, see Theorem 2.36, we may conclude that $w \in W^{1,p}(\mathbb{R}^n)$.

THE END

Bibliography

[1] D.R. Adams and L.I. Hedberg, Function spaces and potential theory, Springer 1996.

[2] R.A. Adams and J.J.F. Fournier, Sobolev spaces (second edition), Academic Press 2003.

[3] L. Diening, P. Harjulehto, P. Hästö and M. Ružička, Lebesgue and Sobolev spaces with variable exponents, Springer 2011.

[4] V.I. Bogachev, Measure Theory 1 and 2, Springer, 2007.

[5] A. Björn and J. Björn, Nonlinear potential theory on metric spaces, European Mathematical Society (EMS) 2011.

[6] E. Di Nezza, G. Palatucci and E. Valdinoci, Hitchhiker's guide to the fractional Sobolev spaces, Bull. Sci. Math. 136 (2012), no. 5, 521–573.

[7] L.C. Evans, Partial differential equations, 2nd edition, American Mathematical Society 2010.

[8] L.C. Evans and R.F. Gariepy, Measure theory and fine properties of functions, CRC Press 1992.

[9] P. Hajłasz, Sobolev inequalities, truncation method, and John domains. Papers on analysis, 109–126, Rep. Univ. Jyväskylä Dep. Math. Stat., 83, 2001.

[10] L. Diening, P. Harjulehto, P. Hästö, Peter and M. Ruzicka, Lebesgue and Sobolev spaces with variable exponents, Springer 2011.

[11] J. Heinonen, Lectures on analysis on metric spaces, Springer 2001.

[12] J. Heinonen, T. Kilpeläinen and O. Martio, Nonlinear potential theory of degenerate elliptic equations. Unabridged republication of the 1993 original. Dover Publications, 2006.

[13] J. Heinonen, P. Koskela, N. Shanmugalingam and J. Tyson, Sobolev spaces on metric measure spaces. An approach based on upper gradients, Cambridge University Press 2015.

[14] G. Leoni, A first course in Sobolev spaces, 2nd edition, American Mathematical Society 2017.

[15] J. Kinnunen, J. Lehrbäck and A. Vähäkangas, Maximal function methods for Sobolev spaces, American Mathematical Society 2021.

[16] V. Maz'ya, Sobolev spaces with applications to elliptic partial differential equation, 2nd edition, Springer 2011.

[17] E. M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton University Press, 1970.

[18] W.P. Ziemer, Weak differentiable functions, Springer 1989.