Function spaces

Lecture notes, Fall 2008

Mikko Salo

Department of Mathematics and Statistics
University of Helsinki

Contents

CHAPTER 1

Introduction

In mathematical analysis one deals with functions which are differentiable (such as continuously differentiable) or integrable (such as square integrable or L^p). It is often natural to combine the smoothness and integrability requirements, which leads one to introduce various spaces of functions.

This course will give a brief introduction to certain function spaces which are commonly encountered in analysis. This will include Hölder, Lipschitz, Zygmund, Sobolev, Besov, and Triebel-Lizorkin type spaces. We will try to highlight typical uses of these spaces, and will also give an account of interpolation theory which is an important tool in their study.

The first part of the course covered integer order Sobolev spaces in domains in $\mathbf{R}^n$, following Evans [**4**, Chapter 5]. These lecture notes contain the second part of the course. Here the emphasis is on Sobolev type spaces where the smoothness index may be any real number. This second part of the course is more or less self-contained, in that we will use the first part mainly as motivation. Also, we will give new proofs in more general settings of certain results in the first part.

Let us describe the structure of these notes. Interpolation theory is covered in Chapter 2. We begin by proving two classical interpolation results for L^p spaces, namely the Riesz-Thorin and Marcinkiewicz theorems. Next the basic setup of abstract interpolation is discussed. The main part of the chapter includes a development of the real interpolation method in the setting of quasinormed Abelian groups, and interpolation results for Banach-valued L^p spaces. The interpolation results proved here or in Chapter 3 are not the most general available (stronger results are found in Bergh-Löfström [**1**] and Triebel [**13**]). Rather, the point is to give an idea of what interpolation theory is about, without spending too much time on technicalities.

Chapter 3 is an introduction to Sobolev, Besov, Triebel-Lizorkin, and Zygmund spaces in $\mathbf{R}^n$. These spaces are most conveniently defined via the Fourier transform, and the first matter is to review (without proofs) some basic facts about Fourier analysis on tempered distributions. It is then not difficult to define fractional Sobolev spaces using the Bessel potentials. We discuss Littlewood-Paley theory to motivate the definition of Besov and Triebel-Lizorkin spaces, which are then defined in terms of suitable Littlewood-Paley partitions of unity.

Zygmund spaces are defined in terms of finite differences and it is shown that these spaces coincide with the L^∞-based Besov spaces, thus giving a Fourier transform characterization of Hölder and Zygmund spaces. In the course of Chapter 3, we establish the following basic relations between the various spaces:

$$
\begin{aligned}
H^{k,p} &= W^{k,p}, \\
B^s_{22} = F^s_{22} &= H^{s,2}, \\
F^s_{p2} &= H^{s,p}, \\
B^s_{\infty\infty} &= C^s_*.
\end{aligned}
$$

Finally, we discuss embedding theorems for these spaces.

There are several notable omissions in these notes. To retain the real analytic flavor of the course, we do not discuss the complex interpolation method. Also, in the discussion of the real interpolation method, we restrict ourselves to the K-functional and do not consider other related concepts such as the J-functional. In the chapter on Sobolev and related spaces, one of the most basic results is the Mihlin multiplier theorem stating the L^p boundedness of certain multipliers on the Fourier side. We use this result extensively, but do not give a proof of it (or of the related Littlewood-Paley theorem). All the results that we need in this context follow from the vector-valued Mihlin theorem (Theorem 3.4.5), which is a standard consequence of the theory of singular integrals (see for instance [**2**]). Also, we mostly restrict ourselves to L^p based spaces with $1 \le p \le \infty$ or $1 < p < \infty$, and do not discuss Hardy or BMO type spaces.

These notes are intended to be accessible to graduate or advanced undergraduate students having some background in real and functional analysis and multivariable calculus (the real analysis part of Rudin [**8**] should be more than sufficient).

References. We will not follow any particular reference in this course, the main reference will be these lecture notes. Useful books on interpolation theory and function spaces are Bergh-Löfström [**1**] and Triebel [**13**]. Function spaces from a harmonic analysis point of view are discussed in Grafakos [**6**] and Stein [**11**].

Notation. We will use multi-index notation: if $\alpha = (\alpha_1, \ldots, \alpha_n)$ is in $\mathbf{N}^n$ and $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$, we write

$$x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \qquad D^\alpha := D_1^{\alpha_1} \cdots D_n^{\alpha_n},$$

where $D_j := \frac{1}{i} \frac{\partial}{\partial x_j}$. We will also use the Japanese bracket

$$\langle x \rangle := (1 + |x|^2)^{1/2}.$$

The equivalence of two (quasi)norms is written as $\|u\|_1 \sim \|u\|_2$. This means that there exists a constant $C > 0$ such that

$$\frac{1}{C}\|u\|_1 \leq \|u\|_2 \leq C\|u\|_1$$

for all relevant functions u. If A and B are two (quasi)normed spaces or groups, we write $A \subseteq B$ to denote that A is contained in B in the set-theoretic sense, and the inclusion map $i : A \to B$ is continuous. We also write $A = B$ to denote that A and B are equal as sets, and they have equivalent (quasi)norms.

CHAPTER 2

Interpolation theory

2.1. Classical results

We will begin by proving two classical interpolation theorems. The first was proved by Riesz (1926) and Thorin (1938), and it forms the basis for the complex interpolation method. The second is due to Marcinkiewicz (1939) and Zygmund (1956), and it is a precursor of the real interpolation method.

Both these results concern interpolation of L^p spaces. Their nature is illustrated by the following special case: if T is a bounded linear operator $L^{p_0} \to L^{p_0}$ and $L^{p_1} \to L^{p_1}$, where $p_0 < p_1$, then T is also bounded $L^p \to L^p$ for $p_0 \leq p \leq p_1$. Thus, it is enough to establish estimates for two end point values of p, and interpolation will then give estimates for all the intermediate values of p for free.

2.1.1. Riesz-Thorin theorem. We will prove the theorem for the spaces $L^p(\mathbf{R}^n)$, although the same proof applies to $L^p(X, \mu)$ and $L^q(Y, \nu)$ where X, Y are measure spaces and μ, ν are σ-finite positive measures.

THEOREM 1. *(Riesz-Thorin interpolation theorem) Suppose T is a complex-linear map from $L^{p_0} + L^{p_1}$ to $L^{q_0} + L^{q_1}$, $1 \leq p_0, p_1, q_0, q_1 \leq \infty$, such that*

$$\begin{aligned} \|Tf\|_{L^{q_0}} &\leq M_0 \|f\|_{L^{p_0}}, \quad f \in L^{p_0}, \\ \|Tf\|_{L^{q_1}} &\leq M_1 \|f\|_{L^{p_1}}, \quad f \in L^{p_1}. \end{aligned}$$

For $0 < \theta < 1$, define p and q by

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \qquad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

If $1 < q < \infty$ then T maps L^p to L^q, and

$$\|Tf\|_{L^q} \leq M_0^{1-\theta} M_1^{\theta} \|f\|_{L^p}, \quad f \in L^p. \tag{1}$$

PROOF. 1. If $p_0 = p_1 = p$, then (1) follows from the interpolation inequality for L^q norms (that is, the Hölder inequality):

$$\|Tf\|_{L^q} \leq \|Tf\|_{L^{q_0}}^{1-\theta} \|Tf\|_{L^{q_1}}^{\theta} \leq M_0^{1-\theta} M_1^{\theta} \|f\|_{L^p}.$$

Thus we may assume $p_0 \neq p_1$, and so $1 < p < \infty$. Also, we may assume that $M_0, M_1 > 0$ since otherwise one would have $T \equiv 0$.

2. Write $\langle f, g \rangle = \int_{\mathbf{R}^n} fg\,dx$. We claim that it is enough to prove

$$|\langle T(f_0 F^{1/p}), g_0 G^{1/q'} \rangle| \leq M_0^{1-\theta} M_1^{\theta} \tag{2}$$

for any measurable f_0, g_0 with $|f_0|, |g_0| \leq 1$, and for any simple functions $F, G \geq 0$ satisfying $\|F\|_{L^1} = \|G\|_{L^1} = 1$.

In fact, since $1 < q < \infty$, (1) is equivalent with

$$|\langle Tf, g \rangle| \leq M_0^{1-\theta} M_1^{\theta} \|f\|_{L^p} \|g\|_{L^{q'}}.$$

Fix $f \in L^p, g \in L^{q'}$ with $f, g \neq 0$, and let

$$f_0(x) := \begin{cases} f(x)/|f(x)| & \text{if } f(x) \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Define g_0 in a similar way. If $F := |f|^p/\|f\|_{L^p}^p$ and $G := |g|^{q'}/\|g\|_{L^{q'}}^{q'}$, then the last estimate may be written as

$$|\langle T(f_0 F^{1/p}), g_0 G^{1/q'} \rangle| \leq M_0^{1-\theta} M_1^{\theta}.$$

Now, there exist nonnegative simple functions $\tilde{F}_j$ with $\|\tilde{F}_j\|_{L^p} = 1$ such that $\tilde{F}_j \to F^{1/p}$ in L^p. Then $F_j := \tilde{F}_j^p \geq 0$ are simple functions with $\|F_j\|_{L^1} = 1$ and $F_j^{1/p} \to F^{1/p}$ in L^1. Using a similar approximation for G, we see that the theorem will follow from the claim (2).

3. To prove (2), let $|f_0|, |g_0| \leq 1$ and let F, G be simple functions

$$F := \sum_{k=1}^{M} a_k \chi_{A_k}, \qquad G := \sum_{l=1}^{N} b_l \chi_{B_l}$$

where $a_k, b_l > 0$ and $\|F\|_{L^1} = \|G\|_{L^1} = 1$. Define

$$\Phi(z) := \langle T(f_0 F^{\frac{1-z}{p_0} + \frac{z}{p_1}}), g_0 G^{\frac{1-z}{q_0'} + \frac{z}{q_1'}} \rangle,$$

for z in the strip $\{z \in \mathbf{C}\,;\, 0 \leq \operatorname{Re} z \leq 1\}$. Here we define $a^z = e^{z \log a}$ for $a > 0$, so that $|a^z| = a^{\operatorname{Re} z}$ and $|a^{it}| = 1$ for $t \in \mathbf{R}$.

We have

$$\Phi(z) = \sum_{k,l} a_k^{\frac{1-z}{p_0}+\frac{z}{p_1}} b_l^{\frac{1-z}{q_0'}+\frac{z}{q_1'}} \langle T(f_0\chi_{A_k}), g_0\chi_{B_l}\rangle.$$

Thus, Φ is analytic for $0 < \operatorname{Re} z < 1$ and bounded and continuous for $0 \leq \operatorname{Re} z \leq 1$. By Hölder's inequality and using the assumptions on T,

$$|\Phi(it)| \leq \|T(f_0 F^{\frac{1-it}{p_0}+\frac{it}{p_1}})\|_{L^{q_0}} \|g_0 G^{\frac{1-it}{q_0'}+\frac{it}{q_1'}}\|_{L^{q_0'}} \leq M_0$$

and similarly

$$|\Phi(1+it)| \leq \|T(f_0 F^{\frac{-it}{p_0}+\frac{1+it}{p_1}})\|_{L^{q_1}} \|g_0 G^{\frac{-it}{q_0'}+\frac{1+it}{q_1'}}\|_{L^{q_1'}} \leq M_1.$$

The claim (2) follows from the three lines theorem below. □

THEOREM 2. *(Three lines theorem) Assume that Φ is analytic in $0 < \operatorname{Re} z < 1$ and bounded and continuous in $0 \leq \operatorname{Re} z \leq 1$. Suppose that for $t \in \mathbf{R}$,*

$$|\Phi(it)| \leq M_0, \qquad |\Phi(1+it)| \leq M_1.$$

Then $|\Phi(\theta + it)| \leq M_0^{1-\theta} M_1^{\theta}$ for $0 < \theta < 1$ and $t \in \mathbf{R}$.

PROOF. We only consider the case where $M_0, M_1 > 0$. The function

$$\Psi(z) := M_0^{-(1-z)} M_1^{-z} \Phi(z)$$

is analytic in $0 < \operatorname{Re} z < 1$, bounded and continuous in $0 \leq \operatorname{Re} z \leq 1$, and satisfies $|\Psi(it)| \leq 1$ and $|\Psi(1+it)| \leq 1$. It is enough to prove that

$$|\Psi(\theta + it)| \leq 1 \tag{3}$$

for $0 < \theta < 1$.

Now, if $\sup_{0\leq\theta\leq 1} |\Psi(\theta + it)| \to 0$ as $t \to \pm\infty$, one obtains (3) by noting that $|\Psi| \leq 1$ on all four sides of the rectangle $[0, 1] \times [-R, R]$ for R large, thus $|\Psi| \leq 1$ inside the rectangle by the maximum principle.

If Ψ does not decay as $\operatorname{Im} z \to \pm\infty$, we consider the functions

$$\Psi_\varepsilon(z) := \exp(\varepsilon(z^2 - 1))\Psi(z), \quad \varepsilon > 0.$$

Since $(s+it)^2 = s^2 - t^2 + 2ist$, we have $|\Psi_\varepsilon(it)| \leq 1$ and $|\Psi_\varepsilon(1+it)| \leq 1$ and also $|\Psi_\varepsilon(z)| \to 0$ as $\operatorname{Im} z \to \pm\infty$. Thus (3) holds for Ψ_ε. Since

$$|\Psi_\varepsilon(\theta + it)| = e^{\varepsilon(\theta^2 - t^2 - 1)} |\Psi(\theta + it)|$$

where $\varepsilon > 0$ is arbitrary, we obtain the result by letting $\varepsilon \to 0$. □

A standard application of the Riesz-Thorin interpolation theorem is the Hausdorff-Young inequality for Fourier transforms. If $f \in L^1(\mathbf{R}^n)$, the Fourier transform of f is the function $\mathscr{F} f = \hat{f} \in L^\infty(\mathbf{R}^n)$ given by

$$\hat{f}(\xi) = \int_{\mathbf{R}^n} e^{-ix\cdot\xi} f(x)\, dx, \quad \xi \in \mathbf{R}^n.$$

It immediately follows that

$$\|\hat{f}\|_{L^\infty} \leq \|f\|_{L^1}. \tag{4}$$

Now, if $f \in L^1 \cap L^2$, the Plancherel (or Parseval) identity states that $\hat{f} \in L^2$ and

$$\|\hat{f}\|_{L^2} = (2\pi)^{n/2} \|f\|_{L^2}. \tag{5}$$

This can be used to extend the Fourier transform to any L^2 function, and the Plancherel identity remains valid.

THEOREM 3. *(Hausdorff-Young inequality) If $1 \leq p \leq 2$ then the Fourier transform maps $L^p(\mathbf{R}^n)$ into $L^{p'}(\mathbf{R}^n)$, and*

$$\|\hat{f}\|_{L^{p'}} \leq (2\pi)^{n/p'} \|f\|_{L^p}.$$

PROOF. We saw above that the Fourier transform is a well defined map $L^1 \to L^\infty$ and $L^2 \to L^2$, thus it maps $L^1 + L^2$ into $L^2 + L^\infty$. We have the norm bounds (4) and (5). Let $0 < \theta < 1$, and define p by

$$\frac{1}{p} = \frac{1-\theta}{1} + \frac{\theta}{2}.$$

Then $\theta = 2/p'$ and conseqently $1/p' = \theta/2 = (1-\theta)/\infty + \theta/2$.

Now, the Riesz-Thorin theorem implies that the Fourier transform maps L^p to $L^{p'}$ with the right norm bound. Since any p with $1 < p < 2$ can be obtained in this way, the result follows. □

REMARK. (Hausdorff-Young for Fourier series) The Fourier coefficients are given by the map $L^2((0, 2\pi)) \to l^2, f \mapsto \hat{f}$, where

$$\hat{f}(k) = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx}\, dx, \quad k \in \mathbf{Z}.$$

One has the norm bounds

$$\|\hat{f}\|_{l^\infty} \leq (2\pi)^{-1} \|f\|_{L^1((0,2\pi))},$$
$$\|\hat{f}\|_{l^2} = (2\pi)^{-1/2} \|f\|_{L^2((0,2\pi))},$$

the second bound coming from Parseval's theorem. The Riesz-Thorin theorem applied to L^p and l^p spaces gives that

$$\|\hat{f}\|_{l^{p'}} \leq (2\pi)^{-1/p}\|f\|_{L^p((0,2\pi))}$$

for $1 \leq p \leq 2$.

2.1.2. Marcinkiewicz theorem. The Riesz-Thorin theorem implies that if a linear operator is known to be bounded on L^{p_0} and L^{p_1}, then it is bounded on L^p for any p between p_0 and p_1. The Marcinkiewicz theorem gives a similar conclusion, but here it is enough to have *weak type* bounds at the endpoints p_0 and p_1 to obtain strong bounds for intermediate p. The weak type bounds are given in terms of the *distribution function*

$$m(|f| > \lambda) := m(\{x \in \mathbf{R}^n\,;\, |f(x)| > \lambda\})$$

where m is the Lebesgue measure on $\mathbf{R}^n$.

Motivation. Let $f \in L^p(\mathbf{R}^n)$. The elementary identity

$$a^p = \int_0^a p\lambda^{p-1}\,d\lambda, \quad a \geq 0,$$

implies

$$|f(x)|^p = \int_0^\infty p\lambda^{p-1}\chi_{\{|f|>\lambda\}}(x)\,d\lambda.$$

Integrating over x, we obtain

$$\int_{\mathbf{R}^n} |f|^p\,dx = \int_0^\infty p\lambda^{p-1}m(|f| > \lambda)\,d\lambda. \tag{6}$$

This indicates why distribution functions might be useful in studying L^p norms. Also, if $T : L^p(\mathbf{R}^n) \to L^q(\mathbf{R}^n)$ is a bounded linear operator, then

$$m(|Tf| > \lambda) = \int_{\{|Tf|>\lambda\}} dx \leq \frac{1}{\lambda^q}\int_{\{|Tf|>\lambda\}} |Tf|^q\,dx \leq \frac{C\|f\|_{L^p}^q}{\lambda^q}.$$

In particular, if T is the identity operator, then

$$\lambda m(|f| > \lambda)^{1/p} \leq \|f\|_{L^p}, \quad f \in L^p.$$

Definition. The *weak L^p space*, written $L^{p\infty}(\mathbf{R}^n)$ $(0 < p < \infty)$, consists of those measurable functions on $\mathbf{R}^n$ for which the expression

$$\|f\|_{L^{p\infty}} := \sup_{\lambda>0} \lambda m(|f| > \lambda)^{1/p}$$

is finite. We define $L^{\infty\infty} = L^\infty$.

We have $L^p \subseteq L^{p\infty} \subseteq L^{p-\varepsilon}_{\text{loc}}$ if $p-\varepsilon > 0$ (exercise). The function $f(x) = |x|^{-\alpha}$, $\alpha > 0$, is in weak L^p iff $p = n/\alpha$, since

$$m(|f| > \lambda) = m(|x| < \lambda^{-1/\alpha}) = C_n \lambda^{-n/\alpha}$$

and $\lambda m(|f| > \lambda)^{1/p} = C_{n,p}\lambda^{1-\frac{n}{\alpha p}}$ is uniformly bounded iff $\alpha p = n$. Note that $\|\cdot\|_{L^{p\infty}}$ is not a norm if $1 \le p < \infty$.

Definition. Let T be an operator (not necessarily linear) from $L^p(\mathbf{R}^n)$ to the space of measurable functions on $\mathbf{R}^n$. We say that

(a) T is *subadditive* if for almost every $x \in \mathbf{R}^n$,

$$|T(f+g)(x)| \le |Tf(x)| + |Tg(x)|,$$

(b) T is *strong type* (p,q) if it maps $L^p(\mathbf{R}^n)$ into $L^q(\mathbf{R}^n)$ and

$$\|Tf\|_{L^q} \le C\|f\|_{L^p}, \quad f \in L^p, \quad \text{and}$$

(c) T is *weak type* (p,q), $q < \infty$, if

$$m(|Tf| > \lambda) \le \left(\frac{C\|f\|_{L^p}}{\lambda}\right)^q, \quad f \in L^p,\ \lambda > 0.$$

Further, T is *weak type* (p,∞) if it is strong type (p,∞).

From the remarks above it follows that any operator of strong type (p,q) is also of weak type (p,q), and that a weak type (p,q) operator maps L^p into weak L^q. The interpolation result is as follows. Note that as in the Riesz-Thorin theorem, we could easily replace $\mathbf{R}^n$ by more general measure spaces on the right and on the left.

