10. Sobolev embedding

The properties of Sobolev spaces are briefly discussed above. If m is a positive integer then $u \in H^m(\mathbb{R}^n)$ 'means' that u has up to m derivatives in $L^2(\mathbb{R}^n)$. The question naturally arises as to the sense in which these 'weak' derivatives correspond to old-fashioned 'strong' derivatives. Of course when m is not an integer it is a little harder to imagine what these 'fractional derivatives' are. However the main result is:

Theorem 10.1 (Sobolev embedding)**.** *If $u \in H^m(\mathbb{R}^n)$ where $m > n/2$ then $u \in \mathcal{C}^0_0(\mathbb{R}^n)$, i.e.,*

$$H^m(\mathbb{R}^n) \subset \mathcal{C}^0_0(\mathbb{R}^n)\,,\ m > n/2\,. \tag{10.1}$$

Proof. By definition, $u \in H^m(\mathbb{R}^n)$ means $v \in \mathcal{S}'(\mathbb{R}^n)$ and $\langle\xi\rangle^m \hat{u}(\xi) \in L^2(\mathbb{R}^n)$. Suppose first that $u \in \mathcal{S}(\mathbb{R}^n)$. The Fourier inversion formula shows that

$$\begin{aligned}(2\pi)^n\,|u(x)| &= \left|\int e^{ix\cdot\xi}\hat{u}(\xi)\,d\xi\right| \\ &\le \left(\int_{\mathbb{R}^n}\langle\xi\rangle^{2m}\,|\hat{u}(\xi)|^2\,d\xi\right)^{1/2}\cdot\left(\sum_{\mathbb{R}^n}\langle\xi\rangle^{-2m}\,d\xi\right)^{1/2}.\end{aligned}$$

Now, if $m > n/2$ then the second integral is finite. Since the first integral is the norm on $H^m(\mathbb{R}^n)$ we see that

$$\sup_{\mathbb{R}^n}|u(x)| = \|u\|_{L^\infty} \le (2\pi)^{-n}\|u\|_{H^m}\,,\ m > n/2\,. \tag{10.2}$$

This is all for $u \in \mathcal{S}(\mathbb{R}^n)$, but $\mathcal{S}(\mathbb{R}^n) \hookrightarrow H^m(\mathbb{R}^n)$ is dense. The estimate (10.2) shows that if $u_j \to u$ in $H^m(\mathbb{R}^n)$, with $u_j \in \mathcal{S}(\mathbb{R}^n)$, then $u_j \to u'$ in $\mathcal{C}^0_0(\mathbb{R}^n)$. In fact $u' = u$ in $\mathcal{S}'(\mathbb{R}^n)$ since $u_j \to u$ in $L^2(\mathbb{R}^n)$ and $u_j \to u'$ in $\mathcal{C}^0_0(\mathbb{R}^n)$ both imply that $\int u_j\varphi$ converges, so

$$\int_{\mathbb{R}^n} u_j\varphi \to \int_{\mathbb{R}^n} u\varphi = \int_{\mathbb{R}^n} u'\varphi\ \forall\ \varphi \in \mathcal{S}(\mathbb{R}^n).$$

□

Notice here the precise meaning of $u = u'$, $u \in H^m(\mathbb{R}^n) \subset L^2(\mathbb{R}^n)$, $u' \in \mathcal{C}^0_0(\mathbb{R}^n)$. When identifying $u \in L^2(\mathbb{R}^n)$ with the corresponding tempered distribution, the values on any set of measure zero 'are lost'. Thus as *functions* (10.1) means that each $u \in H^m(\mathbb{R}^n)$ has a representative $u' \in \mathcal{C}^0_0(\mathbb{R}^n)$.

We can extend this to higher derivatives by noting that

Proposition 10.2. *If $u \in H^m(\mathbb{R}^n)$, $m \in \mathbb{R}$, then $D^\alpha u \in H^{m-|\alpha|}(\mathbb{R}^n)$ and*

$$D^\alpha : H^m(\mathbb{R}^n) \to H^{m-|\alpha|}(\mathbb{R}^n) \tag{10.3}$$

is continuous.

