

Operators on Bochner spaces

One of the central questions in the analysis of Banach space-valued functions is the so-called *L^p -extension problem*: Given a bounded linear operator T defined on $L^p(S)$, when does the prescription

$$(T \otimes I_X)(f \otimes x) := Tf \otimes x$$

define a bounded linear operator $T \otimes I_X$ on $L^p(S; X)$? In Section 2.1 we shall see that this happens in a variety of situations, for instance when X is a Hilbert space or when the operator T is positivity preserving. On the other hand, a necessary condition for the boundedness of $T \otimes I_X$ for every Banach space X is that T be dominated by a positive operator. This spells trouble for operators T whose boundedness on $L^p(S)$ depends on cancellation properties, such as the Fourier transform, the Hilbert transform and more general Calderón–Zygmund operators in harmonic analysis, as well as the Itô isometry in stochastic analysis. Towards the end of Section 2.1 we shall demonstrate that these operators provide counterexamples to the general L^p -extension problem. It will be a major theme in the following chapters to identify, for specific operators T , such as the ones just mentioned, the precise classes of Banach spaces X which have the property that $T \otimes I_X$ is bounded.

In some situations the boundedness question for operators on $L^p(S; X)$ can be treated in full generality. One such situation is considered in Section 2.2, where we present vector-valued versions of the classical interpolation theorems of Riesz–Thorin and Marcinkiewicz, with special attention to the constants involved in the latter case. We also include a self-contained presentation of complex and real interpolation of the spaces $L^p(S; X)$. The next section, Section 2.3, is concerned with the vector-valued versions of some classical results in real analysis such as the Hardy–Littlewood maximal function and the Lebesgue differentiation theorem.

The remainder of the chapter is devoted to three topics which can be read quite independently, although the results will be used frequently in later chapters and in the future volumes. They provide some first illustrations of the

methods developed so far. In Section 2.5 we study differentiability properties of X -valued functions and develop some elements of vector-valued Sobolev spaces and their fractional counterparts.

We continue with a discussion of the vector-valued Fourier transform in Section 2.4. The failure of the Fourier–Plancherel theorem in Bochner spaces $L^2(\mathbb{R}^d; X)$ for all Banach spaces X not isomorphic to a Hilbert space naturally leads to the notion of Fourier type, which is studied in some detail. We also include a brief discussion of vector-valued Schwartz functions and tempered distributions.

In the final section, Section 2.6, we study vector-valued conditional expectations. These are not only of interest in probability theory and stochastic analysis, but also in harmonic analysis, where they appear as averaging operators in approximation schemes; this fact serves as a motivation for setting up the theory in the σ -finite framework.

2.1 The L^p -extension problem

A central issue in Banach space-valued analysis is the following extension problem. Let $(S, \mathcal{A}, \mu)$ be a measure space and suppose that a bounded linear operator T on $L^p(S)$ is given. Let I_X denote the identity operator on X . We may define a linear operator $T \otimes I_X$ on $L^p(S) \otimes X$ by the formula

$$(T \otimes I_X)(f \otimes x) := Tf \otimes x.$$

We leave it to the reader to check that this operator is well defined. For $1 \leq p < \infty$, $L^p(S) \otimes X$ is dense in $L^p(S; X)$ and it makes sense to ask whether $T \otimes I_X$ extends to a bounded operator on $L^p(S; X)$. More generally, the analogous question may be posed for operators acting between two different L^p -spaces. Even for $p = 2$ such a bounded extension generally do not exist without additional assumptions on T or X . In fact, a large part of these volumes is concerned with identifying situations where bounded extensions do exist.

We begin with an extension result for operators $T : L^1(S_1) \rightarrow L^p(S_2)$, which holds for arbitrary Banach spaces X . Its proof is an exercise in using the triangle inequality:

Proposition 2.1.1. *Let $1 \leq p \leq \infty$, let $T : L^1(S_1) \rightarrow L^p(S_2)$ be a bounded linear operator, and let X be a Banach space. Then $T \otimes I_X$ extends uniquely to a bounded operator from $L^1(S_1; X)$ to $L^p(S_2; X)$ and we have*

$$\|T \otimes I_X\| = \|T\|.$$

Proof. For μ_1 -simple functions of the form $\sum_{n=1}^N \mathbf{1}_{A_n} \otimes x_n$ with the sets $A_n \in \mathcal{A}_1$ disjoint and of finite μ_1 -measure, the triangle inequality gives

$$\left\| \sum_{n=1}^N T\mathbf{1}_{A_n} \otimes x_n \right\|_{L^p(S_2; X)} \leq \sum_{n=1}^N \|T\mathbf{1}_{A_n} \otimes x_n\|_{L^p(S_2; X)}$$

$$\begin{aligned}
&= \sum_{n=1}^N \|T \mathbf{1}_{A_n}\|_{L^p(S_2)} \|x_n\| \\
&\leq \|T\| \sum_{n=1}^N \|\mathbf{1}_{A_n}\|_{L^1(S_1)} \|x_n\| \\
&= \|T\| \sum_{n=1}^N \|\mathbf{1}_{A_n} \otimes x_n\|_{L^1(S_1; X)} \\
&= \|T\| \left\| \sum_{n=1}^N \mathbf{1}_{A_n} \otimes x_n \right\|_{L^1(S_1; X)}.
\end{aligned}$$

This proves the boundedness of $T \otimes I_X$ along with the estimate $\|T \otimes I_X\| \leq \|T\|$. The reverse estimate holds trivially. $\square$

Slightly more involved is the following extension result, which nevertheless covers many interesting examples.