Theorem 4. *(Marcinkiewicz interpolation theorem) Suppose that $1 \le p_0 < p_1 \le \infty$, and let T be a subadditive operator from $L^{p_0} + L^{p_1}$ to the space of measurable functions on $\mathbf{R}^n$. Assume that T is weak type (p_0,p_0) and (p_1,p_1). Then T is strong type (p,p) for $p_0 < p < p_1$.*

Proof. 1. First assume $p_1 < \infty$. The proof is based on the idea that any $f \in L^p$, $p_0 < p < p_1$, may be decomposed as $f = f_0 + f_1$ where $f_0 \in L^{p_0}$ and $f_1 \in L^{p_1}$, by taking

$$f_0(x) := \begin{cases} f(x) & \text{if } |f(x)| > \gamma, \\ 0 & \text{otherwise.} \end{cases}$$

Here $\gamma > 0$ is a fixed height. We have

$$\int |f_0|^{p_0}\,dx = \int_{\{|f|>\gamma\}} |f|^{p_0}\,dx \le \gamma^{p_0-p} \int_{\{|f|>\gamma\}} |f|^p\,dx,$$
$$\int |f_1|^{p_1}\,dx = \int_{\{|f|\le\gamma\}} |f|^{p_1}\,dx \le \gamma^{p_1-p} \int_{\{|f|\le\gamma\}} |f|^p\,dx.$$

2. Fix $\lambda > 0$. If $f \in L^p$, we make the decomposition $f = f_0 + f_1$ where $\gamma = \lambda$. By subadditivity we have

$$\{|Tf| > \lambda\} \subseteq \{|Tf_0| > \lambda/2\} \cup \{|Tf_1| > \lambda/2\}.$$

Using the assumptions on T, the distribution function satisfies

$$\begin{aligned} m(|Tf| > \lambda) &\le m(|Tf_0| > \lambda/2) + m(|Tf_1| > \lambda/2) \\ &\le \left(\frac{2C_0\|f_0\|_{L^{p_0}}}{\lambda}\right)^{p_0} + \left(\frac{2C_1\|f_1\|_{L^{p_1}}}{\lambda}\right)^{p_1} \\ &\le \left(\frac{2C_0}{\lambda}\right)^{p_0} \int_{\{|f|>\lambda\}} |f|^{p_0}\,dx + \left(\frac{2C_1}{\lambda}\right)^{p_1} \int_{\{|f|\le\lambda\}} |f|^{p_1}\,dx. \end{aligned}$$

Note that the result in Step 1 would immediately imply that T is weak type (p,p).

3. To obtain strong type bounds, we express the L^p norm in terms of the distribution function as in (6):

$$\begin{aligned} \|Tf\|_{L^p}^p &= \int_0^\infty p\lambda^{p-1} m(|Tf| > \lambda)\,d\lambda \\ &\le p(2C_0)^{p_0} \int_0^\infty \int_{\{|f|>\lambda\}} \lambda^{p-1-p_0} |f|^{p_0}\,dx\,d\lambda \\ &\quad + p(2C_1)^{p_1} \int_0^\infty \int_{\{|f|\le\lambda\}} \lambda^{p-1-p_1} |f|^{p_1}\,dx\,d\lambda. \end{aligned}$$

We estimate the first integral by using Fubini's theorem:

$$\begin{aligned} \int_0^\infty \int_{\{|f|>\lambda\}} \lambda^{p-1-p_0} |f|^{p_0}\,dx\,d\lambda &= \int_{\mathbf{R}^n} \int_0^{|f(x)|} \lambda^{p-1-p_0} |f|^{p_0}\,dx\,d\lambda \\ &= \int_{\mathbf{R}^n} \frac{|f(x)|^{p-p_0}}{p-p_0} |f(x)|^{p_0}\,dx = \frac{1}{p-p_0}\|f\|_{L^p}^p. \end{aligned}$$

Similarly, the second integral is equal to

$$\int_{\mathbf{R}^n} \int_{|f(x)|}^\infty \lambda^{p-1-p_1} |f|^{p_1}\,dx\,d\lambda = \frac{1}{p_1-p}\|f\|_{L^p}^p.$$

Both these computations used the assumption $p_0 < p < p_1$. The theorem is proved in the case $p_1 < \infty$, and the case $p_1 = \infty$ is left as an exercise. $\square$

The Marcinkiewicz theorem is a basic result in real analysis with many applications, such as establishing the L^p mapping properties of the Hilbert transform or the Hardy-Littlewood maximal function. We will instead give a simple result concerning the *Hardy operator*, defined for $f \in L^1_{\mathrm{loc}}(\overline{\mathbf{R}_+})$ by

$$Tf(x) = \frac{1}{x} \int_0^x f(t)\,dt, \quad x > 0.$$

Note that the proof of the Marcinkiewicz theorem works equally well with $\mathbf{R}^n$ replaced by $\mathbf{R}_+$, which is the setting in the next result. The Hardy operator is *not* strong type $(1,1)$, so here it is essential that weak type bounds at the end points are sufficient.

THEOREM 5. *(Mapping properties for T) The Hardy operator maps $L^1(\mathbf{R}_+)$ to weak $L^1(\mathbf{R}_+)$, and is a bounded operator $L^p(\mathbf{R}_+) \to L^p(\mathbf{R}_+)$ for $1 < p \leq \infty$.*

PROOF. Since $|Tf(x)| \leq \frac{1}{x}\|f\|_{L^1(\mathbf{R}_+)}$, we have

$$m(|Tf| > \lambda) \leq m(\{x \in \mathbf{R}_+\,;\, \frac{1}{x}\|f\|_{L^1(\mathbf{R}_+)} > \lambda\}) = \frac{\|f\|_{L^1(\mathbf{R}_+)}}{\lambda}.$$

Thus T is weak type $(1,1)$, and clearly T is strong type (∞,∞). The result follows from Marcinkiewicz interpolation. $\square$

REMARK. The Marcinkiewicz theorem is true also when T is a subadditive operator of weak type (p_0, q_0) and (p_1, q_1), where $p_0 \neq p_1$ and $q_0 \neq q_1$. Then the conclusion is that T is strong type (p,q) if p and q are as in the Riesz-Thorin theorem, and if one has the additional restrictions

$$p_0 \leq q_0, \quad p_1 \leq q_1. \tag{7}$$

Although the Riesz-Thorin and Marcinkiewicz theorems look similar, there are notable differences. In Marcinkiewicz, it is enough to have weak type bounds at the endpoints, and the result applies to certain nonlinear operators. In Riesz-Thorin, the restrictions (7) do not appear, and one can prove L^p bounds in the more general case where T depends analytically on θ.

2.2. Abstract interpolation

We will mostly interpolate Banach spaces. However, for technical reasons it will be convenient to forget a great part of the structure in Banach spaces, and to work with quasinormed Abelian groups.

DEFINITION. Let A be an Abelian group that is written additively. If $c \geq 1$ is a constant, a *c-quasinorm* $\|\cdot\|$ is a real valued function on A satisfying

(1) $\|a\| \geq 0$ for all $a \in A$, and $\|a\| = 0$ iff $a = 0$,
(2) $\|-a\| = \|a\|$ for $a \in A$, and
(3) $\|a + b\| \leq c(\|a\| + \|b\|)$ for $a, b \in A$.

We call $(A, \|\cdot\|)$ a *quasinormed group.*

DEFINITION. A *quasinormed vector space* is a vector space which is a quasinormed group and where the quasinorm satisfies $\|\lambda a\| = |\lambda|\|a\|$ for all scalars λ.

EXAMPLE. If $0 < p < \infty$, let $L^p(\mathbf{R}^n)$ be the set of all measurable functions on $\mathbf{R}^n$ for which the expression

$$\|f\|_{L^p} := \left(\int_{\mathbf{R}^n} |f|^p \, dx\right)^{1/p}$$

is finite. We have

$$\|f + g\|_{L^p} \leq \max(1, 2^{\frac{1-p}{p}})(\|f\|_{L^p} + \|g\|_{L^p}).$$

Thus, if we identify functions which agree outside a set of measure zero, $L^p(\mathbf{R}^n)$ will be a quasinormed vector space if $0 < p < \infty$.

The topology on a quasinormed group A is defined as follows: a set $U \subseteq A$ is open if for all $a \in U$, there is $\varepsilon > 0$ such that the set $\{b \in A\,;\, \|b - a\| < \varepsilon\}$ is contained in U. The following result shows that $(A, \|\cdot\|)$ is in fact a metric space.

LEMMA 1. *(Quasinormed groups are metric spaces) Suppose that A is a c-quasinormed group, and let ρ be defined by the identity $(2c)^\rho = 2$. Then there is a 1-quasinorm $\|\cdot\|^*$ on A such that*

$$\|a\|^* \leq \|a\|^\rho \leq 2\|a\|^*, \quad a \in A.$$

Also, $d(a, b) := \|b - a\|^*$ is a metric defining the topology on A.

PROOF. If $a \in A$, define

$$\|a\|^* := \inf\{\sum_{j=1}^{N} \|a_j\|^\rho \,;\, \sum_{j=1}^{N} a_j = a, N \geq 1\}.$$

Given this definition, we leave the proof as an exercise. $\square$

Since every quasinormed vector space is metrizable, it makes sense to talk about Cauchy sequences and completeness.

DEFINITION. A complete quasinormed vector space will be called a *quasi-Banach space*.

The L^p spaces are quasi-Banach if $0 < p \leq \infty$ (and of course Banach if $1 \leq p \leq \infty$). Also, the weak L^p spaces are quasi-Banach, and if $p > 1$ one can in fact find equivalent norms which make the weak L^p spaces Banach spaces.

We now move to basic concepts of abstract interpolation.

DEFINITION. Let A_0 and A_1 be quasinormed groups. We say that (A_0, A_1) is an *interpolation couple* if there is a quasinormed group $\mathscr{A}$ such that A_0 and A_1 are subgroups of $\mathscr{A}$ and $A_0, A_1 \subseteq \mathscr{A}$ with continuous inclusions.

REMARK. More generally, the reference group $\mathscr{A}$ could be any topological group (such as a space of distributions on $\mathbf{R}^n$).

If (A_0, A_1) is an interpolation couple, then the sets $A_0 \cap A_1$ and $A_0 + A_1$ are subgroups of $\mathscr{A}$, where

$$A_0 + A_1 := \{a \in \mathscr{A} \,;\, a = a_0 + a_1 \text{ for some } a_0 \in A_0, a_1 \in A_1\}.$$

We define

$$\begin{aligned} \|a\|_{A_0\cap A_1} &:= \max(\|a\|_{A_0}, \|a\|_{A_1}), \\ \|a\|_{A_0+A_1} &:= \inf_{a=a_0+a_1, a_j \in A_j} (\|a_0\|_{A_0} + \|a_1\|_{A_1}). \end{aligned}$$

LEMMA 2. *If (A_0, A_1) is an interpolation couple, then $A_0 \cap A_1$ and $A_0 + A_1$ are quasinormed groups.*

PROOF. Clearly $\|\cdot\|_{A_0\cap A_1}$ is nonnegative, satisfies $\|a\|_{A_0\cap A_1} = 0$ iff $a = 0$, and $\|-a\|_{A_0\cap A_1} = \|a\|_{A_0\cap A_1}$. The quasi-triangle inequality follows since

$$\begin{aligned}\|a+b\|_{A_0\cap A_1} &\le \max(c_0(\|a\|_{A_0}+\|b\|_{A_0}), c_1(\|a\|_{A_1}+\|b\|_{A_1}))\\ &\le \max(c_0,c_1)\max(\|a\|_{A_0}+\|b\|_{A_0}, \|a\|_{A_1}+\|b\|_{A_1})\\ &\le \max(c_0,c_1)(\|a\|_{A_0\cap A_1}+\|b\|_{A_0\cap A_1}).\end{aligned}$$

For $\|\cdot\|_{A_0+A_1}$, nonnegativity and invariance under $a\mapsto -a$ are clear. If $\|a\|_{A_0+A_1}=0$, then for any $\varepsilon>0$ there are $a_{j,\varepsilon}\in A_j$ with $a=a_{0,\varepsilon}+a_{1,\varepsilon}$ and $\|a_{0,\varepsilon}\|_{A_0}+\|a_{1,\varepsilon}\|_{A_1}<\varepsilon$. Thus $a_{j,\varepsilon}\to 0$ in A_j as $\varepsilon\to 0$, so also $a=a_{0,\varepsilon}+a_{1,\varepsilon}\to 0$ in $\mathscr{A}$ since $A_j\subseteq\mathscr{A}$ continuously. Finally, if $a,b\in A_0+A_1$ and if $a=a_0+a_1$, $b=b_0+b_1$ with $a_j,b_j\in A_j$, we have

$$\begin{aligned}\|a+b\|_{A_0+A_1} &\le \|a_0+b_0\|_{A_0}+\|a_1+b_1\|_{A_1}\\ &\le \max(c_0,c_1)(\|a_0\|+\|a_1\|+\|b_0\|+\|b_1\|).\end{aligned}$$

Taking the infimum over all such decompositions of a and b, we obtain

$$\|a+b\|_{A_0+A_1}\le \max(c_0,c_1)(\|a\|_{A_0+A_1}+\|b\|_{A_0+A_1}).$$

$\square$

Now, one has the continuous inclusions

$$A_0\cap A_1\subseteq A_j\subseteq A_0+A_1,$$

and $A_0\cap A_1$ and A_0+A_1 are the maximal and minimal groups for which this holds. The idea of interpolation theory is to find groups A_θ such that

$$A_0\cap A_1\subseteq A_\theta\subseteq A_0+A_1,$$

and A_θ should somehow be intermediate between A_0 and A_1.

Also, as in the Riesz-Thorin and Marcinkiewicz theorems, we wish to prove that boundedness of operators between endpoint spaces would imply boundedness between intermediate spaces. We restrict our attention to homomorphisms (that is, operators $T:A\to B$ such that $T(a+b)=Ta+Tb$ and $T(-a)=-Ta$).

Definition. Let A and B be quasinormed groups, and let $T:A\to B$ be a homomorphism. The *quasinorm* of T is

$$\|T\|=\|T\|_{A\to B}:=\sup_{a\in A, a\neq 0}\frac{\|Ta\|_B}{\|a\|_A}.$$

We say that T is *bounded* iff $\|T\|<\infty$.

We are now ready to give the definition of an interpolation functor, which incorporates the idea of intermediate groups and boundedness of operators between them.

Definition. An *interpolation functor* is a rule F such that

(1) F assigns to every interpolation couple (A_0, A_1) a quasinormed group $F(A_0, A_1)$ such that
$$A_0 \cap A_1 \subseteq F(A_0, A_1) \subseteq A_0 + A_1$$
with continuous inclusions, and

(2) if (A_0, A_1) and (B_0, B_1) are any two interpolation couples and $T : A_0 + A_1 \to B_0 + B_1$ is a homomorphism which is bounded $A_0 \to B_0$ and $A_1 \to B_1$, then T is bounded $F(A_0, A_1) \to F(B_0, B_1)$.

Definition. An interpolation functor F is *of type* θ if
$$\|T\|_{F(A_0,A_1)\to F(B_0,B_1)} \le C_\theta \|T\|^{1-\theta}_{A_0\to B_0} \|T\|^{\theta}_{A_1\to B_1}.$$
The functor is said to be *exact* if one can choose $C_\theta = 1$.

If $j = 0, 1$, then $F_j(A_0, A_1) := A_j$ are exact interpolation functors of type j. The main purpose below will be to introduce nontrivial exact interpolation functors obtained from real interpolation, and to study their properties.

2.3. Real interpolation

The real interpolation method is due to Lions and Peetre (late 1950s and early 1960s). There are several equivalent ways of introducing the corresponding interpolation functors. We will use here the K-*functional* for this purpose.

Definition. Let (A_0, A_1) be an interpolation couple. We define, for $t > 0$ and $a \in A_0 + A_1$,
$$K(t, a) = K(t, a; A_0, A_1) := \inf_{a=a_0+a_1, a_j\in A_j} (\|a_0\|_{A_0} + t\|a_1\|_{A_1}).$$

The idea is that for each fixed $t > 0$, $K(t, \cdot)$ is a quasinorm on $A_0 + A_1$ which is equivalent (that is, gives the same topology) to the usual one. We will define interpolation groups as subgroups of $A_0 + A_1$ by imposing restrictions on the behaviour of $K(t, a)$ as $t \to 0$ and $t \to \infty$. The values t correspond to the cut levels γ in the proof of the Marcinkiewicz theorem.

Definition. Let (A_0, A_1) be an interpolation couple. If $0 < \theta < 1$ and $0 < q < \infty$, and if $a \in A_0 + A_1$, we define

$$\|a\|_{(A_0,A_1)_{\theta,q}} := \left(\int_0^\infty [t^{-\theta} K(t,a)]^q \frac{dt}{t} \right)^{1/q} \tag{1}$$

We denote by $(A_0, A_1)_{\theta,q}$ the set of all $a \in A_0 + A_1$ for which the quantity (1) is finite.

Remark. Here θ is the main interpolation parameter, and q can be used to fine-tune the interpolation. One could also consider the case $q = \infty$, but we will not need to do so.

We shall prove that the rule $(A_0, A_1) \mapsto (A_0, A_1)_{\theta,q}$ is an exact interpolation functor of type θ. The first step is a simple lemma.

Lemma 1. *(Properties of the K-functional) Suppose (A_0, A_1) is an interpolation couple. If $a \in A_0 + A_1$ is fixed, then $K(\,\cdot\,, a)$ is a positive, nondecreasing, concave, and continuous function of $t > 0$. If A_j is c_j-quasinormed, we have*

$$K(t, a+b) \le c_0 \left[K(c_1 t/c_0, a) + K(c_1 t/c_0, b)\right], \quad a, b \in A_0 + A_1. \tag{2}$$

Proof. It is clear that $K(\,\cdot\,, a)$ is positive and nondecreasing. We leave as an exercise to check that it is concave and continuous. If $a, b \in A_0 + A_1$ and if $a = a_0 + a_1$, $b = b_0 + b_1$ with $a_j, b_j \in A_j$, then

$$\begin{aligned} K(t, a+b) &\le \|a_0 + b_0\|_{A_0} + t\|a_1 + b_1\|_{A_1} \\ &\le c_0(\|a_0\|_{A_0} + \|b_0\|_{A_0}) + c_1 t(\|a_1\|_{A_1} + \|b_1\|_{A_1}) \\ &\le c_0 \left(\|a_0\|_{A_0} + \frac{c_1 t}{c_0}\|a_1\|_{A_1} + \|b_0\|_{A_0} + \frac{c_1 t}{c_0}\|b_1\|_{A_1} \right). \end{aligned}$$

The result follows by taking the infimum over all decompositions of a and b. □

Next we show that the first requirement for an interpolation functor is satisfied.

Theorem 2. *(Interpolation groups) Suppose (A_0, A_1) is an interpolation couple, $0 < \theta < 1$, and $0 < q < \infty$. Then $(A_0, A_1)_{\theta,q}$ is a quasinormed group, and*

$$A_0 \cap A_1 \subseteq (A_0, A_1)_{\theta,q} \subseteq A_0 + A_1$$

with bounded inclusions.

PROOF. 1. The first step is to prove that $\|\cdot\| := \|\cdot\|_{(A_0,A_1)_{\theta,q}}$ satisfies the conditions for a quasinorm. Clearly $\|a\| \geq 0$ and $\|-a\| = \|a\|$, and if $\|a\| = 0$ then $K(t,a) = 0$ for a.e. $t > 0$, showing that $a = 0$ as in Lemma 2. For the quasi-triangle inequality we use (1) and the quasi-triangle inequality on $L^q := L^q(\mathbf{R}_+, \frac{dt}{t})$:

$$\begin{aligned}\|a+b\| &= \|t^{-\theta}K(t,a+b)\|_{L^q}\\ &\leq c_0 \max(1, 2^{\frac{1-q}{q}})(\|t^{-\theta}K(c_1t/c_0,a)\|_{L^q} + \|t^{-\theta}K(c_1t/c_0,b)\|_{L^q})\\ &\leq c_0^{1-\theta}c_1^{\theta}\max(1,2^{\frac{1-q}{q}})(\|a\|+\|b\|).\end{aligned}$$

In the last step we changed the integration variable t to c_0t/c_1.

2. It immediately follows that $(A_0,A_1)_{\theta,q}$ is a quasinormed group: if $a, b \in (A_0,A_1)_{\theta,q}$ then $\|a\|, \|b\| < \infty$, so by Step 1 also $\|a+b\| < \infty$ and $a + b \in (A_0,A_1)_{\theta,q}$.

3. To show that $A_0 \cap A_1 \subseteq (A_0,A_1)_{\theta,q}$ with bounded inclusion, note that for $a \in A_0 \cap A_1$ we have $K(t,a) \leq \|a\|_{A_0}$ and $K(t,a) \leq t\|a\|_{A_1}$, so

$$K(t,a) \leq \min(1,t)\|a\|_{A_0\cap A_1}.$$

Thus, if $a \in A_0 \cap A_1$, we have

$$\|a\| \leq \|t^{-\theta}\min(1,t)\|_{L^q(\mathbf{R}_+,dt/t)}\|a\|_{A_0\cap A_1} \leq C_{\theta,q}\|a\|_{A_0\cap A_1}.$$

4. Finally, we prove that $(A_0,A_1)_{\theta,q} \subseteq A_0 + A_1$ with bounded inclusion. This is the case $t = 1$ in the more general identity

$$K(t,a) \leq C_{\theta,q}t^{\theta}\|a\|_{(A_0,A_1)_{\theta,q}}, \quad a \in (A_0,A_1)_{\theta,q},\ t > 0.$$

To show the last identity, since $K(\,\cdot\,,a)$ is nondecreasing we have

$$t^{-\theta}K(t,a) = C_{\theta,q}\left(\int_t^{\infty} s^{-\theta q}\frac{ds}{s}\right)^{1/q} K(t,a) \leq C_{\theta,q}\|a\|_{(A_0,A_1)_{\theta,q}}.$$

The result is proved. $\square$

It remains to show the second requirement for an interpolation functor, involving boundedness of operators. The definition of $(A_0,A_1)_{\theta,q}$ is set up so that this will be easy.