Proof. First it is enough to show that each D_j defines a continuous linear map

$$D_j : H^m(\mathbb{R}^n) \to H^{m-1}(\mathbb{R}^n) \ \forall \ j \tag{10.4}$$

since then (10.3) follows by composition.

If $m \in \mathbb{R}$ then $u \in H^m(\mathbb{R}^n)$ means $\hat{u} \in \langle\xi\rangle^{-m}L^2(\mathbb{R}^n)$. Since $\widehat{D_j u} = \xi_j \cdot \hat{u}$, and

$$|\xi_j| \, \langle\xi\rangle^{-m} \le C_m \langle\xi\rangle^{-m+1} \ \forall \ m$$

we conclude that $D_j u \in H^{m-1}(\mathbb{R}^n)$ and

$$\|D_j u\|_{H^{m-1}} \le C_m \|u\|_{H^m} \, .$$

□

Applying this result we see

Corollary 10.3. *If $k \in \mathbb{N}_0$ and $m > \frac{n}{2} + k$ then*

$$H^m(\mathbb{R}^n) \subset \mathcal{C}_0^k(\mathbb{R}^n) \, . \tag{10.5}$$

Proof. If $|\alpha| \le k$, then $D^\alpha u \in H^{m-k}(\mathbb{R}^n) \subset \mathcal{C}_0^0(\mathbb{R}^n)$. Thus the 'weak derivatives' $D^\alpha u$ are continuous. Still we have to check that this means that u is itself k times continuously differentiable. In fact this again follows from the density of $\mathcal{S}(\mathbb{R}^n)$ in $H^m(\mathbb{R}^n)$. The continuity in (10.3) implies that if $u_j \to u$ in $H^m(\mathbb{R}^n)$, $m > \frac{n}{2} + k$, then $u_j \to u'$ in $\mathcal{C}_0^k(\mathbb{R}^n)$ (using its completeness). However $u = u'$ as before, so $u \in \mathcal{C}_0^k(\mathbb{R}^n)$. □

In particular we see that

$$H^\infty(\mathbb{R}^n) = \bigcap_m H^m(\mathbb{R}^n) \subset \mathcal{C}^\infty(\mathbb{R}^n) \, . \tag{10.6}$$

These functions are not in general Schwartz test functions.

Proposition 10.4. *Schwartz space can be written in terms of weighted Sobolev spaces*

$$\mathcal{S}(\mathbb{R}^n) = \bigcap_k \langle x\rangle^{-k} H^k(\mathbb{R}^n) \, . \tag{10.7}$$

Proof. This follows directly from (10.5) since the left side is contained in

$$\bigcap_k \langle x\rangle^{-k}\mathcal{C}_0^{k-n}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n).$$

□

Theorem 10.5 (Schwartz representation)**.** *Any tempered distribution can be written in the form of a finite sum*

$$u = \sum_{\substack{|\alpha|\le m\\|\beta|\le m}} x^\alpha D_x^\beta u_{\alpha\beta}\,,\ u_{\alpha\beta}\in\mathcal{C}_0^0(\mathbb{R}^n). \tag{10.8}$$

or *in the form*

$$u = \sum_{\substack{|\alpha|\le m\\|\beta|\le m}} D_x^\beta(x^\alpha v_{\alpha\beta}),\ \ v_{\alpha\beta}\in\mathcal{C}_0^0(\mathbb{R}^n). \tag{10.9}$$

Thus every tempered distribution is a finite sum of derivatives of continuous functions of poynomial growth.