Proposition 2.1.2. *Let $1 \leq p < \infty$ and let X be isomorphic to a closed subspace of a quotient of a space $L^p(S')$. Then every bounded operator T on $L^p(S)$ has a bounded extension $T \otimes I_X$ to $L^p(S; X)$, with*

$$\|T \otimes I_X\|_{\mathcal{L}(L^p(S; X))} = \|T\|_{\mathcal{L}(L^p(S))}$$

if the isomorphism is isometric.

Examples of closed subspaces of $L^p(0, 1)$ include:

- ℓ^2 , if $1 \leq p < \infty$;
- ℓ^q , if $p \leq q \leq 2$;
- $L^q(0, 1)$, if $p \leq q \leq 2$.

In each of these three examples it is possible to find *isometric* copies of these spaces inside $L^p(0, 1)$. We will show how to find these for ℓ^2 in Subsection 2.1.b and refer to the Notes and Example 2.2.8 for the other two cases.

Proof of Proposition 2.1.2. We prove the proposition in three steps.

Step 1 – Let $X \subseteq L^p(S')$ be a closed subspace and $x_n \in X$, $g_n \in L^p(S)$ for $n = 1, \dots, N$. Since these functions, as well as $Tg_n \in L^p(S)$, are supported on σ -finite subsets of S' or S , we can use Fubini's theorem to estimate the action of $T \otimes I_X$ on $f = \sum_{n=1}^N g_n \otimes x_n$:

$$\begin{aligned}
\|(T \otimes I_X)f\|_{L^p(S; X)}^p &= \int_{S'} \left\| \sum_{n=1}^N Tg_n \otimes x_n(s') \right\|_{L^p(S)}^p d\mu'(s') \\
&\leq \|T\|^p \int_{S'} \left\| \sum_{n=1}^N g_n \otimes x_n(s') \right\|_{L^p(S)}^p d\mu'(s') = \|T\|^p \|f\|_{L^p(S; X)}^p.
\end{aligned}$$

Step 2 – Let $X \approx L^p(S')/Z$, where Z is a closed subspace of $L^p(S')$, and let $q_Z : L^p(S') \rightarrow L^p(S')/Z$ be the quotient mapping. The mapping $I_{L^p(S)} \otimes q_Z$ is bounded (with $\|I_{L^p(S)} \otimes q_Z\| = \|q_Z\|$) and $T \otimes I_{L^p(S')/Z} = (I_{L^p(S)} \otimes q_Z) \circ (T \otimes I_X)$. This gives the extendability to quotients of $L^p(S')$ with an obvious estimate in terms of the norms of q_Z and the isomorphism constant of $X \approx L^p(S')/Z$.

Step 3 – The case where X is isomorphic to a closed subspace of a quotient of $L^p(S')$ follows trivially from Step 2 (with again an obvious estimate for the norm of the extension). $\square$

2.1.a Boundedness of $T \otimes I_X$ for positive operators T

We proceed with an extension result for *positive* operators T (these are operators satisfying $Tf \geq 0$ for all $f \geq 0$).

Theorem 2.1.3. $1 \leq p_1 < \infty$ and $1 \leq p_2 \leq \infty$, let $T : L^{p_1}(S_1) \rightarrow L^{p_2}(S_2)$ be a bounded linear operator, and let X be a Banach space. If T is positive, then $T \otimes I_X$ extends uniquely to a bounded operator from $L^{p_1}(S_1; X)$ to $L^{p_2}(S_2; X)$ and we have

$$\|T \otimes I_X\| = \|T\|.$$

Proof. Let $f \in L^{p_1}(S_1) \otimes X$ be a μ_1 -simple function, say $f = \sum_{n=1}^N \mathbf{1}_{A_n} \otimes x_n$ with the sets $A_n \in \mathcal{A}_1$ disjoint. The positivity of T implies that $|T\mathbf{1}_{A_n}| = T\mathbf{1}_{A_n}$ and

$$\begin{aligned} \left\| (T \otimes I_X) \sum_{n=1}^N \mathbf{1}_{A_n} \otimes x_n \right\|_{L^{p_2}(S_2; X)} &= \left(\int_{S_2} \left\| \sum_{n=1}^N T\mathbf{1}_{A_n} \otimes x_n \right\|^{p_2} d\mu_2 \right)^{1/p_2} \\ &\leq \left(\int_{S_2} \left(\sum_{n=1}^N |T\mathbf{1}_{A_n}| \|x_n\| \right)^{p_2} d\mu_2 \right)^{1/p_2} \\ &= \left(\int_{S_2} \left(T \sum_{n=1}^N \mathbf{1}_{A_n} \|x_n\| \right)^{p_2} d\mu_2 \right)^{1/p_2} \\ &= \left\| T \sum_{n=1}^N \mathbf{1}_{A_n} \|x_n\| \right\|_{L^{p_2}(S_2)} \\ &\leq \|T\| \left\| \sum_{n=1}^N \mathbf{1}_{A_n} \|x_n\| \right\|_{L^{p_1}(S_1)} \\ &= \|T\| \left\| \sum_{n=1}^N \mathbf{1}_{A_n} \otimes x_n \right\|_{L^{p_1}(S_1; X)}, \end{aligned}$$

with obvious modifications for $p_2 = \infty$. Since the μ_1 -simple functions are dense in $L^{p_1}(S_1; X)$, this proves that $T \otimes I$ has a unique extension to a bounded

operator from $L^{p_1}(S_1; X)$ to $L^{p_2}(S_2; X)$ of norm $\|T \otimes I_X\| \leq \|T\|$. Equality $\|T \otimes I_X\| = \|T\|$ is obtained by considering functions f of the form $g \otimes x$ with $g \in L^{p_1}(S_1)$ and $x \in X$ of norm one. $\square$

In order to discuss possible analogues of this theorem for $p_1 = \infty$ we introduce the following terminology. A function $f : S \rightarrow X$ will be called a *countably-valued μ -simple function* if f can be written in the form $f = \sum_{n \geq 1} \mathbf{1}_{A_n} x_n$ with the sets $A_n \in \mathcal{A}$ disjoint and of finite μ -measure.