THEOREM 3. *(Interpolation of operators) Let* (A_0, A_1) *and* (B_0, B_1) *be interpolation couples, and let* $T : A_0 + A_1 \to B_0 + B_1$ *be a homomorphism which is bounded* $A_j \to B_j$ *with quasinorm* M_j $(j = 0, 1)$*. If* $0 < \theta < 1$ *and* $0 < q < \infty$*, then* T *is bounded* $(A_0, A_1)_{\theta,q} \to (B_0, B_1)_{\theta,q}$ *and*

$$\|T\|_{(A_0,A_1)_{\theta,q}\to(B_0,B_1)_{\theta,q}} \le M_0^{1-\theta} M_1^{\theta}.$$

PROOF. First assume $M_0 > 0$, and let $a \in (A_0, A_1)_{\theta,q}$. Then we have $Ta \in B_0 + B_1$, and

$$\begin{aligned} K(t, Ta) &\le \inf_{a=a_0+a_1, a_j\in A_j} (\|Ta_0\|_{B_0} + t\|Ta_1\|_{B_1}) \\ &\le M_0 \inf_{a=a_0+a_1, a_j\in A_j} (\|a_0\|_{A_0} + \frac{M_1 t}{M_0}\|a_1\|_{A_1}) \\ &= M_0 K(M_1 t/M_0, a). \end{aligned}$$

Thus

$$\begin{aligned} \|Ta\|_{(B_0,B_1)_{\theta,q}} &= \|t^{-\theta} K(t, Ta)\|_{L^q(\mathbf{R}_+, dt/t)} \\ &\le M_0 \|t^{-\theta} K(M_1 t/M_0, a)\|_{L^q(\mathbf{R}_+, dt/t)} \\ &\le M_0^{1-\theta} M_1^{\theta} \|a\|_{(A_0,A_1)_{\theta,q}} \end{aligned}$$

by changing the variable t to $M_0 t/M_1$.

If $M_0 = 0$, we can repeat the above argument with M_0 replaced by $\varepsilon > 0$. The result is obtained by letting $\varepsilon \to 0$. □

There is a density result which will be useful when interpolating L^p and Sobolev spaces. It is usually proved via the J*-method*, which is an equivalent way of introducing the spaces $(A_0, A_1)_{\theta,q}$ in real interpolation. Due to lack of time, we will not give the proof here.

THEOREM 4. *(Density of* $A_0 \cap A_1$*) If* (A_0, A_1) *is an interpolation couple and* $0 < \theta < 1$, $0 < q < \infty$*, then* $A_0 \cap A_1$ *is dense in* $(A_0, A_1)_{\theta,q}$.

PROOF. See Bergh-Löfström [**1**, Section 3.11]. □

We end the section by stating some standard properties of real interpolation, and by giving a simple example.

THEOREM 5. *(Elementary properties of real interpolation) Suppose* (A_0, A_1) *is an interpolation couple and* $0 < \theta < 1$, $0 < q < \infty$*. Then*

(1) $(A_0, A_1)_{\theta,q} = (A_1, A_0)_{1-\theta,q}$,
(2) $(A_0, A_1)_{\theta,q} \subseteq (A_0, A_1)_{\theta,\tilde{q}}$ *if* $q \le \tilde{q}$,

(3) *if* $A_0 = A_1$ *then* $(A_0, A_1)_{\theta,q} = A_0 = A_1$,
(4) $\|a\|_{(A_0,A_1)_{\theta,q}} \le C_{\theta,q}\|a\|_{A_0}^{1-\theta}\|a\|_{A_1}^{\theta}$ *for* $a \in A_0 \cap A_1$.
(5) *if* A_0 *and* A_1 *are complete, then so is* $(A_0, A_1)_{\theta,q}$.

Proof. Exercise. □

Example. Let $A_0 = \mathbf{R} \times \{0\}$ and $A_1 = \mathbf{R}^2$, with the usual additive structure and with norms $\|(u, 0)\|_{A_0} = |u|$, $\|(u, v)\|_{A_1} = |u| + |v|$. To determine $(A_0, A_1)_{\theta,q}$, note that if $a = (u, v) \in A_1$ is written as $a = a_0 + a_1$ with $a_j \in A_j$, then $a_0 = (s, 0)$ and $a_1 = (u - s, v)$. Thus

$$K(t, a) = \inf_{s\in\mathbf{R}}(|s| + t(|u - s| + |v|)) = \min(1, t)|u| + t|v|$$

since $s \mapsto |s| + t|u - s|$ is a continuous piecewise linear function going to infinity as $s \to \pm\infty$, which reaches its minimum at $s = 0$ or $s = u$. Now, if $v \neq 0$ then

$$\|a\|_{(A_0,A_1)_{\theta,q}} \ge \left(\int_1^\infty [t^{-\theta}t|v|]^q \frac{dt}{t}\right)^{1/q} = \infty.$$

Thus $(A_0, A_1)_{\theta,q} = A_0 = \mathbf{R} \times \{0\}$ for all θ and q.

2.4. Interpolation of L^p spaces

2.4.1. Interpolation of powers. When interpolating L^p spaces, instead of the L^p norm it will be convenient to consider the quantity $\|f\|_{L^p}^p$ which does not involve the pth root. We will next formalize this idea.

Let $(A, \|\cdot\|)$ be a quasinormed group. If $p > 0$, then by the inequality

$$(x + y)^p \le \max(1, 2^{p-1})(x^p + y^p), \quad x, y \ge 0,$$

the quantity $\|\cdot\|^p$ is also a quasinorm defining the topology on A. We denote by $(A)^p$ the group A equipped with quasinorm $\|\cdot\|^p$. The interpolation result for powers of groups is as follows.

Theorem 1. *(Interpolation of powers) Let* $0 < p_0, p_1 < \infty$ *and let* $0 < \eta < 1$. *If* (A_0, A_1) *is an interpolation couple, then*

$$((A_0)^{p_0}, (A_1)^{p_1})_{\eta,1} = ((A_0, A_1)_{\theta,p})^p$$

with equivalent norms, where

$$p := (1 - \eta)p_0 + \eta p_1, \quad \theta := \frac{p_1}{p}\eta. \tag{1}$$

PROOF. 1. We will use the functional

$$K_\infty(t,a) = K_\infty(t,a;A_0,A_1) := \inf_{a=a_0+a_1, a_j\in A_j} \max(\|a_0\|_{A_0}, t\|a_1\|_{A_1}).$$

This satisfies

$$K_\infty(t,a) \le K(t,a) \le 2K_\infty(t,a),$$

and thus $\|a\|_{(A_0,A_1)_{\theta,q}} \sim \|t^{-\theta}K_\infty(t,a)\|_{L^q(\mathbf{R}_+,dt/t)}$.

2. If $a \in A_0 + A_1 = (A_0)^{p_0} + (A_1)^{p_1}$, then

$$\|a\|_{((A_0)^{p_0},(A_1)^{p_1})_{\eta,1}} \sim \int_0^\infty s^{-\eta} K_\infty(s,a;(A_0)^{p_0},(A_1)^{p_1}) \frac{ds}{s}. \tag{2}$$

We claim that

$$K_\infty(s,a;(A_0)^{p_0},(A_1)^{p_1}) = (K_\infty(t,a;A_0,A_1))^{p_0} \tag{3}$$

if s is defined by

$$s := t^{p_1}(K_\infty(t,a;A_0,A_1))^{p_0-p_1}. \tag{4}$$

To prove (3), we temporarily denote the left and right hand sides by $K_\infty(s)$ and $K_\infty(t)$. Then, using (4),

$$\begin{aligned} K_\infty(s) &= \inf_{a=a_0+a_1} \max(\|a_0\|_{A_0}^{p_0}, s\|a_1\|_{A_1}^{p_1}) \\ &= K_\infty(t)^{p_0} \inf_{a=a_0+a_1} \max\left[\left(\frac{\|a_0\|_{A_0}}{K_\infty(t)}\right)^{p_0}, \left(\frac{t\|a_1\|_{A_1}}{K_\infty(t)}\right)^{p_1}\right]. \end{aligned}$$

Given $\varepsilon > 0$, we have $K_\infty(t) \le \max(\|a_{0,\varepsilon}\|_{A_0}, t\|a_{1,\varepsilon}\|_{A_1}) \le (1+\varepsilon)K_\infty(t)$ for some $a_{j,\varepsilon} \in A_j$. Then

$$1 \le \max\left[\frac{\|a_{0,\varepsilon}\|_{A_0}}{K_\infty(t)}, \frac{t\|a_{1,\varepsilon}\|_{A_1}}{K_\infty(t)}\right] \le 1+\varepsilon.$$

The claim (3) follows.

3. We wish to go back to (2) and make the change of variables (4). To this end, we claim that

$$\begin{cases} s = s(t) \text{ in (4) is a strictly increasing bijective map } \mathbf{R}_+ \to \mathbf{R}_+ \\ \text{which is Lipschitz continuous on bounded intervals.} \end{cases}$$

We start by proving the Lipschitz continuity. If $0 < t_0 \le t_1$, then $K_\infty(t_1) \le \frac{t_1}{t_0} K_\infty(t_0)$ and

$$0 \le K_\infty(t_1) - K_\infty(t_0) \le \frac{t_1 - t_0}{t_0} K_\infty(t_0). \tag{5}$$

Thus, if $I := [1/R, R] \subseteq \mathbf{R}_+$ with R large, then

$$|K_\infty(t_1) - K_\infty(t_0)| \le R \left(\sup_{t \in I} K_\infty(t) \right) |t_1 - t_0|, \quad t_0, t_1 \in I.$$

This shows that $t \mapsto K_\infty(t)$ is Lipschitz continuous on bounded intervals in $\mathbf{R}_+$, and then so is $s = s(t)$.

Whenever the derivative exists, we have

$$\frac{s'(t)}{s(t)} = \left(p_1 + (p_0 - p_1) \frac{t}{K_\infty(t)} K'_\infty(t) \right) \frac{1}{t}.$$

By (5), again at points where the derivative exists,

$$0 \le K'_\infty(t) \le \frac{K_\infty(t)}{t}.$$

Thus, we have

$$\frac{\min(p_0, p_1)}{t} \le \frac{s'(t)}{s(t)} \le \frac{\max(p_0, p_1)}{t}. \tag{6}$$

Consequently $s(t)$ is strictly increasing, and it maps $\mathbf{R}_+$ onto itself.

4. By Step 3, we may change variables in (2) according to (4) (by restricting the integral to the interval $[1/R, R]$ and letting $R \to \infty$). Using (3) and the estimate (6), we obtain

$$\|a\|_{((A_0)^{p_0}, (A_1)^{p_1})_{\eta,1}} \sim \int_0^\infty t^{-p_1 \eta} K_\infty(t)^p \frac{dt}{t}.$$

The right hand side is $\|a\|^p_{(A_0, A_1)_{\theta,p}}$ as required. □

Remark. (Exercise) Let $p_0, p_1 > 0$. If $0 < \eta < 1$ is given, then (1) implies that

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad 1 - \theta = (1 - \eta) \frac{p_0}{p}. \tag{7}$$

Conversely, if $0 < \theta < 1$ is given, then (7) defines p and η so that (1) is satisfied.

2.4.2. Interpolation of sequence spaces. We are now ready to interpolate the sequence spaces l^p. More generally, we will consider the weighted Banach-valued sequence spaces $l^s_p(A)$ where $0 < p < \infty$ and s is a real number. The reason for writing p as a lower index is that later on, s will correspond to the smoothness index in Sobolev type spaces and p will denote the base L^p space. This notation will be consistent with the notation for Sobolev type spaces.

DEFINITION. Let A be a Banach space. We denote by $l_p^s(A)$ the set of all sequences $a = (a_k)_{k=0}^{\infty}$ where $a_k \in A$, for which the quasinorm

$$\|a\|_{l_p^s(A)} := \left(\sum_{k=0}^{\infty} (2^{ks} \|a_k\|_A)^p \right)^{1/p}$$

is finite.

It is easy to see that $l_p^s(A)$ is a quasi-Banach space (Banach if $p \geq 1$), and that $l_p^0(A)$ is just the usual l^p space of A-valued sequences.

THEOREM 2. *(Interpolation of $l_p^s(A)$ spaces) Let $0 < p_0, p_1 < \infty$ and let s_0, s_1 be real numbers. Also let A_0 and A_1 be Banach spaces such that (A_0, A_1) is an interpolation couple. If $0 < \theta < 1$, then*

$$(l_{p_0}^{s_0}(A_0), l_{p_1}^{s_1}(A_1))_{\theta,p} = l_p^s((A_0, A_1)_{\theta,p}) \tag{8}$$

where p and s are defined by

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad s = (1-\theta)s_0 + \theta s_1.$$

PROOF. 1. To make the formulas shorter, we omit writing A_0 and A_1. By Theorem 1 and the remark after it, we have

$$((l_{p_0}^{s_0})^{p_0}, (l_{p_1}^{s_1})^{p_1})_{\eta,1} = ((l_{p_0}^{s_0}, l_{p_1}^{s_1})_{\theta,p})^p$$

where $0 < \eta < 1$ is such that $p = (1-\eta)p_0 + \eta p_1$. Thus, it is enough to study the space $((l_{p_0}^{s_0})^{p_0}, (l_{p_1}^{s_1})^{p_1})_{\eta,1}$.

2. First assume that $a = (a_k)_{k=0}^{\infty}$ is a finite sequence with values in $A_0 \cap A_1$, such that $a_k = 0$ for $k > N$. Then

$$\|a\|_{((l_{p_0}^{s_0})^{p_0}, (l_{p_1}^{s_1})^{p_1})_{\eta,1}} = \int_0^{\infty} t^{-\eta} K(t, a; (l_{p_0}^{s_0})^{p_0}, (l_{p_1}^{s_1})^{p_1}) \frac{dt}{t}$$
$$= \int_0^{\infty} t^{-\eta} \inf_{a=b+c} \left[\|b\|_{l_{p_0}^{s_0}}^{p_0} + t\|c\|_{l_{p_1}^{s_1}}^{p_1} \right] \frac{dt}{t}$$
$$= \int_0^{\infty} t^{-\eta} \inf_{a=b+c} \sum_{k=0}^{N} \left[(2^{ks_0}\|b_k\|)^{p_0} + t(2^{ks_1}\|c_k\|)^{p_1} \right] \frac{dt}{t}$$

where $b \in l^{s_0}_{p_0}$ and $c \in l^{p_1}_{s_1}$. We may interchange the sum and infimum and use Fubini's theorem, and the last expression is equal to

$$
\begin{aligned}
&= \sum_{k=0}^{N} \int_0^\infty t^{-\eta} \inf_{a_k=b_k+c_k} \left[(2^{ks_0}\|b_k\|)^{p_0} + t(2^{ks_1}\|c_k\|)^{p_1} \right] \frac{dt}{t} \\
&= \sum_{k=0}^{N} \int_0^\infty t^{-\eta} 2^{ks_0p_0} K(t2^{k(s_1p_1-s_0p_0)}, a_k; (A_0)^{p_0}, (A_1)^{p_1}) \frac{dt}{t} \\
&= \sum_{k=0}^{N} 2^{k[(1-\eta)s_0p_0+\eta s_1p_1]} \|a_k\|_{((A_0)^{p_0},(A_1)^{p_1})_{\eta,1}},
\end{aligned}
$$

where the last equality follows by changing t to $2^{-k(s_1p_1-s_0p_0)}t$ in the integral.

By Theorem 1 and the remark after it, we have

$$((A_0)^{p_0}, (A_1)^{p_1})_{\eta,1} = ((A_0, A_1)_{\theta,p})^p,$$

and by (7) it also follows that

$$(1-\eta)s_0p_0 + \eta s_1p_1 = (1-\theta)s_0p + \theta s_1 p = sp.$$

Thus, using the result in Step 1, we have proved that

$$\|a\|_{(l^{s_0}_{p_0}(A_0), l^{s_1}_{p_1}(A_1))_{\theta,p}} = \|a\|_{l^s_p((A_0,A_1)_{\theta,p})} \tag{9}$$

for all finite sequences a with values in $A_0 \cap A_1$.

3. We can now prove (8). We first claim that

$$\left\{ \begin{array}{l} \text{the set of finite } A_0 \cap A_1\text{-valued sequences is dense} \\ \text{in both spaces in (8).} \end{array} \right.$$

By Theorem 4 in §2.3, $A_0 \cap A_1$ is dense in $(A_0, A_1)_{\theta,p}$ and thus the claim is true for $l^s_p((A_0,A_1)_{\theta,p})$. In a similar fashion, $l^{s_0}_{p_0}(A_0) \cap l^{s_1}_{p_1}(A_1) = l^{s_0}_{p_0}(A_0 \cap A_1) \cap l^{s_1}_{p_1}(A_0 \cap A_1)$ is dense in $(l^{s_0}_{p_0}, l^{s_1}_{p_1})_{\theta,p}$ so the claim holds also for the latter space.

Let now $a \in (l^{s_0}_{p_0}, l^{s_1}_{p_1})_{\theta,p}$. This means that $a \in l^{s_0}_{p_0} + l^{s_1}_{p_1}$ and that $\|a\|_{(l^{s_0}_{p_0}, l^{s_1}_{p_1})_{\theta,p}}$ is finite. Now, there is a sequence $(a^{(N)})$ of finite $A_0 \cap A_1$-valued sequences such that $a^{(N)} \to a$ in $(l^{s_0}_{p_0}, l^{s_1}_{p_1})_{\theta,p}$. It follows from (9) that $(a^{(N)})$ is a Cauchy sequence in $l^s_p((A_0,A_1)_{\theta,p})$, and by completeness there exists $\tilde{a} \in l^s_p((A_0,A_1)_{\theta,p})$ with $a^{(N)} \to \tilde{a}$ in $l^s_p((A_0,A_1)_{\theta,p})$. It remains to show that $a = \tilde{a}$, but this follows from uniqueness of limits since

$$(l^{s_0}_{p_0}, l^{s_1}_{p_1})_{\theta,p}, l^s_p((A_0, A_1)_{\theta,p}) \subseteq l^{\min(s_0,s_1)}_{\max(p_0,p_1)}(A_0 + A_1)$$

continuously. We have proved that $(l^{s_0}_{p_0}, l^{s_1}_{p_1})_{\theta,p} \subseteq l^s_p((A_0, A_1)_{\theta,p})$, and the other inclusion follows similarly. □

2.4.3. Interpolation of L^p spaces. Finally, we prove the interpolation result for Banach valued L^p spaces. First we need to recall some concepts related to vector-valued integration.

Definition. Let (X, Σ, μ) be a measure space where μ is a positive σ-finite measure, and let $(A, \|\cdot\|)$ be a Banach space.

(1) A *simple function* is any function $s : X \to A$ having the form $s(x) = \sum_{j=1}^N a_j \chi_{B_j}(x)$, where $a_j \in A$ and where B_j are disjoint subsets of X with finite measure.
(2) A function $f : X \to A$ is *strongly measurable* if there exists a sequence (s_k) of simple functions such that $s_k(x) \to f(x)$ for μ-a.e. x as $k \to \infty$.
(3) If $0 < p < \infty$, the space $L^p(A) = L^p(X, \Sigma, \mu; A)$ consists of those strongly measurable functions $f : X \to A$ for which the quasinorm

$$\|f\|_{L^p(A)} := \left(\int_X \|f(x)\|^p \, d\mu(x) \right)^{1/p}$$

is finite. We identify elements of $L^p(X; A)$ which agree μ-a.e.

The last definition is valid since $\|f\|$ is measurable if f is strongly measurable. The spaces $L^p(A)$ are quasi-Banach (Banach if $p \geq 1$), and simple functions are dense.

Theorem 3. *(Interpolation of L^p spaces) Let (X, Σ, μ) be a σ-finite measure space, let $0 < p_0, p_1 < \infty$, and let A_0 and A_1 be Banach spaces so that (A_0, A_1) is an interpolation couple. If $0 < \theta < 1$, then*

$$(L^{p_0}(A_0), L^{p_1}(A_1))_{\theta,p} = L^p((A_0, A_1)_{\theta,p})$$

where p is defined by

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}.$$

PROOF. 1. The proof is analogous to the proof of Theorem 2, and we will give a sketch. Again, we omit writing A_0 and A_1. Theorem 1 and the remark after it imply that

$$((L^{p_0})^{p_0}, (L^{p_1})^{p_1})_{\eta,1} = ((L^{p_0}, L^{p_1})_{\theta,p})^p$$

where $0 < \eta < 1$ is such that $p = (1-\eta)p_0 + \eta p_1$.