Proof. Essentially by definition any $u\in\mathcal{S}'(\mathbb{R}^n)$ is continuous with respect to *one* of the norms $\|\langle x\rangle^k\varphi\|_{\mathcal{C}^k}$. From the Sobolev embedding theorem we deduce that, with $m > k + n/2$,

$$|u(\varphi)| \le C\|\langle x\rangle^k\varphi\|_{H^m}\ \forall\ \varphi\in\mathcal{S}(\mathbb{R}^n).$$

This is the same as

$$\left|\langle x\rangle^{-k}u(\varphi)\right| \le C\|\varphi\|_{H^m}\ \forall\ \varphi\in\mathcal{S}(\mathbb{R}^n).$$

which shows that $\langle x\rangle^{-k}u\in H^{-m}(\mathbb{R}^n)$, i.e., from Proposition 9.8,

$$\langle x\rangle^{-k}u = \sum_{|\alpha|\le m} D^\alpha u_\alpha\,,\ \ u_\alpha\in L^2(\mathbb{R}^n)\,.$$

In fact, choose $j > n/2$ and consider $v_\alpha \in H^j(\mathbb{R}^n)$ defined by $\hat{v}_\alpha = \langle\xi\rangle^{-j}\hat{u}_\alpha$. As in the proof of Proposition 9.14 we conclude that

$$u_\alpha = \sum_{|\beta|\le j} D^\beta u'_{\alpha,\beta}\,,\ u'_{\alpha,\beta}\in H^j(\mathbb{R}^n)\subset\mathcal{C}_0^0(\mathbb{R}^n)\,.$$

Thus,[17]

$$u = \langle x\rangle^k \sum_{|\gamma|\le M} D_\alpha^\gamma v_\gamma\,,\ \ v_\gamma\in\mathcal{C}_0^0(\mathbb{R}^n)\,. \tag{10.10}$$

To get (10.9) we 'commute' the factor $\langle x\rangle^k$ to the inside; since I have not done such an argument carefully so far, let me do it as a lemma.

[17]This is probably the most useful form of the representation theorem!

Lemma 10.6. *For any $\gamma \in \mathbb{N}_0^n$ there are polynomials $p_{\alpha,\gamma}(x)$ of degrees at most $|\gamma - \alpha|$ such that*

$$\langle x\rangle^k D^\gamma v = \sum_{\alpha\le\gamma} D^{\gamma-\alpha}\left(p_{\alpha,\gamma}\langle x\rangle^{k-2|\gamma-\alpha|}v\right) .$$

Proof. In fact it is convenient to prove a more general result. Suppose p is a polynomial of a degree at most j then there exist polynomials of degrees at most $j + |\gamma - \alpha|$ such that

$$p\langle x\rangle^k D^\gamma v = \sum_{\alpha\le\gamma} D^{\gamma-\alpha}(p_{\alpha,\gamma}\langle x\rangle^{k-2|\gamma-\alpha|}v)\,. \tag{10.11}$$

The lemma follows from this by taking $p = 1$.

Furthermore, the identity (10.11) is trivial when $\gamma = 0$, and proceeding by induction we can suppose it is known whenever $|\gamma| \le L$. Taking $|\gamma| = L + 1$,

$$D^\gamma = D_j D^{\gamma'} \ \ |\gamma'| = L.$$

Writing the identity for γ' as

$$p\langle x\rangle^k D^{\gamma'} = \sum_{\alpha'\le\gamma'} D^{\gamma'-\alpha'}(p_{\alpha',\gamma'}\langle x\rangle^{k-2|\gamma'-\alpha'|}v)$$

we may differentiate with respect to x_j. This gives

$$\begin{aligned} p\langle x\rangle^k D^\gamma &= -D_j(p\langle x\rangle^k)\cdot D^{\gamma'}v \\ &+ \sum_{|\alpha'|\le\gamma} D^{\gamma-\alpha'}(p'_{\alpha',\gamma'}\langle x\rangle^{k-2|\gamma-\alpha|+2}v)\,. \end{aligned}$$

The first term on the right expands to

$$(-(D_jp)\cdot\langle x\rangle^k D^{\gamma'}v - \frac{1}{i}kpx_j\langle x\rangle^{k-2}D^{\gamma'}v)\,.$$

We may apply the inductive hypothesis to each of these terms and rewrite the result in the form (10.11); it is only necessary to check the order of the polynomials, and recall that $\langle x\rangle^2$ is a polynomial of degree 2. □

Applying Lemma 10.6 to (10.10) gives (10.9), once negative powers of $\langle x\rangle$ are absorbed into the continuous functions. Then (10.8) follows from (10.9) and Leibniz's formula. □