Lemma 2.1.4. *The bounded countably-valued μ -simple functions are dense in $L^\infty(S; X)$.*

Proof. Fix an arbitrary $f \in L^\infty(S; X)$. Since f is essentially separably valued, we may assume that X is separable. Also, by Proposition 1.1.15, we may assume that μ is σ -finite. Let $(A^{(n)})_{n \geq 1}$ be an exhaustion by disjoint sets of finite μ -measure.

Fix $\varepsilon > 0$ and let $(x_n)_{n \geq 1}$ be a dense sequence in X ; we may assume without loss of generality that $x_j \neq x_k$ if $j \neq k$. For each $s \in S$ let $n(s)$ be the first index $n \geq 1$ such that $\|f(s) - x_n\| < \varepsilon$. For $N \geq 1$ put $A_N := \{n(s) = N\}$. These sets are disjoint, and upon picking a strongly measurable representative of f , from

$$A_N = \{\|f - x_N\| < \varepsilon\} \cap \left(\bigcap_{n=1}^{N-1} \{\|f - x_n\| \geq \varepsilon\} \right) \in \mathcal{A}$$

we see that $g(s) = \sum_{n, N \geq 1} \mathbf{1}_{A_N \cap A^{(n)}} \otimes x_N$ is a countably-valued μ -simple function. Clearly, g is bounded and satisfies $\|f - g\|_\infty \leq \varepsilon$. $\square$

Suppose now that $T : L^\infty(S_1) \rightarrow L^p(S_2)$ is a bounded linear operator, where $1 \leq p \leq \infty$, and let $f \in L^\infty(S_1; X)$ be a countably-valued μ -simple function, say $f = \sum_{n \geq 1} \mathbf{1}_{A_n} \otimes x_n$ with the sets $A_n \in \mathcal{A}_1$ disjoint and of finite measure. By a slight abuse of notation one might try to define

$$(T \otimes I_X)f := \sum_{n \geq 1} T \mathbf{1}_{A_n} \otimes x_n.$$

However, as the next example shows, this is not well defined in general.

Example 2.1.5. Let $S_1 = \mathbb{N}$, $S_2 = \{0\}$, and $X = \mathbb{K}$. Identifying $L^\infty(S_1)$ and $L^p(\{0\})$ with ℓ^∞ and $\mathbb{K}$, respectively, consider the bounded positive operator $T : \ell^\infty \rightarrow \mathbb{K}$, $Tx := \text{Lim } x$, where Lim is a Banach limit. The constant function 1 can be represented as a pointwise convergent sum $\sum_{n \in \mathbb{N}} \mathbf{1}_{\{n\}}$, but

$$T\mathbf{1} = 1 \neq 0 = \sum_{n \in \mathbb{N}} T \mathbf{1}_{\{n\}}.$$

Lemma 2.1.6. *Let $(g_n)_{n \geq 1}$ be a sequence of non-negative functions in $L^\infty(S)$ such that $\sum_{n \geq 1} g_n$ converges almost everywhere to a function $G \in L^\infty(S)$. Let $(x_n)_{n \geq 1}$ be a bounded sequence in X . Then $\sum_{n \geq 1} g_n \otimes x_n$ converges almost everywhere to a function $F \in L^\infty(S; X)$ and we have $F = \sum_{n \geq 1} g_n \otimes x_n$ with convergence in the $\tau(L^\infty(S; X), L^1(S; X^*))$ -topology.*

Proof. Let $M_1, M_2 \geq 0$ be such that $\|x_n\| \leq M_1$ and $0 \leq G \leq M_2$ almost everywhere. Then for almost all $s \in S$,

$$\sum_{n \geq 1} \|g_n(s)x_n\| \leq M_1 \sum_{n \geq 1} g_n(s) \leq M_1 M_2.$$

Therefore, for almost all $s \in S$, the sum $F(s) := \sum_{n \geq 1} g_n(s)x_n$ converges in X and $\|F\|_{L^\infty(S; X)} \leq M_1 M_2$.

To prove the second assertion let $h \in L^1(S; X^*)$. Then

$$\begin{aligned} \limsup_{k \rightarrow \infty} \left| \left\langle \sum_{n=k+1}^{\infty} g_n \otimes x_n, h \right\rangle \right| &\leq \limsup_{k \rightarrow \infty} \int_S \sum_{n=k+1}^{\infty} g_n(s) \|h(s)\| \, d\mu(s) \|x_n\| \\ &\leq M_1 \limsup_{k \rightarrow \infty} \int_S \sum_{n=k+1}^{\infty} g_n(s) \|h(s)\| \, d\mu(s) = 0 \end{aligned}$$

by the dominated convergence theorem. $\square$

Theorem 2.1.7. *Let $(S_1, \mathcal{A}_1, \mu_1)$ and $(S_2, \mathcal{A}_2, \mu_2)$ be σ -finite measure spaces, let $1 \leq p, q \leq \infty$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$, and let $T : L^\infty(S_1) \rightarrow L^q(S_2)$ be a bounded operator which is the adjoint of a bounded operator from $L^p(S_2)$ to $L^1(S_1)$ and which is assumed to be positive if $1 < p \leq \infty$. Then the extension $T \otimes I_X$ to countably-valued μ -simple functions is well defined and extends uniquely to a bounded operator from $L^\infty(S_1; X)$ to $L^q(S_2; X)$ and we have*