2. Let f be a simple function with values in $A_0 \cap A_1$. By Step 1, we have

$$\begin{aligned} \|f\|^p_{(L^{p_0},L^{p_1})_{\theta,p}} &\sim \int_0^\infty t^{-\eta} \inf_{f=f_0+f_1, f_j \in L^{p_j}} (\|f_0\|^{p_0}_{L^{p_0}} + t\|f_1\|^{p_1}_{L^{p_1}}) \frac{dt}{t} \\ &\sim \int_0^\infty t^{-\eta} \inf_{f=f_0+f_1, f_j \in L^{p_j}} \int_X \left[\|f_0(x)\|^{p_0}_{A_0} + t\|f_1(x)\|^{p_1}_{A_1}\right] \, d\mu(x) \, \frac{dt}{t}. \end{aligned}$$

We wish to interchange the infimum and the integral over X. In fact, we claim that

$$\begin{aligned} &\inf_{\substack{f=f_0+f_1 \\ f_j \in L^{p_j}}} \int_X \left[\|f_0(x)\|^{p_0}_{A_0} + t\|f_1(x)\|^{p_1}_{A_1}\right] \, d\mu(x) \\ &\qquad\qquad = \int_X \inf_{\substack{f(x)=f_0(x)+f_1(x) \\ f_j(x) \in A_j}} \left[\|f_0(x)\|^{p_0}_{A_0} + t\|f_1(x)\|^{p_1}_{A_1}\right] \, d\mu(x). \end{aligned}$$

The measurability of the integrand on the right is part of the claim. The proof is left as an exercise. Now, Fubini's theorem implies that

$$\begin{aligned} &\|f\|^p_{(L^{p_0},L^{p_1})_{\theta,p}} \\ &\sim \int_X \int_0^\infty t^{-\eta} \inf_{\substack{f(x)=f_0(x)+f_1(x) \\ f_j(x) \in A_j}} \left[\|f_0(x)\|^{p_0}_{A_0} + t\|f_1(x)\|^{p_1}_{A_1}\right] \, d\mu(x) \, \frac{dt}{t} \\ &\sim \int_X \int_0^\infty t^{-\eta} K(t, f(x); (A_0)^{p_0}, (A_1)^{p_1}) \, \frac{dt}{t} \, d\mu(x) \\ &\sim \int_X \|f(x)\|_{((A_0)^{p_0},(A_1)^{p_1})_{\eta,1}} \, d\mu(x) \sim \int_X \|f(x)\|^p_{(A_0,A_1)_{\theta,p}} \, d\mu(x). \end{aligned}$$

In the last step we used Theorem 1.

3. We have seen that the quasinorms on $(L^{p_0}, L^{p_1})_{\theta,p}$ and $L^p((A_0, A_1)_{\theta,p})$ are equivalent on simple functions with values in $A_0 \cap A_1$. One now argues that such functions are dense in both of the spaces involved, and the result follows as in the proof of Theorem 2. $\square$

CHAPTER 3

Fractional Sobolev spaces, Besov and Triebel spaces

The purpose in this chapter is to extend the theory of integer order Sobolev spaces $W^{k,p}$ to the case of fractional Sobolev spaces $H^{s,p}$, Besov spaces B^s_{pq}, and Triebel-Lizorkin spaces F^s_{pq}. Here, k is a nonnegative integer, s is any real number, and $1 < p, q < \infty$ (or in some cases $0 < p, q \leq \infty$). With these conventions, the various spaces satisfy the relations

$$\begin{aligned} W^{k,p} &= H^{k,p}, \\ H^{s,2} &= B^s_{22} = F^s_{22}, \\ H^{s,p} &= F^s_{p2}. \end{aligned}$$

Because of the first identity, we will write $W^{s,p} := H^{s,p}$.

There are many situations where it is useful to go beyond the $W^{k,p}$ spaces. We will describe a few such situations, in the setting of the Dirichlet problem

$$\begin{cases} -\Delta u = f & \text{in } U, \\ \quad u = g & \text{on } \partial U \end{cases}$$

where U is a bounded open set in $\mathbf{R}^n$ with C^1 boundary.

(1) If $f = 0$ and $g \in L^p(\partial U)$, then there is a unique solution

$$u \in \begin{cases} W^{1/p,p}(U) & \text{if } p \geq 2, \\ B^{1/p}_{p2}(U) & \text{if } 1 < p < 2. \end{cases}$$

In general $u \notin W^{1,p}(U)$.

(2) If one looks for solutions $u \in W^{1,p}(U)$, then the natural space for g is $B^{1-1/p}_{pp}(\partial U) = F^{1-1/p}_{pp}(\partial U)$. It can be shown that the latter space is the image of $W^{1,p}(U)$ under the trace operator.

(3) If $f \in L^p(U)$ and $g = 0$, and if ∂U is C^2, it can be shown that there is a unique solution $u \in W^{2,p}(U)$. However, if ∂U is only C^1, then $u \notin W^{2,p}(U)$ in general (for $p = 2$ the sharp result is that $u \in W^{3/2,2}(U)$).

The above results may be found in Jerison-Kenig [**5**].

Besides PDE, fractional Sobolev type spaces are useful in harmonic analysis and approximation theory, where it is often possible to prove sharper and more general statements by choosing the right spaces. Also, the scales of spaces B^s_{pq} and F^s_{pq} include many other spaces which are common in analysis, such as Hölder, Zygmund, Hardy, and BMO spaces.

Below, we will exclusively consider function spaces in $\mathbf{R}^n$. The theory for function spaces in bounded domains or lower dimensional manifolds can be developed based on the corresponding theory for $\mathbf{R}^n$.

3.1. Fourier analysis

It would be possible to define fractional Sobolev spaces via real interpolation, for instance by $W^{\theta,2}(\mathbf{R}^n) := (L^2(\mathbf{R}^n), W^{1,2}(\mathbf{R}^n))_{\theta,2}$ for $0 < \theta < 1$. We will instead begin with an intrinsic definition based on the Fourier transform, and we will later show that the interpolation definition coincides with the intrinsic one.

Motivation. Let $f \in L^2((0, 2\pi))$ be a 2π-periodic function (that is, an L^2 function on the torus $T^1 := \mathbf{R}/2\pi\mathbf{Z}$). One has the Fourier series

$$f(x) = \sum_{k=-\infty}^{\infty} \hat{f}(k) e^{ikx}, \tag{1}$$

with $\hat{f}(k) = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx}\, dx$. Also, by Parseval's theorem

$$\|\hat{f}\|_{l^2} = (2\pi)^{-1/2} \|f\|_{L^2((0,2\pi))}.$$

If $D = \frac{1}{i}\frac{d}{dx}$, differentiating (1) formally gives

$$D^m f(x) = \sum_{k=-\infty}^{\infty} k^m \hat{f}(k) e^{ikx}.$$

This suggests the following notion of a fractional derivative: if $s \geq 0$, then for any $f \in L^2((0, 2\pi))$ we define

$$|D|^s f(x) := \sum_{k=-\infty}^{\infty} |k|^s \hat{f}(k) e^{ikx},$$

provided that the last series converges in some sense. If f is such that $(|k|^s \hat{f}(k))$ is square summable, then $|D|^s f$ is an L^2 function by Parseval's theorem. This motivates the definition

$$W^{s,2}(T^1) := \{f \in L^2((0, 2\pi))\,;\, (|k|^s \hat{f}(k)) \in l^2\}, \quad s \geq 0,$$

which is the right definition for Sobolev spaces on the torus.

A similar idea works for Sobolev spaces in $\mathbf{R}^n$, but one needs to replace Fourier series by the Fourier transform. We have already defined the Fourier transform in §2.1 for functions in $L^1(\mathbf{R}^n)$, but here a more general definition is required.

In the following, we will give a brief review of the Fourier transform in the general setting of tempered distributions. For proofs we refer to Rudin [**9**, Chapter 7]. We begin by considering a test function space designed for the purposes of Fourier analysis.

Definition. The *Schwartz space* $\mathscr{S}(\mathbf{R}^n)$ is the set of all infinitely differentiable functions $f : \mathbf{R}^n \to \mathbf{C}$ such that the seminorms

$$\|f\|_{\alpha,\beta} := \|x^\alpha D^\beta f(x)\|_{L^\infty(\mathbf{R}^n)}$$

are finite for all multi-indices $\alpha, \beta \in \mathbf{N}^n$. If $(f_j)_{j=1}^\infty$ is a sequence in $\mathscr{S}$, we say that $f_j \to f$ in $\mathscr{S}$ if $\|f_j - f\|_{\alpha,\beta} \to 0$ for all α, β.

It follows from the definition that a smooth function f is in $\mathscr{S}(\mathbf{R}^n)$ iff for all β and N there exists $C > 0$ such that

$$|D^\beta f(x)| \leq C\langle x\rangle^{-N}, \quad x \in \mathbf{R}^n.$$

Based on this, Schwartz space is sometimes called the space of *rapidly decreasing test functions.*

Example. Any function in $C_c^\infty(\mathbf{R}^n)$ is in Schwartz space, and functions like $e^{-\gamma|x|^2}$, $\gamma > 0$, also belong to $\mathscr{S}$. The function $e^{-\gamma|x|}$ is not in Schwartz space because it is not smooth at the origin, and also $\langle x\rangle^{-N}$ is not in $\mathscr{S}$ because it doesn't decrease sufficiently rapidly at infinity.

There is a metric defining the topology on $\mathscr{S}$, and then $\mathscr{S}$ is a complete metric space. The operations $f \mapsto Pf$ and $f \mapsto D^\beta f$ are continuous maps $\mathscr{S} \to \mathscr{S}$, if P is any polynomial and β any multi-index. More generally, let

$$\begin{aligned} \mathscr{O}_M(\mathbf{R}^n) := \{f \in C^\infty(\mathbf{R}^n) \ ; \text{ for all } \beta \text{ there exist } C, N > 0 \\ \text{such that } |D^\beta f(x)| \leq C\langle x\rangle^N\}. \end{aligned}$$

It is easy to see that the map $f \mapsto af$ is continuous $\mathscr{S} \to \mathscr{S}$ if $a \in \mathscr{O}_M$.

Definition. If $f \in \mathscr{S}(\mathbf{R}^n)$, then the *Fourier transform* of f is the function $\mathscr{F} f = \hat{f} : \mathbf{R}^n \to \mathbf{C}$ defined by

$$\hat{f}(\xi) := \int_{\mathbf{R}^n} e^{-ix\cdot\xi} f(x)\, dx, \quad \xi \in \mathbf{R}^n.$$

The importance of Schwartz space is based on the fact that it is invariant under the Fourier transform.

Theorem 1. *(Fourier inversion formula) The Fourier transform is an isomorphism from $\mathscr{S}(\mathbf{R}^n)$ onto $\mathscr{S}(\mathbf{R}^n)$. A Schwartz function f may be recovered from its Fourier transform by the inversion formula*

$$f(x) = \mathscr{F}^{-1}\hat{f}(x) = (2\pi)^{-n} \int_{\mathbf{R}^n} e^{ix\cdot\xi} \hat{f}(\xi)\, d\xi, \quad x \in \mathbf{R}^n.$$

After introducing the Fourier transform on nicely behaving functions, it is possible to define it in a very general setting by duality.

Definition. Let $\mathscr{S}'(\mathbf{R}^n)$ be the set of continuous linear functionals $\mathscr{S}(\mathbf{R}^n) \to \mathbf{C}$. The elements of $\mathscr{S}'$ are called *tempered distributions*, and their action on test functions is written as

$$\langle T, \varphi\rangle := T(\varphi), \qquad T \in \mathscr{S}',\ \varphi \in \mathscr{S}.$$

If $(T_j)_{j=1}^\infty$ is a sequence in $\mathscr{S}'$ and if $T \in \mathscr{S}'$, we say that $T_j \to T$ in $\mathscr{S}'$ if $\langle T_j, \varphi\rangle \to \langle T, \varphi\rangle$ for all $\varphi \in \mathscr{S}(\mathbf{R}^n)$.

More precisely, $\mathscr{S}'$ is the dual space of $\mathscr{S}$, and it is equipped with the weak-* topology. The elements in $\mathscr{S}'$ are somewhat loosely also called *distributions of polynomial growth*, which is justified by the following examples.

Examples. 1. Let $u : \mathbf{R}^n \to \mathbf{C}$ be a measurable function, such that for some $C, N > 0$ one has

$$|u(x)| \leq C\langle x\rangle^N, \quad \text{for a.e. } x.$$

Then $T_u \in \mathscr{S}'(\mathbf{R}^n)$, where

$$\langle T_u, \varphi \rangle := \int_{\mathbf{R}^n} u(x)\varphi(x)\,dx, \quad \varphi \in \mathscr{S}(\mathbf{R}^n).$$

We will always identify a function u with the tempered distribution T_u defined in this way.

2. In a similar way, any function $u \in L^p(\mathbf{R}^n)$ with $1 \leq p \leq \infty$ is a tempered distribution (with the identification $u = T_u$).

3. Let μ be a positive Borel measure in $\mathbf{R}^n$ such that

$$\int_{\mathbf{R}^n} \langle x \rangle^{-N}\,d\mu(x) < \infty$$

for some $N > 0$. Then T_μ is a tempered distribution, where

$$\langle T_\mu, \varphi \rangle := \int_{\mathbf{R}^n} \varphi(x)\,d\mu(x), \quad \varphi \in \mathscr{S}(\mathbf{R}^n).$$

In particular, the Dirac measure δ_{x_0} at $x_0 \in \mathbf{R}^n$ is in $\mathscr{S}'$, and

$$\langle \delta_{x_0}, \varphi \rangle = \varphi(x_0), \quad \varphi \in \mathscr{S}(\mathbf{R}^n).$$

4. The function $e^{\gamma x}$ is not in $\mathscr{S}'(\mathbf{R})$ if $\gamma \neq 0$, since it is not possible to define $\int_{\mathbf{R}} e^{\gamma x}\varphi(x)\,dx$ for all Schwartz functions φ. However, $e^{\gamma x}\cos(e^{\gamma x})$ belongs to $\mathscr{S}'$ since it is the distributional derivative (see below) of the bounded function $\sin(e^{\gamma x}) \in \mathscr{S}'$.

We now wish to extend some operations, defined earlier for Schwartz functions, to the case of tempered distributions. This is possible via duality. For instance, let $a \in \mathscr{O}_M(\mathbf{R}^n)$. If $u, \varphi \in \mathscr{S}(\mathbf{R}^n)$ we have

$$\langle au, \varphi \rangle = \langle u, a\varphi \rangle.$$

Since $\varphi \mapsto a\varphi$ is continuous on $\mathscr{S}$, we may for any $T \in \mathscr{S}'$ define the product aT as the tempered distribution given by

$$\langle aT, \varphi \rangle := \langle T, a\varphi \rangle, \quad \varphi \in \mathscr{S}.$$

Similarly, motivated by the corresponding property for functions in $\mathscr{S}$, if $T \in \mathscr{S}'$ then the (distributional) derivative $D^\beta T$ is the tempered distribution given by

$$\langle D^\beta T, \varphi \rangle := (-1)^{|\beta|}\langle T, D^\beta \varphi \rangle, \quad \varphi \in \mathscr{S}.$$

Finally, we can define the Fourier transform of any $T \in \mathscr{S}'$ as the tempered distribution $\mathscr{F}T = \hat{T}$ with

$$\langle \hat{T}, \varphi \rangle := \langle T, \hat{\varphi} \rangle, \quad \varphi \in \mathscr{S}.$$

We state the basic properties of these operations as a theorem.

THEOREM 2. *(Operations on $\mathscr{S}'$) If $a \in \mathscr{O}_M$ and β is any multi-index, then the multiplication $T \mapsto aT$ and the distributional derivative $T \mapsto D^\beta T$ are continuous maps $\mathscr{S}' \to \mathscr{S}'$. Also, the Fourier transform $T \mapsto \hat{T}$ is an isomorphism from $\mathscr{S}'$ onto $\mathscr{S}'$, and one has the inversion formula*

$$\langle T, \varphi \rangle = (2\pi)^{-n} \langle \hat{T}, \hat{\varphi}(-\,\cdot\,) \rangle, \quad \varphi \in \mathscr{S}.$$

If $T \in \mathscr{S}$ then these operations coincide with the usual ones on $\mathscr{S}$.

The preceding result shows that one may differentiate or Fourier transform any tempered distribution (thus, also any L^p function or measurable polynomially bounded function) any amount of times, and the result will be a tempered distribution. We will give some examples of this very general (and useful) phenomenon.

EXAMPLES. 1. Let $u(x) := |x|,\ x \in \mathbf{R}$. Since u is continuous and polynomially bounded, we have $u \in \mathscr{S}'(\mathbf{R})$. We claim the one has the distributional derivatives

$$\begin{aligned} u' &= \operatorname{sgn}(x), \\ u'' &= 2\delta_0. \end{aligned}$$

In fact, if $\varphi \in \mathscr{S}(\mathbf{R})$, one has

$$\begin{aligned} \langle u', \varphi \rangle = -\langle u, \varphi' \rangle &= \int_{-\infty}^0 x\varphi'(x)\,dx - \int_0^\infty x\varphi'(x)\,dx \\ &= \int_{\mathbf{R}} \operatorname{sgn}(x)\varphi(x)\,dx = \langle \operatorname{sgn}(x), \varphi \rangle, \end{aligned}$$

using integration by parts and the rapid decay of φ. Similarly,

$$\begin{aligned} \langle u'', \varphi \rangle = -\langle u', \varphi' \rangle &= \int_{-\infty}^0 \varphi'(x)\,dx - \int_0^\infty \varphi'(x)\,dx \\ &= 2\varphi(0) = \langle 2\delta_0, \varphi \rangle. \end{aligned}$$

2. The Fourier transform of δ_0 is the constant 1, since

$$\begin{aligned}\langle \widehat{\delta_0}, \varphi \rangle = \langle \delta_0, \hat{\varphi} \rangle = \hat{\varphi}(0) &= \int_{-\infty}^{\infty} \varphi(x)\,dx \\ &= \langle 1, \varphi \rangle.\end{aligned}$$

If $u \in L^2(\mathbf{R}^n)$ then u is a tempered distribution, and the Fourier transform $\hat{u}$ is another element of $\mathscr{S}'$. The Plancherel theorem (which is the exact analog of Parseval's theorem for Fourier series) states that in fact $\hat{u} \in L^2$, and that the Fourier transform is an isometry on L^2 up to a constant.

Theorem 3. *(Plancherel theorem) The Fourier transform is an isomorphism from $L^2(\mathbf{R}^n)$ onto $L^2(\mathbf{R}^n)$, and one has*

$$\|\hat{u}\|_{L^2} = (2\pi)^{n/2} \|u\|_{L^2}.$$

We end this section by noting the identities

$$\begin{aligned}(D^\alpha u)\hat{} &= \xi^\alpha \hat{u}, \\ (x^\alpha u)\hat{} &= (-D_\xi)^\alpha \hat{u},\end{aligned}$$

which hold for Schwartz functions u by a direct computation, and remain true for tempered distributions u by duality. This shows that the Fourier transform converts derivatives into multiplication by polynomials, and vice versa. This will be our route for defining fractional Sobolev spaces: it is easy to define fractional derivatives of u by multiplying $\hat{u}$ with fractional powers of polynomials on the Fourier side.

3.2. Fractional Sobolev spaces

We are now in a position to give the Fourier transform definition of fractional Sobolev spaces.

Motivation. We can characterize $W^{k,2}(\mathbf{R}^n)$ by using the Fourier transform and the Plancherel theorem by

$$\begin{aligned}u \in W^{k,2}(\mathbf{R}^n) \ &\Leftrightarrow \ D^\alpha u \in L^2(\mathbf{R}^n) \text{ for all } |\alpha| \leq k \\ &\Leftrightarrow \ \xi^\alpha \hat{u} \in L^2(\mathbf{R}^n) \text{ for all } |\alpha| \leq k.\end{aligned}$$

Thus, a function is in $W^{k,2}(\mathbf{R}^n)$ iff its Fourier transform, multiplied by any polynomial of degree $\leq k$, is in $L^2(\mathbf{R}^n)$. In fact, one polynomial is sufficient: we have

$$u \in W^{k,2}(\mathbf{R}^n) \ \Leftrightarrow \ \langle \xi \rangle^k \hat{u} \in L^2(\mathbf{R}^n).$$

To see this, it is enough to note that $|\xi^\alpha| \leq |\xi|^{|\alpha|} \leq \langle\xi\rangle^k$ for $|\alpha| \leq k$, and that

$$\langle\xi\rangle^{2k} = (1 + \xi_1^2 + \ldots + \xi_n^2)^k = \sum_{|\gamma| \leq 2k} c_\gamma \xi^\gamma.$$

This shows that $\langle\xi\rangle^k \hat{u} \in L^2$ iff $\xi^\alpha \hat{u} \in L^2$ for $|\alpha| \leq k$, as desired.

Here and later, it will be convenient to introduce a notation for multipliers on the Fourier side.

Definition. Let $m(\xi) \in \mathscr{O}_M(\mathbf{R}^n)$. We define the *Fourier multiplier operator*

$$m(D)u := \mathscr{F}^{-1}\{m(\xi)\hat{u}(\xi)\}$$

for any tempered distribution $u \in \mathscr{S}'(\mathbf{R}^n)$.

The maps $u \mapsto \hat{u} \mapsto m\hat{u} \mapsto \mathscr{F}^{-1}\{m\hat{u}\}$ are continuous on $\mathscr{S}'$, so the operator $m(D)$ is a well defined, continuous linear operator on $\mathscr{S}'$. Since $m(\xi) = \langle\xi\rangle^s$ is in $\mathscr{O}_M$ for any $s \in \mathbf{R}$, also the next definition is a valid one.

Definition. Let $1 < p < \infty$ and let s be a real number. The *fractional Sobolev space* $H^{s,p}(\mathbf{R}^n)$ is the set of all tempered distributions $u \in \mathscr{S}'(\mathbf{R}^n)$ such that

$$\langle D\rangle^s u \in L^p(\mathbf{R}^n).$$

The norm is given by $\|u\|_{H^{s,p}(\mathbf{R}^n)} := \|\langle D\rangle^s u\|_{L^p(\mathbf{R}^n)}$.

More explicitly, the space $H^{s,p}(\mathbf{R}^n)$ consists of those tempered distributions u in $\mathbf{R}^n$, for which the tempered distribution $\mathscr{F}^{-1}\{\langle\xi\rangle^s\hat{u}(\xi)\}$ happens to be an L^p function. The operator $\langle D\rangle^s$ is a *Bessel potential*, and it is an isometric isomorphism from $H^{s,p}(\mathbf{R}^n)$ onto $L^p(\mathbf{R}^n)$. The spaces $H^{s,p}$ are also called *Bessel potential spaces* for this reason.

Theorem 1. *(Sobolev spaces as function spaces) If s is a real number and $1 < p < \infty$, then $H^{s,p}(\mathbf{R}^n)$ is a Banach space and $\mathscr{S}(\mathbf{R}^n)$ is a dense subset.*

Proof. If $u, v \in H^{s,p}$ and λ, μ are scalars, then $\langle D\rangle^s(\lambda u + \mu v) = \lambda\langle D\rangle^s u + \mu\langle D\rangle^s v$ is an L^p function, showing that $\lambda u + \mu v \in H^{s,p}$. Thus $H^{s,p}$ is a vector space.

To show that $\|\cdot\|_{H^{s,p}}$ is a norm, all the other properties are clear except that fact that $\|u\|_{H^{s,p}} = 0$ implies $u = 0$. But if $\|u\|_{H^{s,p}} = 0$ then $\langle D\rangle^s u = 0$ as an L^p function, thus also as a tempered distribution.

Taking Fourier transforms, we obtain $\langle \xi \rangle^s \hat{u} = 0$, which implies $\hat{u} = 0$ upon multiplying by the function $\langle \xi \rangle^{-s} \in \mathscr{O}_M$. Thus $u = 0$.

For completeness, let (u_j) be a Cauchy sequence in $H^{s,p}$. It follows that $(\langle D \rangle^s u_j)$ is a Cauchy sequence in L^p, and there exists $\tilde{u} \in L^p$ such that

$$\langle D \rangle^s u_j \to \tilde{u} \quad \text{in } L^p.$$

Let $u := \langle D \rangle^{-s} \tilde{u}$. Then $u \in H^{s,p}$, and $u_j \to u$ in $H^{s,p}$ as required.

The Schwartz space is a subset of $H^{s,p}$, since $\mathscr{F}^{-1}\{\langle \xi \rangle^s \hat{f}\}$ belongs to $\mathscr{S}$ for $f \in \mathscr{S}$ (recall that the Fourier transform and multiplication by $\mathscr{O}_M$ functions are continuous maps $\mathscr{S} \to \mathscr{S}$). To show density, let $u \in H^{s,p}$. Then $u_0 := \langle D \rangle^s u$ is an L^p function, and since $1 < p < \infty$ there exists a sequence $(v_j) \subseteq C_c^\infty(\mathbf{R}^n)$ with $v_j \to u_0$ in L^p. Let

$$u_j := \langle D \rangle^{-s} v_j.$$

Since $v_j \in \mathscr{S}$, we have $u_j \in \mathscr{S}$. Also,

$$\|u_j - u\|_{H^{s,p}} = \|\langle D \rangle^s u_j - \langle D \rangle^s u\|_{L^p} = \|v_j - u_0\|_{L^p} \to 0$$

as $j \to \infty$, showing that $\mathscr{S}$ is dense in $H^{s,p}$. □

We next wish to show that $H^{k,p}(\mathbf{R}^n) = W^{k,p}(\mathbf{R}^n)$ if k is a nonnegative integer. The case $p = 2$ follows from the argument in the beginning of this section. In general, if $u \in H^{k,p}$ and $|\alpha| \leq k$, we have

$$D^\alpha u = D^\alpha \langle D \rangle^{-k} (\langle D \rangle^k u)$$

where $\langle D \rangle^k u \in L^p$ by definition. If one could show that $D^\alpha \langle D \rangle^{-k}$ maps L^p to L^p if $|\alpha| \leq k$, then it would follow that $D^\alpha u \in L^p$ as required.

The following Fourier multiplier result will be used several times below. For a proof, see [**2**].

THEOREM 2. *(Mihlin multiplier theorem) Assume that $m = m(\xi)$ is a bounded C^∞ complex valued function in $\mathbf{R}^n$, which satisfies*

$$|D^\alpha m(\xi)| \leq M \langle \xi \rangle^{-|\alpha|}, \quad \xi \in \mathbf{R}^n,$$

for $|\alpha| \leq \left\lfloor \frac{n}{2} \right\rfloor + 1$. Then $m(D)$ is a bounded linear operator $L^p(\mathbf{R}^n) \to L^p(\mathbf{R}^n)$ for $1 < p < \infty$, and

$$\|m(D)u\|_{L^p} \leq C_p M \|u\|_{L^p}.$$

In the next proof, notice how Fourier multipliers make it possible to manipulate derivatives in an efficient way.

THEOREM 3. *(The space $H^{s,p}$ for s integer) If $1 < p < \infty$ and if k is a nonnegative integer, then $H^{k,p}(\mathbf{R}^n) = W^{k,p}(\mathbf{R}^n)$.*

PROOF. 1. If $u \in H^{k,p}(\mathbf{R}^n)$, then $D^\alpha u = m(D)(\langle D\rangle^k u)$ where

$$m(\xi) := \frac{\xi^\alpha}{\langle \xi\rangle^k}.$$

If $|\alpha| \leq k$ then m satisfies the conditions of the Mihlin multiplier theorem. Consequently $D^\alpha u \in L^p$ for $|\alpha| \leq k$, showing that $u \in W^{k,p}$.

2. Let $u \in W^{k,p}(\mathbf{R}^n)$. Then $D^\alpha u \in L^p$ for $|\alpha| \leq k$, and we need to prove that $\langle D\rangle^k u \in L^p$. Let $\chi(\xi) \in C^\infty(\mathbf{R})$ with $0 \leq \chi \leq 1$, $\chi = 0$ for $|\xi| \leq 1$, and $\chi = 1$ for $|\xi| \geq 2$. We write

$$\langle D\rangle^k u = \mathscr{F}^{-1}\{\langle \xi\rangle^k \hat{u}\} = \mathscr{F}^{-1}\{\tilde{m}(\xi)(1 + \sum_{j=1}^n \chi(\xi_j)|\xi_j|^k)\hat{u}\}$$

where

$$\tilde{m}(\xi) := \frac{\langle \xi\rangle^k}{1 + \sum_{j=1}^n \chi(\xi_j)|\xi_j|^k}.$$

This shows that

$$\begin{aligned}\langle D\rangle^k u &= \tilde{m}(D)u + \sum_{j=1}^n \mathscr{F}^{-1}\{\tilde{m}(\xi)\chi(\xi_j)|\xi_j|^k \hat{u}\} \\ &= \tilde{m}(D)u + \sum_{j=1}^n \tilde{m}(D)m_j(D_j)(D_j^k u)\end{aligned}$$

where

$$m_j(\xi_j) := \chi(\xi_j)|\xi_j|^k \xi_j^{-k} = \begin{cases} \chi(\xi_j), & \xi_j \geq 0, \\ (-1)^k \chi(\xi_j), & \xi_j < 0. \end{cases}$$

3. We claim that $\tilde{m}(\xi)$ satisfies the conditions of the Mihlin multiplier theorem. Note that $\tilde{m}$ is smooth because $\chi = 0$ near 0. Also,

$$\langle \xi\rangle^k \leq (1 + n \max_j |\xi_j|^2)^{k/2} \leq C(1 + \max_j |\xi_j|^k) \leq C(1 + \sum_{j=1}^n |\xi_j|^k),$$

which shows that $\tilde{m}$ is bounded. The derivatives satisfy $|D^\alpha \tilde{m}(\xi)| \leq C\langle \xi\rangle^{-|\alpha|}$, so the Mihlin condition holds. We obtain

$$\|\langle D\rangle^k u\|_{L^p} \leq C\left(\|u\|_{L^p} + \sum_{j=1}^n \|m_j(D_j)(D_j^k u)\|_{L^p}\right).$$

4. Clearly $m_j(\xi_j)$ satisfies the conditions in Mihlin's theorem for $n = 1$. We have

$$\begin{aligned}\|m_j(D_j)(D_j^k u)\|_{L^p(\mathbf{R}^n)} &= \| \, \|m_j(D_j)(D_j^k u)\|_{L^p(\mathbf{R})} \, \|_{L^p(\mathbf{R}^{n-1})} \\ &\leq C\| \, \|D_j^k u\|_{L^p(\mathbf{R})} \, \|_{L^p(\mathbf{R}^{n-1})} = C\|D_j^k u\|_{L^p(\mathbf{R}^n)}.\end{aligned}$$

Since $D_j^k u \in L^p$ for all j, we obtain $\langle D \rangle^k u \in L^p$ as required. $\square$

DEFINITION. We will sometimes write $W^{s,p}(\mathbf{R}^n) := H^{s,p}(\mathbf{R}^n)$.

Note that $u \in W^{k,p}$ iff $u \in W^{k-1,p}$ and $D_j u \in W^{k-1,p}$ for all j. Mihlin's theorem allows to prove a similar result for $H^{s,p}$.

THEOREM 4. *(Monotonicity of $H^{s,p}$ spaces) Let $1 < p < \infty$ and let s be real. Then*

$$H^{s',p}(\mathbf{R}^n) \subseteq H^{s,p}(\mathbf{R}^n) \quad \text{for } s' \geq s,$$

and one has

$$u \in H^{s,p}(\mathbf{R}^n) \;\Leftrightarrow\; u \in H^{s-1,p}(\mathbf{R}^n),\, D_j u \in H^{s-1,p}(\mathbf{R}^n) \;\; (1 \leq j \leq n).$$

PROOF. Exercise. $\square$

The preceding result implies in particular that $H^{s,p}$ is a subspace of L^p if $s \geq 0$, and that $W^{k,p} \subseteq H^{s,p}$ if $k \geq s$. If s is negative then $H^{s,p}$ will contain distributions which are not functions.

EXAMPLES. 1. We have

$$\delta_0 \in H^{s,2} \quad \text{iff} \quad s < -n/2.$$

Indeed, since $\widehat{\delta_0} = 1$, one has $\delta_0 \in H^{s,2}$ iff $\mathscr{F}^{-1}\{\langle \xi \rangle^s\} \in L^2$ iff $\langle \xi \rangle^s \in L^2$ by Plancherel. The last condition holds iff $s < -n/2$.

2. If $1 < p < \infty$, a more careful analysis of $\mathscr{F}^{-1}\{\langle \xi \rangle^s\}$ shows that $\delta_0 \in H^{s,p}$ iff $s < -n/p'$.

3. Consider the function $u : \mathbf{R} \to \mathbf{R}$ with

$$u(x) := \begin{cases} 1 - |x| & \text{for } |x| < 1, \\ 0 & \text{for } |x| \geq 1. \end{cases}$$

We claim that $u \in H^{s,2}(\mathbf{R})$ iff $s < 3/2$. This can be proved by taking the Fourier transform, or alternatively by considering distributional derivatives:

$$u' = \begin{cases} 1 & \text{for} \quad -1 < x < 0, \\ -1 & \text{for} \quad 0 < x < 1, \\ 0 & \text{for} \quad |x| \geq 1, \end{cases}$$

$$u'' = \delta_{-1} - 2\delta_0 + \delta_1.$$

By part 1 in this example, $u'' \in H^{s,2}(\mathbf{R})$ for $s < -1/2$. Then by Theorem 4, u' is in $H^{s,2}(\mathbf{R})$ for $s < 1/2$, and consequently $u \in H^{s,2}(\mathbf{R})$ if $s < 3/2$. The 'only if' part follows by showing that $u'' \notin H^{-1/2,2}(\mathbf{R})$ (exercise).

We will continue to develop the theory of fractional Sobolev spaces (embedding theorems, interpolation, equivalent characterizations) in the more general setting of Triebel spaces later. Here, we will conclude the section by determining the dual space of $H^{s,p}$ (in the dual pairing of $\mathscr{S}$ and $\mathscr{S}'$).

Theorem 5. *(Duality of $H^{s,p}$ spaces) Let $1 < p < \infty$ and let s be a real number. Then*

$$(H^{s,p}(\mathbf{R}^n))' = H^{-s,p'}(\mathbf{R}^n).$$

More precisely, given a continuous linear functional $T : H^{s,p} \to \mathbf{C}$, there is a unique $u \in H^{-s,p'}$ such that $T(\varphi) = \langle u, \varphi\rangle$ for $\varphi \in \mathscr{S}$. Conversely, if $u \in H^{-s,p'}$, then u (acting on $\mathscr{S}$) has a unique extension as a continuous linear functional on $H^{s,p}$.

Proof. Let $T : H^{s,p} \to \mathbf{C}$ be a continuous linear functional. Then $|T(v)| \leq C\|\langle D\rangle^s v\|_{L^p}$ for all $v \in H^{s,p}$, and we may define a functional

$$\tilde{T} : L^p \to \mathbf{C}, \quad \tilde{T}(w) := T(\langle D\rangle^{-s} w).$$

This satisfies $|\tilde{T}(w)| \leq C\|w\|_{L^p}$, and the duality for L^p spaces implies that there is a unique function $u_0 \in L^{p'}$ such that

$$\tilde{T}(w) = \langle u_0, w\rangle, \quad w \in L^p.$$

We let $u := \langle D\rangle^s u_0$. Then $u \in H^{-s,p'}$ and for $\varphi \in \mathscr{S}$ (expressed as $\varphi = \hat{\psi}$ for $\psi \in \mathscr{S}$) one has

$$\begin{aligned} \langle u, \varphi\rangle &= \langle \langle D\rangle^s u_0, \hat{\psi}\rangle = \langle \langle \xi\rangle^s \hat{u}_0, \psi\rangle = \langle \hat{u}_0, \langle \xi\rangle^s \psi\rangle = \langle u_0, (\langle \xi\rangle^s \psi)\hat{}\,\rangle \\ &= \langle u_0, \langle D\rangle^s \varphi\rangle \end{aligned}$$

since $\mathscr{F}\{\langle D\rangle^s\varphi\} = \langle\xi\rangle^s\mathscr{F}^2\{\psi\} = \mathscr{F}^2\{\langle\xi\rangle^s\psi\}$. This shows that

$$\langle u, \varphi\rangle = \tilde{T}(\langle D\rangle^s\varphi) = T(\varphi)$$

as required.

Conversely, if $u \in H^{-s,p'}$, then $u = \langle D\rangle^s u_0$ with $u_0 \in L^{p'}$. Similarly as above, we have for $\varphi \in \mathscr{S}$

$$\langle u, \varphi\rangle = \langle\langle D\rangle^s u_0, \varphi\rangle = \langle u_0, \langle D\rangle^s\varphi\rangle$$

and consequently

$$|\langle u, \varphi\rangle| \leq \|u_0\|_{L^{p'}}\|\langle D\rangle^s\varphi\|_{L^p} = \|u\|_{H^{-s,p'}}\|\varphi\|_{H^{s,p}}.$$

Thus u has an extension (by continuity) which is a continuous linear functional on $H^{s,p}$. $\square$

3.3. Littlewood-Paley theory

For the discussion of Besov and Triebel spaces, it will be useful to introduce some more properties for the Fourier transform in L^p. We have seen many times that the L^2 theory is simpler than the L^p theory, the reason being that in the L^2 case one can exploit orthogonality and the Plancherel theorem. We will next introduce a partial substitute for these properties in the L^p case.

Motivation. We will, as usual, use Fourier series in 1D to gain intuition on the problem. If f is an L^2 function in $(0, 2\pi)$, one has the Parseval theorem

$$\int_0^{2\pi} |f(x)|^2\,dx = 2\pi\sum_{k=-\infty}^{\infty}|\hat{f}(k)|^2.$$

We would like to express the L^p norm of f in terms of its Fourier coefficients. For simplicity, we consider the case $p = 4$ and the function

$$f(x) = \sum_{k=0}^{N} a_k e^{ikx}.$$

To study the L^4 norm of f, we compute

$$\begin{aligned}\int_0^{2\pi} |f(x)|^4\,dx &= \int_0^{2\pi} \left(\sum_{k=0}^{N} a_k e^{ikx}\right)^2 \left(\sum_{l=0}^{N} \bar{a}_l e^{-ilx}\right)^2 dx \\ &= \sum_{0\le k_1,k_2,l_1,l_2\le N} a_{k_1}a_{k_2}\bar{a}_{l_1}\bar{a}_{l_2} \int_0^{2\pi} e^{i(k_1+k_2-l_1-l_2)x}\,dx \\ &= 2\pi \sum_{\substack{0\le k_1,k_2,l_1,l_2\le N\\ k_1+k_2=l_1+l_2}} a_{k_1}a_{k_2}\bar{a}_{l_1}\bar{a}_{l_2}.\end{aligned}$$

For given k_1 and k_2, there are many l_1 and l_2 satisfying $k_1+k_2 = l_1+l_2$, and it is hard to see any orthogonality in the above expression. However, in the special case where the Fourier series is *sparse* (or *lacunary*) in the sense that

$$a_k = 0 \text{ unless } k = 2^p \text{ for some integer } p,$$

it follows that

$$\begin{aligned}\int_0^{2\pi} |f(x)|^4\,dx &= 2\pi \sum_{\substack{k_j=2^{p_j},\,l_j=2^{q_j}\\ 2^{p_1}+2^{p_2}=2^{q_1}+2^{q_2}}} a_{k_1}a_{k_2}\bar{a}_{l_1}\bar{a}_{l_2} \\ &= 2\pi \sum_{k=2^p} |a_k|^4 + 4\pi \sum_{k_j=2^{p_j},\,k_1\ne k_2} |a_{k_1}|^2|a_{k_2}|^2.\end{aligned}$$

The last equality follows since $2^{p_1}+2^{p_2} = 2^{q_1}+2^{q_2}$ implies $\{p_1,p_2\} = \{q_1,q_2\}$. Finally, we obtain

$$2\pi\left(\sum_{k=0}^{N}|a_k|^2\right)^2 \le \int_0^{2\pi}|f(x)|^4\,dx \le 4\pi\left(\sum_{k=0}^{N}|a_k|^2\right)^2,$$

which can be written as

$$\|f\|_{L^4} \sim \left(\sum_{k=-\infty}^{\infty} |\hat{f}(k)|^2\right)^{1/2}.$$

The above example indicates that the functions $e^{i2^j x}$ and $e^{i2^{j+1}x}$ are somehow orthogonal in an L^p sense (this has to do with the fact that the first function is roughly constant on the period of the second one). A similar property remains true for e^{ikx} and e^{ilx} if $k \in \Delta_j$ and $l \in \Delta_{j+1}$, where

$$\Delta_j = \begin{cases} \{0\}, & j = 0, \\ \{2^{j-1}, 2^{j-1}+1, \ldots, 2^j - 1\}, & j > 0, \\ -\Delta_{-j}, & j < 0. \end{cases}$$

We introduce the *dyadic* (or *Littlewood-Paley*) decomposition of f, given in terms of

$$S_j f := \sum_{k \in \Delta_j} \hat{f}(k) e^{ikx}.$$

We denote by Sf the vector

$$Sf := (S_j f)_{j \in \mathbf{Z}}.$$

The *Littlewood-Paley square function* is the map

$$x \mapsto \|Sf(x)\|_{l^2} = \left(\sum_{j=-\infty}^{\infty} |S_j f(x)|^2 \right)^{1/2}.$$

The heart of Littlewood-Paley theory is the fact that the L^p norm of a function is comparable to the L^p norm of the corresponding square function.