$$\|T \otimes I_X\| = \|T\|.$$

Proof. By assumption, $T = V^*$ for some bounded operator V from $L^p(S_2)$ to $L^1(S_1)$. Clearly, V is positive if T is, and therefore by Theorem 2.1.3 (if $1 < p \leq \infty$) and Proposition 2.1.1 (if $p = 1$) the operator $V \otimes I_{X^*}$ has a unique extension to a bounded operator from $L^p(S_2; X^*)$ to $L^1(S_1; X^*)$. For $k = 1, 2$, the spaces $L^\infty(S_k; X)$ are isometrically contained as closed subspaces in $(L^1(S_k; X^*))^*$ in a natural way. Under these identifications, we claim that the adjoint operator $U := (V \otimes I_{X^*})^*$ maps $L^\infty(S_1; X)$ into $L^q(S_2; X)$. This restricted operator extends $T \otimes I_X$.

By Lemma 2.1.4 it suffices to show that for any countably-valued μ -simple function $f \in L^\infty(S_1; X)$ one has $Uf \in L^q(S_2; X)$. Fix such a function, say $f = \sum_{n \geq 1} \mathbf{1}_{A_n} x_n$ with $(A_n)_{n \geq 1}$ a sequence of disjoint sets in $\mathcal{A}$ of finite measure, and let $(x_n)_{n \geq 1}$ be a bounded sequence in X . Let $f_k = \sum_{n=1}^k \mathbf{1}_{A_n} \otimes x_n$ be the corresponding partial sums. Then $f_k \rightarrow f$ almost everywhere in X

and, by Lemma 2.1.6, in the weak* topology of $L^1(S_1; X^*)^*$. Since U , being an adjoint operator, is weak*-continuous, it follows that $Uf_k \rightarrow Uf$ in the weak*-topology of $L^p(S_1; X^*)^*$. On the other hand, $Uf_k \in L^q(S_2; X)$ and $Uf_k = \sum_{n=1}^k T(\mathbf{1}_{A_n}) \otimes x_n$. Note that $T(\mathbf{1}_{A_n}) \geq 0$ and, almost everywhere,

$$\sum_{n=1}^N T(\mathbf{1}_{A_n}) = T \sum_{n=1}^N \mathbf{1}_{A_n} \leq T(\mathbf{1}).$$

This implies that the sum $\sum_{n \geq 1} T(\mathbf{1}_{A_n})$ converges almost everywhere and its limit is in $L^q(S_2)$. Again from Lemma 2.1.6, we obtain the convergence $\lim_{k \rightarrow \infty} Uf_k = F$ both almost everywhere and in the weak*-topology of $L^p(S_1; X^*)^*$, where $F = \sum_{n \geq 1} T(\mathbf{1}_{A_n}) \otimes x_n$. Therefore, $Uf = F$ and the result follows. $\square$

2.1.b Boundedness of $T \otimes I_H$ for Hilbert spaces H

For Hilbert spaces H , no positivity requirement on T is needed for the L^p -boundedness of $T \otimes I_H$. Historically this result, proved in the 1930's by Paley, Marcinkiewicz, and Zygmund, provided one of the first instances of the use of Gaussian methods in functional analysis. While a systematic treatment of such techniques will be postponed to Volume II, we here provide an elementary introduction sufficient for the present needs.

Gaussian random variables

Let us first introduce the relevant terminology. A *random variable* is a measurable scalar-valued function defined on a probability space $(\Omega, \mathbb{P})$. The Lebesgue integral of an integrable random variable is called its *expectation* and is denoted by

$$\mathbb{E}\xi := \int_{\Omega} \xi \, d\mathbb{P}.$$

Random variables $\xi_1, \xi_2, \dots$ are said to be *independent* if for all $N \geq 1$ and all Borel subsets $B_1, \dots, B_N$ we have

$$\mathbb{P}\left(\bigcap_{n=1}^N \{\xi_n \in B_n\}\right) = \prod_{n=1}^N \mathbb{P}(\{\xi_n \in B_n\}).$$

If $(\xi_n)_{n=1}^N$ is a finite sequence of independent integrable random variables with values in $\mathbb{K}$, then their product $\prod_{n=1}^N \xi_n$ is integrable and we have $\mathbb{E} \prod_{n=1}^N \xi_n = \prod_{n=1}^N \mathbb{E}\xi_n$.

A random variable γ is called *centred Gaussian* with *variance* $\sigma^2 > 0$ if its distribution is given by

$$\mathbb{P}(\gamma \in B) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_B \exp(-t^2/2\sigma^2) \, dt$$

(for all Borel sets $B \subseteq \mathbb{R}$, when $\mathbb{K} = \mathbb{R}$), respectively

$$\mathbb{P}(\gamma \in B) = \frac{1}{\pi\sigma^2} \int_B \exp(-|z|^2/\sigma^2) dz$$

(for all Borel sets $B \subseteq \mathbb{C}$, when $\mathbb{K} = \mathbb{C}$). In both cases, γ is called *standard Gaussian* if $\sigma^2 = 1$.

It is easy to check that centred Gaussian random variables belong to $L^p(\Omega)$ for all $1 \leq p < \infty$; in the real case their L^p -norms can be calculated explicitly as

$$\|\gamma\|_p^p = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} |x|^p e^{-x^2/2\sigma^2} dx = \frac{2^{p/2}\sigma^p}{\sqrt{\pi}} \Gamma((p+1)/2), \quad (2.1)$$

where $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ is the Euler gamma function.