THEOREM 1. *(Littlewood-Paley theorem for Fourier series) If f is an L^p function in $(0, 2\pi)$, and if $1 < p < \infty$, then $Sf \in L^p(l^2)$ and*

$$\|f\|_{L^p} \sim \|Sf\|_{L^p(l^2)}.$$

PROOF. See [**3**]. □

The theorem is saying that the Littlewood-Paley pieces $S_j f$ of an L^p function f are somewhat independent of each other. For instance, one can multiply each such piece by ± 1 and the resulting function will still be in L^p. When proving L^p results in harmonic analysis it is often enough to prove the result for a single Littlewood-Paley piece, and the general case follows by summing over all such pieces.

The purpose in the remainder of this section is to discuss the counterpart of Theorem 1 for the Fourier transform in $\mathbf{R}^n$. This will involve a smooth Littlewood-Paley decomposition of a function f in $L^p(\mathbf{R}^n)$.

DEFINITION. Let $\eta = \eta(\xi)$ be a fixed radial function in $C_c^\infty(\mathbf{R}^n)$ such that $0 \leq \eta \leq 1$, $\eta = 1$ for $|\xi| \leq 1/4$, and $\eta = 0$ for $|\xi| \geq 1/2$. We define nonnegative functions φ, φ_j by $\varphi(\xi)^2 = \eta(\xi/2) - \eta(\xi)$, and

$$\varphi_j(\xi)^2 = \varphi(2^{-j}\xi)^2 \qquad (j \geq 1),$$
$$\varphi_0(\xi)^2 = 1 - \sum_{j=1}^{\infty} \varphi_j(\xi)^2.$$

We collect some properties of the functions φ_j.

THEOREM 2. *(Littlewood-Paley partition of unity) The functions φ_j are smooth in $\mathbf{R}^n$, they satisfy $0 \leq \varphi_j \leq 1$, and they are supported in frequency annuli of radius $\sim 2^j$:*

$$\mathrm{supp}(\varphi_j) \subseteq \{2^{j-2} \leq |\xi| \leq 2^j\}, \quad \mathrm{supp}(\varphi_0) \subseteq \{|\xi| \leq 1\}.$$

Also, they form a partition of unity in the sense that

$$1 = \sum_{j=0}^{\infty} \varphi_j^2.$$

PROOF. Since η is radial and nonnegative, it is not hard to see that $\varphi(\xi) := (\eta(\xi/2) - \eta(\xi))^{1/2}$ is C^∞ and that $0 \leq \varphi \leq 1$. Also, φ is supported in $\{1/4 \leq |\xi| \leq 1\}$, which shows the support condition for φ_j if $j \geq 1$. We have

$$\sum_{j=1}^{N} \varphi_j(\xi)^2 = \sum_{j=1}^{N} \left[\eta(2^{-j-1}\xi) - \eta(2^{-j}\xi)\right] = \eta(2^{-N-1}\xi) - \eta(2^{-1}\xi).$$

The last function is supported in $\{1/2 \leq |\xi| \leq 2^N\}$ and is equal to 1 in $\{1 \leq |\xi| \leq 2^{N-1}\}$. This shows that φ_0 is supported in the unit ball, as required. $\square$

DEFINITION. If $f \in \mathscr{S}'(\mathbf{R}^n)$ and $j \geq 0$, we define

$$S_j f := \varphi_j(D) f,$$
$$S f := (S_j f)_{j=0}^{\infty}.$$

The functions $S_j f$ are called the dyadic (or Littlewood-Paley) pieces of f. They are always C^∞ functions in $\mathbf{R}^n$, as Fourier transforms of compactly supported distributions (see Rudin [**9**, Chapter 7]). The square function in this case is the map

$$x \mapsto \|Sf(x)\|_{l^2} = \left(\sum_{j=0}^{\infty} |S_j f(x)|^2 \right)^{1/2}.$$

The Littlewood-Paley theorem states that the $L^p(\mathbf{R}^n)$ norm of f is comparable to the $L^p(\mathbf{R}^n)$ norm of the square function.

THEOREM 3. *(Littlewood-Paley theorem in $\mathbf{R}^n$) Let $1 < p < \infty$. If $f \in L^p(\mathbf{R}^n)$, then $Sf \in L^p(l^2)$ and*

$$\|f\|_{L^p} \sim \|Sf\|_{L^p(l^2)}.$$

If $p = 2$ then $\|f\|_{L^2} = \|Sf\|_{L^2(l^2)}$.

PROOF. We prove the case $p = 2$, which is a direct consequence of the Plancherel theorem: if $f \in L^2(\mathbf{R}^n)$ then

$$\begin{aligned}
\|f\|_{L^2}^2 = (2\pi)^{-n}\|\hat{f}\|_{L^2}^2 &= (2\pi)^{-n}\sum_{j=0}^{\infty}\int \psi_j(\xi)^2|\hat{f}(\xi)|^2\,d\xi \\
&= \sum_{j=0}^{\infty}\int |S_jf(x)|^2\,dx = \int \|Sf(x)\|_{l^2}^2\,dx \\
&= \|Sf\|_{L^2(l^2)}^2.
\end{aligned}$$

The case $p \neq 2$ is a nontrivial result. A standard proof, as in [**2**], is based on the theory of (Hilbert-valued) singular integrals, where it is shown that operators such as S map $L^1(\mathbf{R}^n)$ into weak $L^1(l^2)$. Then a version of the Marcinkiewicz interpolation theorem gives that

$$\|Sf\|_{L^p(l^2)} \leq C\|f\|_{L^p}$$

for $1 < p < 2$, and the case where $2 < p < \infty$ follows by duality.

For the converse inequality, we use the identity $\|f\|_{L^2}^2 = \|Sf\|_{L^2(l^2)}^2$, which gives upon polarization that

$$\int_{\mathbf{R}^n} f\bar{g}\,dx = \int_{\mathbf{R}^n}\sum_{j=0}^{\infty} S_jf\overline{S_jg}\,dx.$$

Then

$$\begin{aligned}
\|f\|_{L^p} &= \sup\{\left|\int f\bar{g}\,dx\right| ;\ \|g\|_{L^{p'}} = 1\} \\
&= \sup\{\left|\int \sum S_jf\overline{S_jg}\,dx\right| ;\ \|g\|_{L^{p'}} = 1\} \\
&\leq \sup\{\|Sf\|_{L^p(l^2)}\|Sg\|_{L^{p'}(l^2)} ;\ \|g\|_{L^{p'}} = 1\} \\
&\leq C\sup\{\|Sf\|_{L^p(l^2)}\|g\|_{L^{p'}} ;\ \|g\|_{L^{p'}} = 1\} \\
&\leq C\|Sf\|_{L^p(l^2)}.
\end{aligned}$$

□

3.4. Besov and Triebel spaces

3.4.1. Definition of Besov and Triebel spaces. The discussion of Littlewood-Paley theory in §3.3 motivates the definition of Besov and Triebel spaces. As always, we begin by considering the L^2 case.

MOTIVATION. If $f \in \mathscr{S}(\mathbf{R}^n)$, let $S_j f = \varphi_j(D)f$ be the Littlewood-Paley pieces of f as in §3.3. Since $\langle\xi\rangle \sim 2^j$ on $\mathrm{supp}(\varphi_j)$, the $H^{s,2}$ norm of f can be expressed as

$$\begin{aligned}
\|f\|_{H^{s,2}}^2 &= \|\mathscr{F}^{-1}\{\langle\xi\rangle^s \hat{f}\}\|_{L^2}^2 = (2\pi)^{-n}\|\langle\xi\rangle^s \hat{f}\|_{L^2}^2 \\
&= (2\pi)^{-n}\sum_{j=0}^{\infty}\int_{\mathbf{R}^n}\varphi_j(\xi)^2\langle\xi\rangle^{2s}|\hat{f}(\xi)|^2\,d\xi \sim \sum_{j=0}^{\infty}2^{2js}\|\varphi_j(\xi)\hat{f}\|_{L^2}^2 \\
&\sim \sum_{j=0}^{\infty}2^{2js}\|\varphi_j(D)f\|_{L^2}^2 \\
&\sim \left\|\left(\sum_{j=0}^{\infty}2^{2js}|\varphi_j(D)f(x)|^2\right)^{1/2}\right\|_{L^2}^2 .
\end{aligned}$$

This shows that

$$\|f\|_{H^{s,2}} \sim \|(\varphi_j(D)f)\|_{l_2^s(L^2)} \sim \|(\varphi_j(D)f)\|_{L^2(l_2^s)}, \tag{1}$$

where l_p^s are the weighted sequence spaces introduced in §2.4.

The Besov and Triebel spaces are obtained by replacing the exponents 2 by p and q in the norms appearing in (1). For later purposes, we will need to use more general Littlewood-Paley partitions of unity in the definition of the spaces.

DEFINITION. If N is a positive integer, let Ψ_N be the set of all sequences $(\psi_j)_{j=0}^{\infty}$ of functions which satisfy

(a) $\psi_j \in C_c^\infty(\mathbf{R}^n)$ and $\psi_j \geq 0$ for all j,
(b) ψ_j is supported in $\{2^{j-N} \leq |\xi| \leq 2^{j+N}\}$ for $j \geq 1$, and ψ_0 is supported in $\{|\xi| \leq 2^N\}$,
(c) there exists $c > 0$ such that $\sum_{j=0}^{\infty}\psi_j(\xi) \geq c$, and
(d) for any multi-index α there exists $C_\alpha > 0$ such that

$$|D^\alpha\psi_j(\xi)| \leq C_\alpha\langle\xi\rangle^{-|\alpha|}, \quad j = 1, 2, \ldots .$$

We let $\Psi = \bigcup_{N=1}^{\infty}\Psi_N$.

Example. It is easy to check that $(\varphi_j)_{j=0}^{\infty}$ is a sequence in Ψ. We can generate more sequences in Ψ by choosing a nonnegative smooth function ψ supported in $\{2^{-N} \leq |\xi| \leq 2^N\}$ with $\psi > 0$ for $1/\sqrt{2} \leq |\xi| \leq \sqrt{2}$, and by letting $\psi_j(\xi) := \psi(2^{-j}\xi)$ for $j \geq 1$ and choosing ψ_0 in a suitable way. The normalized functions

$$\chi_j(\xi) := \frac{\psi_j(\xi)}{\sum_{k=-\infty}^{\infty} \psi(2^{-k}\xi)} \quad (j \geq 1), \quad \chi_0(\xi) := \frac{\sum_{k=-\infty}^{0} \psi(2^{-k}\xi)}{\sum_{k=-\infty}^{\infty} \psi(2^{-k}\xi)},$$

yield a sequence $(\chi_j) \in \Psi$ with $\sum_{j=0}^{\infty} \chi_j \equiv 1$. Further, if $(\psi_j) \in \Psi$ is given, we may use the above idea to find a sequence $(\tilde{\psi}_j) \in \Psi$ satisfying $\tilde{\psi}_j = 1$ on $\mathrm{supp}(\psi_j)$ for all $j \geq 0$.

We will use the constructions in the previous example several times below. After these preparations, we are ready to give the definition of Besov and Triebel spaces.

Definition. Suppose that $(\psi_j)_{j=0}^{\infty}$ is a sequence in Ψ, let s be a real number, and let $1 < p, q < \infty$.

(a) The *Besov space* $B_{pq}^s(\mathbf{R}^n)$ is the set of all $f \in \mathscr{S}'(\mathbf{R}^n)$ for which $(\psi_j(D)f) \in l_q^s(L^p)$. The norm is

$$\|f\|_{B_{pq}^s} := \|(\psi_j(D)f)\|_{l_q^s(L^p)} = \left(\sum_{j=0}^{\infty} [2^{js} \|\psi_j(D)f\|_{L^p}]^q \right)^{1/q}.$$

(b) The *Triebel space* (or *Triebel-Lizorkin space*) $F_{pq}^s(\mathbf{R}^n)$ consists of those $f \in \mathscr{S}'(\mathbf{R}^n)$ for which $(\psi_j(D)f) \in L^p(l_q^s)$. The norm is

$$\|f\|_{F_{pq}^s} := \|(\psi_j(D)f)\|_{L^p(l_q^s)} = \left\| \left(\sum_{j=0}^{\infty} \left[2^{js} |\psi_j(D)f|\right]^q \right)^{1/q} \right\|_{L^p}.$$

Remarks. 1. More precisely, the Besov space B_{pq}^s consists of those tempered distributions f for which each $\psi_j(D)f$ is an L^p function, and the sum $\sum [2^{js} \|\psi_j(D)f\|_{L^p}]^q$ converges.

2. For the Triebel spaces, we note that if $f \in \mathscr{S}'$ then any $\psi_j(D)f$ is a C^∞ function as the inverse Fourier transform of a compactly supported distribution. It follows that the map

$$x \mapsto \left(\sum_{j=0}^{\infty} \left[2^{js} |\psi_j(D)f(x)|\right]^q \right)^{1/q}$$

is nonnegative and measurable (as the qth root of the supremum of continuous functions). Then f is in F^s_{pq} iff the last function is in $L^p(\mathbf{R}^n)$.

3. One could also consider the case $0 < p, q \leq 1$, but we will restrict our attention to $1 < p, q < \infty$ for simplicity. Later we will discuss the Besov spaces also when $p = q = \infty$.

We now prove that the spaces defined above do not depend on the choice of the sequence (ψ_j). This will allow considerable flexibility when proving statements about these spaces. Below, we will usually give the details for Besov spaces and leave the Triebel case as an exercise.

THEOREM 1. *(Independence of partition of unity) Let $(\psi_j), (\tilde{\psi}_j)$ be two sequences in Ψ. Then for all admissible s, p, q,*

$$\|(\psi_j(D)f)\|_{l^s_q(L^p)} \sim \|(\tilde{\psi}_j(D)f)\|_{l^s_q(L^p)},$$
$$\|(\psi_j(D)f)\|_{L^p(l^s_q)} \sim \|(\tilde{\psi}_j(D)f)\|_{L^p(l^s_q)}.$$

Thus, any sequence $(\psi_j) \in \Psi$ in the definition of B^s_{pq} and F^s_{pq} will result in the same spaces with equivalent norms.

PROOF. 1. Let $f \in \mathscr{S}'$ be such that $(\psi_j(D)f) \in l^s_q(L^p)$. We wish to show that $(\tilde{\psi}_j(D)f) \in l^s_q(L^p)$, and

$$\|(\tilde{\psi}_j(D)f)\|_{l^s_q(L^p)} \leq C\|(\psi_j(D)f)\|_{l^s_q(L^p)}$$

with C independent of f.

2. Choose N so that $(\psi_j), (\tilde{\psi}_j) \in \Psi_N$. We write

$$\tilde{\psi}_j(D)f = \sum_{k=0}^{\infty} \tilde{\psi}_j(D)\chi(D)\psi_k(D)f$$

where

$$\chi(\xi) := \left(\sum_{k=0}^{\infty} \psi_k(\xi)\right)^{-1}.$$

Here χ is a bounded C^∞ function, and χ and $\tilde{\psi}_j$ satisfy the conditions in Mihlin's theorem with bounds independent of j by (a)–(d) above. Since $\tilde{\psi}_j(\xi)\psi_k(\xi) = 0$ if $|j - k| \geq 2N$, we have (with $\psi_k := 0$ for $k < 0$)

$$\begin{aligned} 2^{js}\|\tilde{\psi}_j(D)f\|_{L^p} &\leq C \sum_{j-2N \leq k \leq j+2N} 2^{js}\|\psi_k(D)f\|_{L^p} \\ &\leq C \sum_{j-2N \leq k \leq j+2N} 2^{ks}\|\psi_k(D)f\|_{L^p}. \end{aligned}$$

It follows that

$$
\begin{aligned}
\|(\tilde{\psi}_j(D)f)\|^q_{l^s_q(L^p)} &= \sum_{j=0}^{\infty} \left[2^{js}\|\tilde{\psi}_j(D)f\|_{L^p}\right]^q \\
&\le C\sum_{j=0}^{\infty}\left[\sum_{j-2N\le k\le j+2N} 2^{ks}\|\psi_k(D)f\|_{L^p}\right]^q \\
&\le C\sum_{j=0}^{\infty}\sum_{j-2N\le k\le j+2N}\left[2^{ks}\|\psi_k(D)f\|_{L^p}\right]^q \\
&\le C\sum_{j=0}^{\infty}\left[2^{js}\|\psi_j(D)f\|_{L^p}\right]^q .
\end{aligned}
$$

This proves the result for the Besov case.

3. The Triebel case is left as an exercise. It uses the vector-valued Mihlin multiplier theorem, which will be stated (for $q = 2$) below. □

3.4.2. Properties of Besov and Triebel spaces. We begin by proving that B^s_{pq} and F^s_{pq} are Banach spaces. To this end, we first state an abstract result which describes how the function spaces sit inside $l^s_q(L^p)$ and $L^p(l^s_q)$.

Definition. Let A and B be normed spaces. We say that a bounded linear operator $R : A \to B$ is a *retraction* is there is another bounded linear operator $S : B \to A$ with

$$RSb = b \quad \text{for all } b \in B.$$

Then S is called a *coretraction* for R.

Lemma 2. *If A is Banach and B is normed, and if $S : B \to A$ is a coretraction for $R : A \to B$, then B is Banach and S is an isomorphism from B onto a closed subspace of A.*

Proof. If $Sb = 0$ then $b = RSb = 0$, showing that S is injective. We claim that $SR : A \to A$ is a projection and the range of S is equal to the range of SR. Clearly $(SR)^2a = S(RS)Ra = SRa$ for $a \in A$ and $SR(A) \subseteq S(B)$. Also, any Sb can be written as $Sb = SR(Sb)$, so that $S(B) \subseteq SR(A)$. Then $S(B)$ is closed as the range of a projection.

If (b_j) is a Cauchy sequence in B, then (Sb_j) is Cauchy in A and there exists $Sb \in S(B)$ with $Sb_j \to Sb$ in A. Applying R, we see that

$b_j \to b$ in B and that B is complete. The map $S : B \to S(B)$ is an isomorphism by the open mapping theorem. $\square$

THEOREM 3. *(B^s_{pq} and F^s_{pq} as function spaces) B^s_{pq} and F^s_{pq} are Banach spaces, and $\mathscr{S} \subseteq B^s_{pq} \subseteq \mathscr{S}'$ and $\mathscr{S} \subseteq F^s_{pq} \subseteq \mathscr{S}'$ with continuous inclusions. Further, $\mathscr{S}$ is dense in B^s_{pq} and F^s_{pq}.*

PROOF. 1. We first show that B^s_{pq} is a normed space. Let (ψ_j) be a sequence in Ψ with $\sum_{j=0}^{\infty} \psi_j = 1$. If $f, g \in B^s_{pq}$ then

$$\begin{aligned} \|\lambda f + \mu g\|_{B^s_{pq}} &= \|\lambda(\psi_j(D)f) + \mu(\psi_j(D)g)\|_{l^s_q(L^p)} \\ &\leq |\lambda|\|f\|_{B^s_{pq}} + |\mu|\|g\|_{B^s_{pq}}, \end{aligned}$$

showing that B^s_{pq} is a vector space and $\|\cdot\|_{B^s_{pq}}$ obeys the triangle inequality. Also the other conditions for a norm are satisfied, for instance if f is in B^s_{pq} and $\|f\|_{B^s_{pq}} = 0$, then $\psi_j(D)f = 0$ for all j which implies that $\hat{f} = \sum_j \psi_j(\xi)\hat{f} = 0$ and $f = 0$.

2. To show that B^s_{pq} is complete, we define the maps

$$\begin{aligned} &S : B^s_{pq} \to l^s_q(L^p), \quad f \mapsto (\psi_j(D)f), \\ &R : l^s_q(L^p) \to B^s_{pq}, \quad (g_j) \mapsto \sum_{j=0}^{\infty} \tilde{\psi}_j(D)g_j, \end{aligned}$$

where $(\tilde{\psi}_j)$ is a sequence in Ψ with $\tilde{\psi}_j = 1$ on $\mathrm{supp}(\psi_j)$ for all $j \geq 0$. We leave as an exercise to show that the sum defining R converges in B^s_{pq}, and that S and R are bounded linear operators (this uses the argument in Theorem 1). We have that

$$RSf = \sum \tilde{\psi}_j(D)\psi_j(D)f = \sum \psi_j(D)f = f,$$

so S is a coretraction. Since $l^s_q(L^p)$ is Banach, Lemma 2 shows that B^s_{pq} is a Banach space.