We say that $(\gamma_n)_{n \geq 1}$ is a *Gaussian sequence* if it consists of independent standard Gaussian random variables, defined on some probability space $(\Omega, \mathbb{P})$. It follows that

$$\mathbb{P}((\gamma_1, \dots, \gamma_N) \in B) = \frac{1}{(2\pi)^{N/2}} \int_B \exp(-|t|^2/2\sigma^2) dt \quad (2.2)$$

for all Borel set $B \subseteq \mathbb{R}^N$, when $\mathbb{K} = \mathbb{R}$. If $Q \in \mathcal{L}(\mathbb{R}^N)$ is an orthogonal transformation, then

$$\mathbb{P}(Q(\gamma_1, \dots, \gamma_N) \in B) = \frac{1}{(2\pi)^{N/2}} \int_{Q^*B} \exp(-|t|^2/2\sigma^2) dt,$$

and this agrees with the right hand side of (2.2) by a change of variables. This shows that $Q(\gamma_1, \dots, \gamma_N)$, having the same joint distribution as the original $(\gamma_1, \dots, \gamma_N)$ is another sequence of independent standard real Gaussian variables. In particular, for $a_n \in \mathbb{R}$ with $\sum_{n=1}^N a_n^2 = 1$, we find that $\sum_{n=1}^N a_n \gamma_n$ is another standard Gaussian variable.

Analogous results with a similar proofs are valid for complex Gaussian sequences and unitary transformations $U \in \mathcal{L}(\mathbb{C}^N)$ and numbers $a_n \in \mathbb{C}_N$ with $\sum_{n=1}^N |a_n|^2 = 1$.

A simple consequence of these invariance considerations is the following result mentioned after Proposition 2.1.2.

Proposition 2.1.8. *Let Ω be a probability space supporting a Gaussian sequence $(\gamma_n)_{n \geq 1}$, and let $p \in [1, \infty)$. Then $L^p(\Omega)$ has an isometric copy of ℓ^2 as a closed subspace, and in fact this subspace may be realised as the closure in $L^p(\Omega)$ of all finite sums $\sum_{n \geq 1} a_n \gamma_n$.*

Proof. By the results just discussed, $\|a\|_{\ell^2}^{-1} \sum_{n \geq 1} a_n \gamma_n$ is a standard Gaussian variable; hence

$$\left\| \sum_{n \geq 1} a_n \gamma_n \right\|_{L^p(\Omega)} = \|\gamma\|_p \|a\|_{\ell^2}. \quad (2.3)$$

As a consequence, the mapping

$$a \mapsto \frac{1}{\|\gamma\|_p} \sum_{n \geq 1} a_n \gamma_n$$

establishes an isometric embedding of ℓ^2 onto a closed subspace of $L^p(\Omega)$. $\square$

Extension theorem for Hilbert spaces

We are now prepared for the main result of this subsection:

Theorem 2.1.9 (Paley–Marcinkiewicz–Zygmund). *Let $(S_1, \mathcal{A}_1, \mu_1)$ and $(S_2, \mathcal{A}_2, \mu_2)$ be measure spaces, let $1 \leq p_1, p_2 < \infty$, and consider a bounded linear operator T from $L^{p_1}(S_1)$ to $L^{p_2}(S_2)$. Let H be a Hilbert space. Then $T \otimes I_H$ uniquely extends to a bounded operator from $L^{p_1}(S_1; H)$ to $L^{p_2}(S_2; H)$, and its norm satisfies*

$$\|T\| \leq \|T \otimes I_H\| \leq M_{p_1, p_2} \|T\|$$

with

$$M_{p_1, p_2} = \max \left\{ \frac{\|\gamma\|_{p_1}}{\|\gamma\|_{p_2}}, 1 \right\}.$$

Proof. We begin with a simple reduction. Since boundedness of an operator can be tested on sequences, and since each $f \in L^{p_1}(S_1) \otimes H$ takes its values in a finite-dimensional subspace of H , there is no loss of generality in assuming H to be separable. Under this assumption H is isometrically isomorphic to the Hilbert space $G := G^2(\Omega)$ constructed in Proposition 2.1.8 as the closure in $L^2(\Omega)$ of all random variables of the form $\sum_{n=1}^N a_n \gamma_n$, with $(\gamma_n)_{n \geq 1}$ a given sequence of independent standard Gaussian random variables.

It thus suffices to show that $T \otimes I_G$ is bounded with norm $\leq M_{p_1, p_2} \|T\|$. Let $f \in L^{p_1}(S_1; G)$ be arbitrary. Since $(\gamma_n)_{n \geq 1}$ is an orthonormal basis for G , we can write $f = \sum_{n \geq 1} \gamma_n f_n$ with $f_n = \mathbb{E}(\bar{\gamma}_n f)$ (in the real case, the conjugation is of course superfluous). It follows from (2.3) and Fubini's theorem (in the version of Corollary 1.2.23) that

$$\begin{aligned} \|(T \otimes I_G)f\|_{L^{p_2}(S_2; G)} &= \left\| \sum_{n \geq 1} \gamma_n T f_n \right\|_{L^{p_2}(S_2; G)} \\ &= \|\gamma\|_{p_2}^{-1} \left\| \sum_{n \geq 1} \gamma_n T f_n \right\|_{L^{p_2}(S_2; L^{p_2}(\Omega))} \\ &= \|\gamma\|_{p_2}^{-1} \left\| \sum_{n \geq 1} \gamma_n T f_n \right\|_{L^{p_2}(\Omega; L^{p_2}(S_2))} \\ &\leq \|\gamma\|_{p_2}^{-1} \|T\| \left\| \sum_{n \geq 1} \gamma_n f_n \right\|_{L^{p_2}(\Omega; L^{p_1}(S_1))}. \end{aligned}$$

Now we consider two cases.