3. Next we show that $\mathscr{S} \subseteq B^s_{pq}$. If $f \in \mathscr{S}$ then

$$\begin{aligned}
\|\psi_j(D)f\|_{L^p} &\leq C\|\langle x\rangle^{2n}\psi_j(D)f\|_{L^\infty} \\
&\leq C\|\mathscr{F}^{-1}\{\langle D_\xi\rangle^{2n}[\psi_j(\xi)\hat{f}(\xi)]\}\|_{L^\infty} \\
&\leq C\|\langle D_\xi\rangle^{2n}[\psi_j(\xi)\hat{f}(\xi)]\|_{L^1} \\
&\leq C\|\langle \xi\rangle^{2n}\langle D_\xi\rangle^{2n}[\psi_j(\xi)\hat{f}(\xi)]\|_{L^\infty} \\
&\leq C\sum_{|\alpha|,|\beta|\leq 2n}\|\langle \xi\rangle^{2n}D^\alpha\psi_j(\xi)D^\beta\hat{f}(\xi)\|_{L^\infty} \\
&\leq C2^{-j(s+1)}\sum_{|\beta|\leq 2n}\|\langle \xi\rangle^{2n+s+1}D^\beta\hat{f}(\xi)\|_{L^\infty}.
\end{aligned}$$

Thus $\|(2^{js}\|\psi_j(D)f\|_{L^p})\|_{l^q} \leq C\sum_{|\beta|\leq 2n}\|\langle \xi\rangle^{2n+s+1}D^\beta\hat{f}(\xi)\|_{L^\infty}$, showing that $\mathscr{S}$ is continuously embedded in B^s_{pq}. The embedding $B^s_{pq} \subseteq \mathscr{S}'$ is an exercise.

4. It remains to show that $\mathscr{S}$ is dense in B^s_{pq}. If $f \in B^s_{pq}$, we first note that the functions

$$f_N := \sum_{j=0}^N \psi_j(D)f$$

converge to f in B^s_{pq} as $N \to \infty$. This follows from the argument in the proof of Theorem 1. Thus it is enough to approximate each f_N by Schwartz functions.

Fix $\varepsilon > 0$, and let R and S be as in Step 2. If f_N is given, there is some $N' > 0$ such that $\psi_j(D)f_N = 0$ for $j > N'$. Since each $\psi_j(D)f_N$ is in L^p, we may find $\varphi_j \in \mathscr{S}$ such that

$$\|\psi_j(D)f_N - \varphi_j\|_{L^p} \leq 2^{-j(s+1)}\varepsilon, \quad 0 \leq j \leq N'.$$

Set $\varphi_j := 0$ for $j > N'$, and let $\varphi := R(\varphi_j) = \sum\tilde{\psi}_j(D)\varphi_j$. Since the sum is finite, we have $\varphi \in \mathscr{S}$, and

$$\begin{aligned}
\|f_N - \varphi\|_{B^s_{pq}} = \|RSf_N - R(\varphi_j)\|_{B^s_{pq}} &\leq C\|(\psi_j(D)f_N) - (\varphi_j)\|_{l^s_q(L^p)} \\
&= \left(\sum_{j=0}^{N'}\left[2^{js}\|\psi_j(D)f_N - \varphi_j\|_{L^p}\right]^q\right)^{1/q} \leq C_q\varepsilon.
\end{aligned}$$

This completes the proof. □

The following result gives some simple relations between the Besov and Triebel spaces.

THEOREM 4. *(Monotonicity of B^s_{pq} and F^s_{pq}) Let s be a real number and $1 < p, q < \infty$.*

(a) *The Besov spaces satisfy*

$$\begin{aligned} B^{s_2}_{pq} &\subseteq B^{s_1}_{pq} && \text{if } s_1 \le s_2, \\ B^{s}_{pq_1} &\subseteq B^{s}_{pq_2} && \text{if } q_1 \le q_2. \end{aligned}$$

These identities are true for the F-spaces as well.

(b) *If $p = q$, one has $B^s_{pp} = F^s_{pp}$.*

(c) *One has*

$$\begin{aligned} B^s_{pq} \subseteq F^s_{pq} \subseteq B^s_{pp} \quad &\text{if} \quad 1 < q \le p < \infty, \\ B^s_{pp} \subseteq F^s_{pq} \subseteq B^s_{pq} \quad &\text{if} \quad 1 < p \le q < \infty. \end{aligned}$$

PROOF. 1. The first two identities follow immediately from the inclusions $l^{s_2}_q \subseteq l^{s_1}_q$ and $l^{q_1} \subseteq l^{q_2}$:

$$\begin{aligned} \|f\|_{B^{s_1}_{pq}} &= \left(\sum_{j=0}^{\infty} \left[2^{js_1} \|\psi_j(D) f\|_{L^p} \right]^q \right)^{1/q} \\ &\le \left(\sum_{j=0}^{\infty} \left[2^{js_2} \|\psi_j(D) f\|_{L^p} \right]^q \right)^{1/q} = \|f\|_{B^{s_2}_{pq}}, \\ \|f\|_{B^{s}_{pq_1}} &= \left(\sum_{j=0}^{\infty} \left[2^{js} \|\psi_j(D) f\|_{L^p} \right]^{q_1} \right)^{1/q_1} \\ &\le \left(\sum_{j=0}^{\infty} \left[2^{js} \|\psi_j(D) f\|_{L^p} \right]^{q_2} \right)^{1/q_2} = \|f\|_{B^{s}_{pq_2}}, \end{aligned}$$

since if $(x_j) \in l^{q_1}$ then

$$\sum |x_j|^{q_2} \le \left(\sup |x_j|^{q_1}\right)^{\frac{q_2 - q_1}{q_1}} \sum |x_j|^{q_1} \le \left(\sum |x_j|^{q_1} \right)^{q_2/q_1}.$$

The proof for the F-spaces is similar.

2. If $f \in \mathscr{S}$ then

$$\begin{aligned} \|f\|^p_{B^s_{pp}} &= \sum 2^{jsp} \|\psi_j(D) f\|^p_{L^p} = \int \sum 2^{jsp} |\psi_j(D) f|^p \, dx \\ &= \left\| \left(\sum \left[2^{js} |\psi_j(D) f| \right]^p \right)^{1/p} \right\|^p_{L^p} = \|f\|_{F^s_{pp}}. \end{aligned}$$

3. In (c), the second inclusion on the first line and first inclusion on the second line follow from (a) and (b). To show the first inclusion, if $f \in B^s_{pq}$ and $1 < q \le p < \infty$ we have

$$\begin{aligned}\|f\|_{F^s_{pq}} &= \left\| \left(\sum \left[2^{js} |\psi_j(D) f| \right]^q \right)^{1/q} \right\|_{L^p} = \left\| \sum \left[2^{js} |\psi_j(D) f| \right]^q \right\|^{1/q}_{L^{p/q}} \\ &\le \left(\sum 2^{jsq} \|\psi_j(D) f\|^q_{L^p} \right)^{1/q} = \|f\|_{B^s_{pq}}.\end{aligned}$$

The proof of the last inclusion is analogous. □

Next we consider the relation of Besov and Triebel spaces to the Bessel potential spaces. Note that $\|f\|_{F^0_{p2}} = \|(\psi_j(D)f)\|_{L^p(l^2)}$, so the Littlewood-Paley theorem implies that $F^0_{p2} = L^p$ with equivalent norms. To prove a corresponding result for $s \neq 0$, we need a generalization of Mihlin's theorem. See [**2**] for a proof.

Theorem 5. *(Vector-valued Mihlin multiplier theorem) Suppose $m(\xi) := (m_{jk}(\xi))_{j,k=0}^\infty$ is a matrix of C^∞ complex valued functions on $\mathbf{R}^n$, which for any multi-index α satisfies*

$$\left(\sum_{j,k=0}^\infty |D^\alpha m_{jk}(\xi)|^2 \right)^{1/2} \le C_\alpha \langle \xi \rangle^{-|\alpha|}.$$

Then the operator $m(D)$, defined for sequences $F = (f_j)_{j=0}^\infty$ of Schwartz functions with $f_j = 0$ for j large, by

$$(m(D)F)_j := \sum_{k=0}^\infty m_{jk}(D) f_k, \quad j = 0, 1, 2, \ldots,$$

has a bounded extension $L^p(l^2) \to L^p(l^2)$ for any $1 < p < \infty$, and

$$\|m(D)F\|_{L^p(l^2)} \le C \|F\|_{L^p(l^2)}, \quad F \in L^p(l^2),$$

where C only depends on p, n, and finitely many C_α.

The inequality $\|(\varphi_j(D)f)\|_{L^p(l^2)} \le C\|f\|_{L^p}$ in the Littlewood-Paley theorem is an immediate consequence of the last result, if one takes $m_{j0}(\xi) := \varphi_j(\xi)$ and $m_{jk}(\xi) := 0$ for $k \ge 1$, and applies $m(D)$ to $F := (f, 0, 0, \ldots)$. The following proof uses a similar idea.

THEOREM 6. *(Fractional Sobolev spaces as Triebel spaces) If s is a real number and $1 < p < \infty$, then*

$$F^s_{p2} = H^{s,p},$$
$$F^s_{22} = B^s_{22} = H^{s,2}.$$

PROOF. 1. We show that $\|f\|_{F^s_{p2}} \le C\|f\|_{H^{s,p}}$ for any Schwartz function f. Let $f_0 := \langle D\rangle^s f$, let (ψ_j) be a sequence in Ψ, and define

$$m_{j0}(\xi) := 2^{js}\langle\xi\rangle^{-s}\psi_j(\xi),$$
$$m_{jk}(\xi) := 0 \quad (k \ge 1),$$
$$F := (f_0, 0, 0, \ldots).$$

Since $|D^\alpha m_{j0}(\xi)| \le C_\alpha\langle\xi\rangle^{-|\alpha|}\chi_{\mathrm{supp}(\psi_j)}(\xi)$ by the properties of (ψ_j), the matrix $m = (m_{jk})$ satisfies the conditions in the vector-valued Mihlin theorem. Consequently

$$\begin{aligned}\|f\|_{F^s_{p2}} &= \|(2^{js}\psi_j(D)f)\|_{L^p(l^2)} = \|(m_{j0}(D)f_0)\|_{L^p(l^2)} = \|m(D)F\|_{L^p(l^2)} \\ &\le C\|F\|_{L^p(l^2)} = C\|f_0\|_{L^p} = C\|f\|_{H^{s,p}}.\end{aligned}$$

2. Let us prove the converse inequality, $\|f\|_{H^{s,p}} \le C\|f\|_{F^s_{p2}}$ for $f \in \mathscr{S}$. If (ψ_k) is a sequence in Ψ with $\sum \psi_k = 1$, we have

$$\|f\|_{H^{s,p}} = \|\langle D\rangle^s f\|_{L^p} = \|\sum_k \psi_k(D)\langle D\rangle^s f\|_{L^p}.$$

The last sum converges in L^p by the Littlewood-Paley theorem for instance. We choose another sequence $(\tilde\psi_k)$ in Ψ with $\tilde\psi_k = 1$ on $\mathrm{supp}(\psi_k)$, and define

$$m_{0k}(\xi) := 2^{-ks}\langle\xi\rangle^s\psi_k(\xi),$$
$$m_{jk}(\xi) := 0 \quad (j \ge 1),$$
$$F := (2^{ks}\tilde\psi_k(D)f)_{k=0}^\infty.$$

Again, m satisfies the conditions in Mihlin's theorem. Since $\psi_k(D) = \psi_k(D)\tilde\psi_k(D)$, we have

$$\begin{aligned}\|f\|_{H^{s,p}} &\le \|\sum_k m_{0k}(D)F_k\|_{L^p} = \|m(D)F\|_{L^p(l^2)} \\ &\le C\|F\|_{L^p(l^2)} = C\|f\|_{F^s_{p2}}.\end{aligned}$$

3. We have proved that $\|f\|_{F^s_{p2}} \sim \|f\|_{H^{s,p}}$ for $f \in \mathscr{S}$. The standard density argument then shows that $F^s_{p2} = H^{s,p}$ with equivalent norms.

The second identity in the theorem is a consequence of the first one and the fact that $F^s_{pp} = B^s_{pp}$. □

Finally, we consider the interpolation of various function spaces. The following abstract result will be used to pass from interpolation results on L^p and l^s_q spaces to similar results for Besov spaces.

LEMMA 7. *Let (A_0, A_1) and (B_0, B_1) be two interpolation couples of Banach spaces, and assume that $R : A_0 + A_1 \to B_0 + B_1$ and $S : B_0 + B_1 \to A_0 + A_1$ are linear operators such that R is a retraction $A_j \to B_j$ with coretraction $S : B_j \to A_j$ $(j = 0, 1)$. If F is any interpolation functor, then S is a coretraction $F(B_0, B_1) \to F(A_0, A_1)$ which is an isomorphism onto a closed subspace of $F(A_0, A_1)$.*

PROOF. Since R is bounded $A_j \to B_j$, the definition of interpolation functor implies that R is bounded $F(A_0, A_1) \to F(B_0, B_1)$. Similarly, S is bounded $F(B_0, B_1) \to F(A_0, A_1)$. We also have $RSb = b$ for $b \in B_0$ and $b \in B_1$, thus also for $b \in F(B_0, B_1)$. Then R is a retraction on $F(A_0, A_1)$ with coretraction S, and the result follows from Lemma 2 above. □

The following interpolation results are not the sharpest possible (for better results see Triebel [**13**]), but they give an idea of what can be done.

THEOREM 8. *(Interpolation of function spaces) Let s_0, s_1 be real numbers, and let $1 < p_0, p_1, q_0, q_1 < \infty$. If $0 < \theta < 1$, let s, p, q be defined by*

$$s = (1-\theta)s_0 + \theta s_1, \quad \frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

(a) The Besov spaces satisfy

$$\begin{aligned} (B^{s_0}_{pq_0}, B^{s_1}_{pq_1})_{\theta,q} &= B^s_{pq} && \text{if } p_0 = p_1 = p, \\ (B^{s_0}_{p_0q_0}, B^{s_1}_{p_1q_1})_{\theta,q} &= B^s_{pq} && \text{if } p = q. \end{aligned}$$

(b) The Triebel spaces satisfy

$$\begin{aligned} (F^{s_0}_{p_0q_0}, F^{s_1}_{p_1q_1})_{\theta,p} &= B^s_{pp} && \text{if } s_0 \neq s_1, \\ (F^{s}_{p_0q_0}, F^{s}_{p_1q_1})_{\theta,p} &= B^s_{pp} && \text{if } s_0 = s_1 = s \text{ and } p = q, \\ (F^{s}_{p_0r}, F^{s}_{p_1r})_{\theta,p} &= F^s_{pr} && \text{if } s_0 = s_1 = s \text{ and } 1 < r < \infty. \end{aligned}$$

(c) The fractional Sobolev spaces satisfy

$$\begin{aligned} (H^{s_0,p}, H^{s_1,p})_{\theta,r} &= B^s_{pr} && \text{if } s_0 \neq s_1,\ p_0 = p_1 = p,\ 1 < r < \infty, \\ (H^{s,p_0}, H^{s,p_1})_{\theta,p} &= H^{s,p} && \text{if } s_0 = s_1 = s. \end{aligned}$$

PROOF. We prove (a). Consider the coretraction S given in the proof of Theorem 3,

$$S : B^{s_j}_{p_j q_j} \to l^{s_j}_{q_j}(L^{p_j}),\ \ f \mapsto (\psi_k(D)f)_{k=0}^{\infty} \quad (j = 0, 1).$$

By Lemma 7, S is an isomorphism from $(B^{s_0}_{p_0q_0}, B^{s_1}_{p_1q_1})_{\theta,q}$ onto a closed subspace of $(l^{s_0}_{q_0}(L^{p_0}), l^{s_1}_{q_1}(L^{p_1}))_{\theta,q}$. But by Theorems 2 and 3 in §2.4, we have with equivalent norms

$$(l^{s_0}_{q_0}(L^{p_0}), l^{s_1}_{q_1}(L^{p_1}))_{\theta,q} = l^s_q((L^{p_0}, L^{p_1})_{\theta,q}) = l^s_q(L^p)$$

if either $p_0 = p_1 = p$ or $p = q$. In these cases, since S is an isomorphism we have

$$\|f\|_{(B^{s_0}_{p_0q_0}, B^{s_1}_{p_1q_1})_{\theta,q}} \sim \|Sf\|_{l^s_q(L^p)} \sim \|f\|_{B^s_{pq}}.$$

The second and third identities in (b) follow in a similar way from a retraction argument. However, for the first identity one needs a more general interpolation result for the l^s_q spaces than the one proved in §2.4; see Triebel [**13**]. For the first identity in (c) we also refer to [**13**], and the second identity in (c) follows from the last identity in (b). □

3.5. Hölder and Zygmund spaces

In this section, we will see how the Hölder spaces fit into the framework of Besov and Triebel spaces. First, we extend the definition of Besov spaces to the case $p = q = \infty$. We use a sequence $(\psi_j)_{j=0}^{\infty}$ in Ψ such that $\sum_{j=0}^{\infty} \psi_j = 1$, $\psi_j(\xi) = \psi(2^{-j}\xi)$ for $j \geq 1$, and $\psi_j(-\xi) = \psi_j(\xi)$ for $j \geq 0$.

DEFINITION. If s is a real number, the space $B^s_{\infty\infty}(\mathbf{R}^n)$ consists of those $f \in \mathscr{S}'(\mathbf{R}^n)$ for which the sequence $(2^{js}\|\psi_j(D)f\|_{L^\infty})$ is bounded. The norm is

$$\|f\|_{B^s_{\infty\infty}} := \sup_{j\geq 0} 2^{js}\|\psi_j(D)f\|_{L^\infty}.$$

We will show that if $0 < s < 1$, the space $B^s_{\infty\infty}$ is precisely the space of Hölder continuous functions of exponent s. However, the equality breaks down if $s = 1$. To obtain a characterization for integer s, we will need to consider second order differences.

Definition. The *Zygmund space* $C_*^s(\mathbf{R}^n)$ for $s > 0$ is defined as follows: if $s = k + \gamma$ where k is a nonnegative integer and $0 < \gamma < 1$, we define $C_*^s(\mathbf{R}^n) := C^{k,\gamma}(\mathbf{R}^n)$ and

$$\|f\|_{C_*^s} := \|f\|_{C^{k,\gamma}} = \|f\|_{C^k} + \sum_{|\alpha|=k} \sup_{x,h} \frac{|\partial^\alpha f(x+h) - \partial^\alpha f(x)|}{|h|^\gamma}.$$

Further, if $s = k$ where k is a positive integer, then $C_*^k(\mathbf{R}^n)$ is the set of all $f \in C^{k-1}(\mathbf{R}^n)$ for which the following norm is finite:

$$\|f\|_{C_*^k} := \|f\|_{C^{k-1}} + \sum_{|\alpha|=k} \sup_{x,h} \frac{|\partial^\alpha f(x+h) - 2\partial^\alpha f(x) + \partial^\alpha f(x-h)|}{|h|}.$$

Example. The space C_*^1 consists of those continuous and bounded functions f for which $|f(x+h) - 2f(x) + f(x-h)| \leq C|h|$ for all x, h. Clearly one has the inclusions

$$C^1 \subseteq C^{0,1} \subseteq C_*^1.$$

These inclusions are strict: an example of a function in C_*^1 which is not Lipschitz continuous is given by the lacunary Fourier series

$$\sum_{k=1}^\infty 2^{-k} e^{i2^k x}.$$

Below, it will be useful to write $\psi_j(D)f$ as the convolution of f against a function. If $m \in C_c^\infty(\mathbf{R}^n)$ and $f \in L^1(\mathbf{R}^n)$ we have

$$\begin{aligned} m(D)f(x) &= \mathscr{F}^{-1}\{m(\xi)\hat{f}(\xi)\} = (2\pi)^{-n} \int_{\mathbf{R}^n} e^{ix\cdot\xi} m(\xi)\hat{f}(\xi)\,d\xi \\ &= \int_{\mathbf{R}^n} \left[(2\pi)^{-n} \int_{\mathbf{R}^n} e^{i(x-y)\cdot\xi} m(\xi)\,d\xi\right] f(y)\,dy \\ &= \int_{\mathbf{R}^n} \check{m}(x-y) f(y)\,dy. \end{aligned}$$

Thus $m(D)f = \check{m} * f$, where we have written $\check{m} := \mathscr{F}^{-1}m$. If $m \in \mathscr{O}_M$ and $f \in \mathscr{S}'$, with correct interpretations (see Schwartz [**10**]) it is possible to write $m(D)f = \check{m} * f$ where $\check{m}$ is a rapidly decreasing distribution.

Theorem 1. *(Hölder spaces as Besov spaces) If $s > 0$ is not an integer, we have*

$$C_*^s = B_{\infty\infty}^s.$$

PROOF. 1. Let us first assume $f \in C^\gamma_* = C^{0,\gamma}$ with $0 < \gamma < 1$. If $j \geq 1$ then $\int \check{\psi}_j(y)\,dy = \psi_j(0) = 0$ and

$$|\psi_j(D)f(x)| = \left|\int \check{\psi}_j(y)f(x-y)\,dy\right| = \left|\int \check{\psi}_j(y)[f(x-y)-f(x)]\,dy\right|$$
$$\leq \|f\|_{C^{0,\gamma}} \int |y|^\gamma |\check{\psi}_j(y)|\,dy.$$

Here $\check{\psi}_j(y) = \mathscr{F}^{-1}\{\psi(2^{-j}\cdot)\} = 2^{jn}\check{\psi}(2^j y)$, so

$$|\psi_j(D)f(x)| \leq 2^{-j\gamma}\|f\|_{C^{0,\gamma}} \int |y|^\gamma |\check{\psi}(y)|\,dy \leq C 2^{-j\gamma}\|f\|_{C^{0,\gamma}}.$$

Since $\|\psi_0(D)f\|_{L^\infty} \leq \|\check{\psi}_0\|_{L^1}\|f\|_{L^\infty}$ we get $2^{j\gamma}\|\psi_j(D)f\|_{L^\infty} \leq C\|f\|_{C^{0,\gamma}}$ for all $j \geq 0$, so that $C^{0,\gamma} \subseteq B^\gamma_{\infty\infty}$.

2. Let now $f \in C^{k+\gamma}_* = C^{k,\gamma}$ with $k \in \mathbf{Z}_+$ and $0 < \gamma < 1$. We use the Taylor expansion with remainder in integral form,

$$f(x-y) = \sum_{|\alpha|<k} \frac{\partial^\alpha f(x)}{\alpha!}(-y)^\alpha + \sum_{|\alpha|=k} k \int_0^1 (1-t)^{k-1}\frac{\partial^\alpha f(x-ty)}{\alpha!}(-y)^\alpha\,dt.$$

Since $\int (-y)^\alpha \check{\psi}_j(y)\,dy = D^\alpha \psi_j(0) = 0$ for $j \geq 1$, we have

$$\psi_j(D)f(x) = \int \check{\psi}_j(y)f(x-y)\,dy$$
$$= \int \check{\psi}_j(y) \sum_{|\alpha|=k} k \int_0^1 (1-t)^{k-1}\frac{\partial^\alpha f(x-ty)}{\alpha!}(-y)^\alpha\,dt\,dy$$
$$= \int \check{\psi}_j(y) \sum_{|\alpha|=k} k \int_0^1 (1-t)^{k-1}\frac{\partial^\alpha f(x-ty) - \partial^\alpha f(x)}{\alpha!}(-y)^\alpha\,dt\,dy.$$

Taking absolute values, we obtain

$$|\psi_j(D)f(x)| \leq C\left(\int |y|^{k+\gamma}|\check{\psi}_j(y)|\,dy\right)\|f\|_{C^{k,\gamma}} \leq C2^{-j(k+\gamma)}\|f\|_{C^{k,\gamma}}.$$

Thus $C^{k,\gamma} \subseteq B^{k+\gamma}_{\infty\infty}$.

3. Assume now that $f \in B^s_{\infty\infty}$ where $0 < s < 1$. We need to show that $f \in C^0$ and $|f(x+h) - f(x)| \leq C|h|^s$. Since $\sum \psi_j = 1$, we have

$$\|f\|_{L^\infty} \leq \sum_{j=0}^\infty \|\psi_j(D)f\|_{L^\infty} \leq C_s\|f\|_{B^s_{\infty\infty}}.$$

More precisely, the sum $\sum_{j=0}^\infty \psi_j(D)f$ converges in C^0, and since it also converges in $\mathscr{S}'$ to f we must have $f \in C^0$.

Let now $h \in \mathbf{R}^n$ with $|h| \leq 1$ (if $|h| > 1$ the claim is trivial). Choose k so that $2^{-k-1} < |h| \leq 2^{-k}$. Also, choose a sequence $(\chi_j)_{j=0}^{\infty}$ with $\chi_j = 1$ on $\mathrm{supp}(\psi_j)$, and $\chi_j(\xi) = \chi(2^{-j}\xi)$ for $j \geq 1$ where $\chi \in C_c^{\infty}(\mathbf{R}^n)$. Let $f_j := \psi_j(D)f = \chi_j(D)\psi_j(D)f$. Then for $j \geq 1$,

$$\begin{aligned} f_j(x+h) - f_j(x) &= \int [\check{\chi}_j(x+h-y) - \check{\chi}_j(x-y)]\psi_j(D)f(y)\,dy \\ &= \int [\check{\chi}_j(y+h) - \check{\chi}_j(y)]\psi_j(D)f(x-y)\,dy \\ &= \int [\check{\chi}(y+2^j h) - \check{\chi}(y)]\psi_j(D)f(x-2^{-j}y)\,dy. \end{aligned}$$

Using that $\check{\chi}(y+z) - \check{\chi}(y) = \int_0^1 \nabla\check{\chi}(y+tz)\cdot z\,dt$, we estimate

$$\int |\check{\chi}(y+2^j h) - \check{\chi}(y)|\,dy \leq \begin{cases} C2^j|h|, & j \leq k, \\ C, & j > k. \end{cases}$$

It follows that

$$\begin{aligned} |f(x+h) - f(x)| &\leq \sum_{j=0}^{k} |f_j(x+h) - f_j(x)| + \sum_{j=k+1}^{\infty} |f_j(x+h) - f_j(x)| \\ &\leq C\sum_{j=0}^{k} 2^j|h|\|\psi_j(D)f\|_{L^\infty} + C\sum_{j=k+1}^{\infty} \|\psi_j(D)f\|_{L^\infty} \\ &\leq C|h|^s\|f\|_{B^s_{\infty\infty}}\left(\sum_{j=0}^{k} 2^{j(1-s)}|h|^{1-s} + |h|^{-s}\sum_{j=k+1}^{\infty} 2^{-js}\right) \\ &\leq C|h|^s\|f\|_{B^s_{\infty\infty}}, \end{aligned}$$

since $|h| \sim 2^{-k}$.

4. Finally, assume $f \in B^s_{\infty\infty}$ where $s = k+\gamma$ and $k \in \mathbf{Z}_+$, $0 < \gamma < 1$. We first show that $f \in C^k$. If (χ_j) is as in Step 3, we have for $j \geq 1$

$$\begin{aligned} \chi_j(D)D^\alpha f &= \check{\chi}_j * D^\alpha f = D^\alpha\check{\chi}_j * f \\ &= 2^{j|\alpha|}2^{jn}(D^\alpha\check{\chi})(2^j\,\cdot\,) * f. \end{aligned}$$

Therefore, if $|\alpha| \leq k$ then

$$\begin{aligned} \|D^\alpha f\|_{L^\infty} &\leq \sum_{j=0}^{\infty} \|\psi_j(D)D^\alpha f\|_{L^\infty} \leq \sum_{j=0}^{\infty} \|\chi_j(D)D^\alpha\psi_j(D)f\|_{L^\infty} \\ &\leq C\sum_{j=0}^{\infty} 2^{j|\alpha|}\|\psi_j(D)f\|_{L^\infty} \leq C\|f\|_{B^s_{\infty\infty}}. \end{aligned}$$

The fact that $D^\alpha f \in C^{0,\gamma}$ for $|\alpha| = k$ follows by repeating the argument in Step 3 with $f_j := \psi_j(D)D^\alpha f$. $\square$

In the next result, we consider the case of integer s.

Theorem 2. *(Zygmund spaces as Besov spaces) If $s > 0$ is an integer, then $C_*^s = B_{\infty\infty}^s$.*

Proof. 1. We shall only prove that $C_*^1 = B_{\infty\infty}^1$. Assume that f is a function in C_*^1. If $j \geq 1$ we have $\int \check{\psi}_j(y)\,dy = 0$ and $\check{\psi}_j(-y) = \check{\psi}_j(y)$ (because $\psi_j(-\xi) = \psi_j(\xi)$), so that

$$\begin{aligned}
|\psi_j(D)f(x)| &= \left| \int \check{\psi}_j(y) f(x-y)\,dy \right| \\
&= \left| \frac{1}{2} \int \check{\psi}_j(y)[f(x-y) - 2f(x) + f(x+y)]\,dy \right| \\
&\leq \frac{1}{2} \|f\|_{C_*^1} \int |y||\check{\psi}_j(y)|\,dy \leq C 2^{-j} \|f\|_{C_*^1}.
\end{aligned}$$

Also $|\psi_0(D)f(x)| \leq \|\check{\psi}_0\|_{L^1} \|f\|_{L^\infty} \leq C\|f\|_{C_*^1}$, so we have $C_*^1 \subseteq B_{\infty\infty}^1$.

2. Let now $f \in B_{\infty\infty}^1$. Then $f \in B_{\infty\infty}^{1-\varepsilon} \subseteq C^0$ by Theorem 1. We need to show that

$$|f(x+h) - 2f(x) + f(x-h)| \leq C|h|, \quad x, h \in \mathbf{R}^n.$$

It is enough to show this for $|h| \leq 1$. As in the proof of Theorem 1, choose k so that $2^{-k-1} < |h| \leq 2^{-k}$, and choose $(\chi_j)_{j=0}^\infty \in \Psi$ with $\chi_j = 1$ on $\mathrm{supp}(\psi_j)$, $\chi_j(\xi) = \chi(2^{-j}\xi)$ for $j \geq 1$. Letting $f_j := \psi_j(D)f = \chi_j(D)\psi_j(D)f$, we have

$$\begin{aligned}
&f_j(x+h) - 2f_j(x) + f_j(x-h) \\
&= \int [\check{\chi}(y + 2^j h) - 2\check{\chi}(y) + \check{\chi}(y - 2^j h)] \psi_j(D) f(x - 2^{-j}y)\,dy.
\end{aligned}$$

We use the estimate

$$\int |\check{\chi}(y + 2^j h) - 2\check{\chi}(y) + \check{\chi}(y - 2^j h)|\,dy \leq \left\{ \begin{array}{ll} C 2^{2j} |h|^2, & j \leq k, \\ C, & j > k. \end{array} \right.$$

The second part of this follows from the triangle inequality, and the first part is a consequence of the Taylor expansion

$$h(z) = h(0) + \nabla h(0) \cdot z + \sum_{|\alpha|=2} 2 \int_0^1 (1-t) \frac{\partial^\alpha h(tz)}{\alpha!} z^\alpha\,dt$$

applied to $h(z) := \check{\chi}(y+z) - 2\check{\chi}(y) + \check{\chi}(y-z)$. We obtain

$$
\begin{aligned}
|f(x+h) - 2f(x) + f(x-h)| &\leq \sum_{j=0}^{k} |f_j(x+h) - 2f_j(x) + f_j(x-h)| \\
&\quad + \sum_{j=k+1}^{\infty} |f_j(x+h) - 2f_j(x) + f_j(x-h)| \\
&\leq C \sum_{j=0}^{k} 2^{2j} |h|^2 \|\psi_j(D) f\|_{L^\infty} + C \sum_{j=k+1}^{\infty} \|\psi_j(D) f\|_{L^\infty} \\
&\leq C|h| \|f\|_{B^1_{\infty\infty}} \left(\sum_{j=0}^{k} 2^j |h| + |h|^{-1} \sum_{j=k+1}^{\infty} 2^{-j} \right) \\
&\leq C|h| \|f\|_{B^1_{\infty\infty}},
\end{aligned}
$$

since $|h| \sim 2^{-k}$. □

Definition. We will write $C^s_* := B^s_{\infty\infty}$ for $s \leq 0$.

Finally, we remark that it is possible to characterize the Besov spaces B^s_{pq} in terms of difference quotients also when $p, q < \infty$. Define

$$
\begin{aligned}
\omega_p(t, f) &:= \sup_{|y|<t} \|f(\,\cdot\, + y) - f(\,\cdot\,)\|_{L^p}, \\
\omega_p^2(t, f) &:= \sup_{|y|<t} \|f(\,\cdot\, + y) - 2f(\,\cdot\,) + f(\,\cdot\, - y)\|_{L^p}.
\end{aligned}
$$

The following result may be found, also in a more general form, in Bergh-Löfström [**1**].

Theorem 3. *(Finite difference characterization of Besov spaces) Let $s > 0$ and $1 < p, q < \infty$. If $s = k + \gamma$ where k is a nonnegative integer and $0 < \gamma < 1$, then*

$$
\|f\|_{B^s_{pq}} \sim \|f\|_{L^p} + \sum_{j=1}^{n} \left(\int_0^\infty [t^{-\gamma} \omega_p(t, D_j^k f)]^q \frac{dt}{t} \right)^{1/q}.
$$

If $s = k + 1$ where k is a nonnegative integer, then

$$
\|f\|_{B^s_{pq}} \sim \|f\|_{L^p} + \sum_{j=1}^{n} \left(\int_0^\infty [t^{-1} \omega_p^2(t, D_j^k f)]^q \frac{dt}{t} \right)^{1/q}.
$$

3.6. Embedding theorems

We will finish the chapter with a number of embedding theorems for the different spaces. The following L^p estimates for convolutions will be useful for this purpose. Part (c) is called Young's inequality.

THEOREM 1. *(L^p estimates for convolution) Let $1 \le p \le \infty$.*

(a) *If $f \in L^1$ and $g \in L^p$, then $f * g \in L^p$ and*

$$\|f * g\|_{L^p} \le \|f\|_{L^1} \|g\|_{L^p}.$$

(b) *If $f \in L^p$ and $g \in L^{p'}$, then $f * g \in L^\infty$ and*

$$\|f * g\|_{L^\infty} \le \|f\|_{L^p} \|g\|_{L^{p'}}.$$

(c) *If $f \in L^p$ and $g \in L^q$, where $1 \le q \le p'$, then $f * g \in L^r$ and*

$$\|f * g\|_{L^r} \le \|f\|_{L^p} \|g\|_{L^q}$$

if r is defined by $\frac{1}{r} = \frac{1}{p} - \frac{1}{q'}$.

PROOF. (a) It is proved in [**7**] (or Rudin [**8**]) that the convolution of two L^1 functions is an L^1 function, and $\|f * g\|_{L^1} \le \|f\|_{L^1} \|g\|_{L^1}$. If $f \in L^1$ and $g \in L^\infty$, we have the trivial bound $\|f * g\|_{L^\infty} \le \|f\|_{L^1} \|g\|_{L^\infty}$. Thus, if $f \in L^1$ we have a map $T : L^1 + L^\infty \to L^1 + L^\infty$, $Tg = f * g$. Since

$$\begin{aligned} T &: L^1 \to L^1, \qquad && \|T\| \le \|f\|_{L^1}, \\ T &: L^\infty \to L^\infty, \qquad && \|T\| \le \|f\|_{L^1}, \end{aligned}$$

interpolation gives that T maps L^p to L^p with $\|Tg\|_{L^p} \le \|f\|_{L^1} \|g\|_{L^p}$.

(b) This is Hölder's inequality:

$$|f * g(x)| = \left| \int f(x-y) g(y) \, dy \right| \le \|f(x - \,\cdot\,)\|_{L^p} \|g\|_{L^{p'}} = \|f\|_{L^p} \|g\|_{L^{p'}}.$$

(c) If $f \in L^p$, by (a) and (b) the map $T : g \mapsto f * g$ satisfies

$$\begin{aligned} T &: L^1 \to L^p, \qquad && \|T\| \le \|f\|_{L^p}, \\ T &: L^{p'} \to L^\infty, \qquad && \|T\| \le \|f\|_{L^p}. \end{aligned}$$

Let $1 < q < p'$, and let $0 < \theta < 1$ be such that $\frac{1}{q} = \frac{1-\theta}{1} + \frac{\theta}{p'}$. Then $\theta = \frac{p}{q'}$ and $\frac{1-\theta}{p} + \frac{\theta}{\infty} = \frac{1}{p} - \frac{1}{q'} = \frac{1}{r}$. Thus interpolation shows that $T : L^q \to L^r$ with $\|Tg\|_{L^r} \le \|f\|_{L^p} \|g\|_{L^q}$, as required. $\square$

THEOREM 2. *(Embeddings for function spaces) Assume that* $s \geq s_1$ *and* $1 < p, p_1 < \infty$*, and*

$$s - \frac{n}{p} = s_1 - \frac{n}{p_1}.$$

Then for any $1 < r < \infty$ *one has*

$$\begin{aligned} B^s_{pr} &\subseteq B^{s_1}_{p_1 r}, \\ F^s_{pr} &\subseteq F^{s_1}_{p_1 r}, \\ H^{s,p} &\subseteq H^{s_1,p_1}. \quad . \end{aligned}$$

Further, if t *is any real number and* $1 < p, r < \infty$*, then*

$$\begin{aligned} B^t_{pr} &\subseteq C^{t-n/p}_*, \\ F^t_{pr} &\subseteq C^{t-n/p}_*, \\ H^{t,p} &\subseteq C^{t-n/p}_*. \end{aligned}$$

PROOF. 1. We prove the first embedding for Besov spaces. If $s = s_1$ then $p = p_1$ and the claim is trivial. We may assume $s > s_1$, and then $p < p_1 < \infty$. Let $(\psi_j)_{j=0}^\infty$ and $(\chi_j)_{j=0}^\infty$ be two sequences in Ψ with $\chi_j = 1$ on $\mathrm{supp}(\psi_j)$, and $\chi_j(\xi) = \chi(2^{-j}\xi)$ for $j \geq 1$. Then

$$\begin{aligned} \|\psi_j(D)f\|_{L^{p_1}} &= \|\chi_j(D)\psi_j(D)f\|_{L^{p_1}} = \|\check{\chi}_j * \psi_j(D)f\|_{L^{p_1}} \\ &\leq \|\check{\chi}_j\|_{L^q}\|\psi_j(D)f\|_{L^p} \end{aligned}$$

by Young's inequality, if q is chosen so that $\frac{1}{p_1} = \frac{1}{p} - \frac{1}{q'}$. If $j \geq 1$ we have

$$\|\check{\chi}_j\|_{L^q} = \|2^{jn}\check{\chi}(2^j\,\cdot\,)\|_{L^q} = C2^{jn/q'} = C2^{j(s-s_1)}$$

since

$$\frac{1}{q'} = \frac{1}{p} - \frac{1}{p_1} = \frac{s-s_1}{n}.$$

If $j = 0$ then $\|\psi_0(D)f\|_{L^{p_1}} \leq \|\check{\chi}_0\|_{L^q}\|\psi_0(D)f\|_{L^p} = C\|\psi_0(D)f\|_{L^p}$. Consequently, we obtain

$$\|f\|_{B^{s_1}_{p_1 r}} = \|(2^{js_1}\|\psi_j(D)f\|_{L^{p_1}})\|_{l^r} \leq C\|(2^{js}\|\psi_j(D)f\|_{L^p})\|_{l^r} = C\|f\|_{B^s_{pr}}.$$

This holds for Schwartz functions f, and the first embedding follows since Schwartz functions are dense.

2. For the second embedding for Besov spaces, we note that

$$2^{j(t-n/p)}\|\psi_j(D)f\|_{L^\infty} \leq 2^{j(t-n/p)}\|\check{\chi}_j\|_{L^{p'}}\|\psi_j(D)f\|_{L^p} \leq C2^{jt}\|\psi_j(D)f\|_{L^p},$$

since $\|\check{\chi}_j\|_{L^{p'}} = C2^{jn/p}$. Consequently

$$\|f\|_{C_*^{t-n/p}} \sim \sup_{j\geq 0} 2^{j(t-n/p)}\|\psi_j(D)f\|_{L^\infty} \leq C\|f\|_{B^t_{pr}}$$

for any r.

3. The first embedding for Triebel spaces uses the Hardy-Littlewood-Sobolev inequality, and we refer to Triebel [**13**] for the proof. The second embedding however follows directly from the corresponding embedding for Besov spaces, upon noting that by Theorem 4 in §3.4, for any r there is some r_1 such that

$$F^t_{pr} \subseteq B^t_{pr_1} \subseteq C_*^{t-n/p}.$$

4. The embedding results for fractional Sobolev spaces follow from the corresponding results for Triebel spaces, using the fact that $H^{t,p} = F^t_{p2}$ by Theorem 6 in §3.4. □

Remark. Considering the spaces B^s_{pq}, F^s_{pq}, and $H^{s,p}$, the number $s - \frac{n}{p}$ is called the *differential dimension*. For the Zygmund space $C^s_* = B^s_{\infty\infty}$, the differential dimension is s. The embeddings considered here can be memorized by noting that the differential dimension is preserved while the smoothness index becomes smaller.

Bibliography

[1] J. Bergh and J. Löfström, *Interpolation spaces*, Springer, 1976.

[2] J. Duoandikoetxea, *Fourier analysis*, AMS, 2001.

[3] R. Edwards and G. Gaudry, *Littlewood-Paley and multiplier theory*, Springer-Verlag, Berlin, 1977.

[4] L. C. Evans, *Partial differential equations*, AMS, 1998.

[5] D. Jerison and C. E. Kenig, *The inhomogeneous Dirichlet problem in Lipschitz domains*, J. Funct. Anal. **130** (1995), p. 161–219.

[6] L. Grafakos, *Classical and modern Fourier analysis*, Pearson Education, 2004.

[7] I. Holopainen, *Reaalianalyysi 1*, lecture notes (in Finnish), available at the webpage http://www.helsinki.fi/∼iholopai/ReAn02.pdf.

[8] W. Rudin, *Real and complex analysis*, 3rd edition, McGraw-Hill, 1986.

[9] W. Rudin, *Functional analysis*, 2nd edition, McGraw-Hill, 1991.

[10] L. Schwartz, *Théorie des distributions*, Hermann, 1950.

[11] E. Stein, *Singular integrals and differentiability properties of functions*, Princeton University Press, 1970.

[12] E. Stein, *Harmonic analysis*, Princeton University Press, 1993.

[13] H. Triebel, *Interpolation theory, function spaces, differential operators*, North-Holland, Amsterdam, 1978.