If $p_2 < p_1$, then by Hölder's inequality we have $\mathbb{E}\xi^{p_2/p_1} \leq (\mathbb{E}\xi)^{p_2/p_1}$ and therefore

$$\begin{aligned} \left\| \sum_{n \geq 1} \gamma_n f_n \right\|_{L^{p_2}(\Omega; L^{p_1}(S_1))} &\leq \left\| \sum_{n \geq 1} \gamma_n f_n \right\|_{L^{p_1}(\Omega; L^{p_1}(S_1))} \\ &= \left\| \sum_{n \geq 1} \gamma_n f_n \right\|_{L^{p_1}(S_1; L^{p_1}(\Omega))} \\ &= \|\gamma\|_{p_1} \left\| \sum_{n \geq 1} \gamma_n f_n \right\|_{L^{p_1}(S_1; G)} = \|\gamma\|_{p_1} \|f\|_{L^{p_1}(S_1; G)}. \end{aligned}$$

This proves that $\|T \otimes I_G\| \leq \|\gamma\|_{p_1} \|\gamma\|_{p_2}^{-1} \|T\|$.

If $p_2 \geq p_1$, then we can apply Minkowski's inequality with exponent $p_2/p_1 \geq 1$ to obtain

$$\begin{aligned} \left\| \sum_{n \geq 1} \gamma_n f_n \right\|_{L^{p_2}(\Omega; L^{p_1}(S_1))} &\leq \left\| \sum_{n \geq 1} \gamma_n f_n \right\|_{L^{p_1}(S_1; L^{p_2}(\Omega))} \\ &= \|\gamma\|_{p_2} \left\| \sum_{n \geq 1} \gamma_n f_n \right\|_{L^{p_1}(S_1; G)} = \|\gamma\|_{p_2} \|f\|_{L^{p_1}(S_1; G)}. \end{aligned}$$

Therefore, $\|T \otimes I_G\| \leq \|T\|$. □

Extension results for complexifications of real Banach spaces

Incidentally, the proof technique just introduced also leads to a useful extension result in the context of *complexification* of a real Banach space X that we now discuss. We recall from Appendix B.4 that $X_{\mathbb{C}}$ is the space $X \times X$ equipped with the complex scalar multiplication

$$(a + ib)(x_1, x_2) := (ax_1 - bx_2, bx_1 + ax_2),$$

which makes it a complex vector space. Instead of (x_1, x_2) we shall write $x_1 + ix_2$. There are various norms on the complexification of a Banach space which turn it into a Banach space in such a way that if $T : X \rightarrow Y$ is a bounded operator, then its complexification $T_{\mathbb{C}}(x_1 + ix_2) := Tx_1 + iTx_2$ has the same norm as X . One simple choice of such a norm has been discussed in Appendix B.4. In the context of the L^p -extension problem it is natural to ask for a little more, viz. that if T is a bounded operator between two L^p spaces, then $(T \otimes I_X)_{\mathbb{C}}$, as a bounded operator the spaces $L^p(X_{\mathbb{C}})$, has the same norm as $T \otimes I_X$. The following lemma describes a norm which has exactly this property.

Lemma 2.1.10. *Let X be a real Banach space and $p \in [1, \infty)$. Equipped with the norm*

$$\|x_1 + ix_2\|_{X_{\mathbb{C}}^{\gamma, p}} := \frac{1}{\|\gamma_1\|_p} (\mathbb{E} \|\gamma_1 x_1 + \gamma_2 x_2\|_X^p)^{1/p},$$

where γ_1, γ_2 are independent real standard Gaussian random variables, $X_{\mathbb{C}}$ becomes a complex Banach space, which we denote by $X_{\mathbb{C}}^{\gamma, p}$.

Proof. All other properties are reasonably routine, except perhaps the compatibility of the norm with the scalar multiplication. Thus we need to check that

$$\|(a + ib)(x_1 + ix_2)\|_{X_{\mathbb{C}}^{\gamma,p}} = |a + ib| \|x_1 + ix_2\|_{X_{\mathbb{C}}^{\gamma,p}}.$$

We may assume that $|a + ib| = 1$. Then observe that

$$\gamma_1(ax_1 - bx_2) + \gamma_2(bx_1 + ax_2) = (a\gamma_1 + b\gamma_2)x_1 + (-b\gamma_1 + a\gamma_2)x_2.$$

Since $Q := \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ is an orthogonal matrix, it follows that $Q \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}$ is also a pair of independent standard Gaussian random variables, and thus

$$\mathbb{E}\|(a\gamma_1 + b\gamma_2)x_1 + (-b\gamma_1 + a\gamma_2)x_2\|_X^p = \mathbb{E}\|\gamma_1x_1 + \gamma_2x_2\|_X^p.$$

This gives the asserted norm identity. $\square$

Remark 2.1.11. For $X = \mathbb{R}$, we have $\|x_1 + ix_2\|_{\mathbb{R}_{\mathbb{C}}^{\gamma,p}} = (x_1^2 + x_2^2)^{1/2}$ for all $p \in [1, \infty)$. Indeed, if the right side is one, then $x_1\gamma_1 + x_2\gamma_2$ is another standard Gaussian random variable so that $(\mathbb{E}|\gamma_1x_1 + \gamma_2x_2|^p)^{1/p} = \|\gamma_1\|_p$, and the general case follows from this by scaling. Thus $\mathbb{R}_{\mathbb{C}}^{\gamma,p}$ coincides with $\mathbb{C}$ with its usual norm.

Proposition 2.1.12. *Let X and Y be real Banach spaces, let $p \in [1, \infty)$, and let $(S_1, \mathcal{A}_1, \mu_1)$ and $(S_2, \mathcal{A}_2, \mu_2)$ be σ -finite measure spaces. Given $T \in \mathcal{L}(L^p(S_1; X), L^p(S_2; Y))$, define the operator $T_{\mathbb{C}}$ by*

$$T_{\mathbb{C}}(f + ig) := Tf + iTg \quad \forall (f + ig) \in L^p(S_1; X_{\mathbb{C}}^{\gamma,p}).$$

Then $T_{\mathbb{C}} \in \mathcal{L}(L^p(S_1; X_{\mathbb{C}}^{\gamma,p}), L^p(S_2; Y_{\mathbb{C}}^{\gamma,p}))$ and

$$\|T_{\mathbb{C}}\|_{\mathcal{L}(L^p(S_1; X_{\mathbb{C}}^{\gamma,p}), L^p(S_2; Y_{\mathbb{C}}^{\gamma,p}))} = \|T\|_{\mathcal{L}(L^p(S_1; X), L^p(S_2; Y))}.$$

Proof. Observe that $Jf := f + i0$ identifies $L^p(S_1; X)$ isometrically as a (real-linear) subspace of $L^p(S_1; X_{\mathbb{C}}^{\gamma,p})$. Since $T_{\mathbb{C}}J = JT$, it is immediate that $\|T_{\mathbb{C}}\| \geq \|T\|$. The converse direction is a consequence of the following computation using only definitions and Fubini's theorem. Let Ω be the probability space supporting the random variables γ_1 and γ_2 . Then

$$\begin{aligned} \|\gamma\|_p \|T_{\mathbb{C}}(f + ig)\|_{L^p(S_2; Y_{\mathbb{C}}^{\gamma,p})} &= \|\gamma_1 Tf + \gamma_2 Tg\|_{L^p(S_2; L^p(\Omega; Y))} \\ &= \|T(\gamma_1 f + \gamma_2 g)\|_{L^p(\Omega; L^p(S_2; Y))} \\ &\leq \|T\| \|\gamma_1 f + \gamma_2 g\|_{L^p(\Omega; L^p(S_1; X))} \\ &= \|T\| \|\gamma_1 f + \gamma_2 g\|_{L^p(S_1; L^p(\Omega; X))} \\ &= \|\gamma\|_p \|T\| \|f + ig\|_{L^p(S_1; X_{\mathbb{C}}^{\gamma,p})}. \end{aligned}$$

$\square$

Corollary 2.1.13. *Let X be a real Banach space, let $p \in [1, \infty)$, and let $(S_1, \mathcal{A}_1, \mu_1)$ and $(S_2, \mathcal{A}_2, \mu_2)$ be σ -finite measure spaces. Suppose the operator $T \in \mathcal{L}(L^p(S_1; \mathbb{R}), L^p(S_2; \mathbb{R}))$ has a bounded extension $T \otimes I_X \in \mathcal{L}(L^p(S_1; X), L^p(S_2; X))$. Then $T_{\mathbb{C}} \in \mathcal{L}(L^p(S_1; \mathbb{C}), L^p(S_2; \mathbb{C}))$ has a bounded extension*

$$T_{\mathbb{C}} \otimes I_{X_{\mathbb{C}}^{\gamma, p}} \in \mathcal{L}(L^p(S_1; X_{\mathbb{C}}^{\gamma, p}), L^p(S_2; X_{\mathbb{C}}^{\gamma, p}))$$

of the same norm as $T \otimes I_X$.

Here $T_{\mathbb{C}}$ is the complex extension of T given by Proposition 2.1.12 in the case that $X = Y = \mathbb{R}$.

Proof. Recall that, as sets, $X_{\mathbb{C}}^{\gamma, p} = X_{\mathbb{C}}$. It is a routine exercise to verify that

$$L^p(S_1; \mathbb{C}) \otimes X_{\mathbb{C}} = (L^p(S_1; \mathbb{R}) \otimes X)_{\mathbb{C}},$$

where $Y_{\mathbb{C}}$ is the complexification of the real vector space Y for various choices of Y above; and, moreover, that

$$T_{\mathbb{C}} \otimes I_{X_{\mathbb{C}}} = (T \otimes I_X)_{\mathbb{C}}$$

as operators on the algebraic tensor product above. But the latter operator has a bounded extension in $\mathcal{L}(L^p(S_1; X_{\mathbb{C}}^{\gamma, p}), L^p(S_2; X_{\mathbb{C}}^{\gamma, p}))$ by Proposition 2.1.12 (applied to $T \otimes I_X$ in place of T), of the same norm as $T \otimes I_X \in \mathcal{L}(L^p(S_1; X_{\mathbb{C}}^{\gamma, p}), L^p(S_2; X_{\mathbb{C}}^{\gamma, p}))$, and hence, by the identity, so has $T_{\mathbb{C}} \otimes I_{X_{\mathbb{C}}}$. $\square$

Remark 2.1.14. Both Proposition 2.1.12 and Corollary 2.1.13 remain valid with each $L^p(S_1; Z)$, $Z \in \{\mathbb{R}, \mathbb{C}, X, X_{\mathbb{C}}^{\gamma, p}\}$ replaced by its subspace

$$L_0^p(S_1; Z) := \left\{ f \in L^p(S_1; Z) : \int_{S_1} f \, d\mu_1 = 0 \right\}$$

in the case that $(S_1, \mathcal{A}_1, \mu_1)$ is a finite measure space. It is easily verified that the same proofs apply in this situation, noting that a product space-valued function has mean zero if and only if both its component have.

2.1.c Counterexamples

We now turn to some basic explicit examples of operators whose vector-valued extensions may fail to be bounded. The first such example goes back to the very beginning days of the theory of the Bochner integral and was published in 1933 by Bochner, in the same year in which he introduced his integral.

Example 2.1.15 (Fourier transform). The Fourier transform

$$\mathcal{F}f(\xi) := \int_{\mathbb{R}} f(x) e^{-2\pi i x \cdot \xi} \, dx, \quad f \in L^1(\mathbb{R}),$$

is bounded from $L^1(\mathbb{R})$ to $L^\infty(\mathbb{R})$. Also, by the Plancherel theorem (Theorem 2.4.9), its restriction to $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ extends to an isometry on $L^2(\mathbb{R})$ (the so-called *Fourier–Plancherel transform*). Hence, by the Riesz–Thorin theorem (Theorem 2.2.1 below), it interpolates to a bounded operator of norm ≤ 1 from $L^p(\mathbb{R})$ to $L^{p'}(\mathbb{R})$, for all $p \in [1, 2]$ and $\frac{1}{p} + \frac{1}{p'} = 1$; this result is known as the *Hausdorff–Young theorem*. By an easy dilation argument one furthermore shows that if $\mathcal{F}$ extends to a bounded operator from $L^p(\mathbb{R})$ to $L^q(\mathbb{R})$, then $q = p'$.

Trivially, the Fourier transform extends to a bounded operator $\mathcal{F} \otimes I_X : L^1(\mathbb{R}; X) \rightarrow L^\infty(\mathbb{R}; X)$. Since the X -valued step functions with values are dense in $L^1(\mathbb{R}; X)$ and the Fourier transform of such functions belong to $C_0(\mathbb{R}; X)$, this extension actually maps $L^1(\mathbb{R}; X)$ into $C_0(\mathbb{R}; X)$ (see Lemma 2.4.3). We will show here that the Fourier transform does not extend to a bounded operator from $L^p(\mathbb{R}; \ell^r)$ to $L^{p'}(\mathbb{R}; \ell^r)$ whenever $1 \leq r < p \leq 2$.

Let $f \in C_c(\mathbb{R})$ be a non-zero function with support in the interval $(0, 1)$. For $N \geq 1$ define the functions $f_N \in C_c(\mathbb{R}; \ell^r)$ by

$$f_N := \sum_{n=0}^N f(\cdot + n) \otimes e_{n+1}.$$

Then, by the disjointness of the supports of the functions $f(\cdot + n)$,

$$\begin{aligned} \|f_N\|_{L^p(\mathbb{R}; \ell^r)} &= \left(\int_{-\infty}^{\infty} \left(\sum_{n=1}^N |f(x+n)|^r \right)^{p/r} dx \right)^{1/p} \\ &= \left(\int_{-\infty}^{\infty} \sum_{n=1}^N |f(x+n)|^p dx \right)^{1/p} = N^{1/p} \|f\|_{L^p(\mathbb{R})}. \end{aligned}$$

On the other hand, $(\mathcal{F} \otimes I_{\ell^r})f_N(\xi) = \sum_{n=0}^N e^{2\pi i n \xi} \mathcal{F}f(\xi) \otimes e_{n+1}$ and therefore

$$\begin{aligned} \|(\mathcal{F} \otimes I_{\ell^r})f_N\|_{L^{p'}(\mathbb{R}; \ell^r)} &= \left(\int_{-\infty}^{\infty} \left(\sum_{n=1}^N |e^{2\pi i n \xi} \mathcal{F}f(\xi)|^r \right)^{p'/r} d\xi \right)^{1/p'} \\ &= N^{1/r} \left(\int_{-\infty}^{\infty} |\mathcal{F}f(\xi)|^{p'} d\xi \right)^{1/p'} = N^{1/r} \|\mathcal{F}f\|_{L^{p'}(\mathbb{R})}, \end{aligned}$$

with obvious modifications if $p' = \infty$. Hence,

$$\|(\mathcal{F} \otimes I_{\ell^r})\| \geq N^{\frac{1}{r} - \frac{1}{p}} \frac{\|\mathcal{F}f\|_{L^{p'}(\mathbb{R})}}{\|f\|_{L^p(\mathbb{R})}}.$$

Since we are assuming that $r < p$, this shows that $\mathcal{F} \otimes I_{\ell^r}$ does not extend to a bounded operator from $L^p(\mathbb{R}; \ell^r)$ to $L^{p'}(\mathbb{R}; \ell^r)$.

For $1 \leq p \leq 2$ and $p' < r \leq \infty$, the Fourier transform fails to extend to a bounded operator from $L^p(\mathbb{R}; \ell^r)$ to $L^{p'}(\mathbb{R}; \ell^r)$ as well. The proof runs along similar lines, this time using the functions $f_N(x) = \sum_{n=0}^N e^{-2\pi i k x} f(x) \otimes e_{n+1}$.

In the remaining parameter range $1 \leq p \leq 2$ and $p \leq r \leq p'$ the Fourier transform *does* extend to a bounded operator from $L^p(\mathbb{R}; \ell^r)$ to $L^{p'}(\mathbb{R}; \ell^r)$. In this range the space $L^p(\mathbb{R}; \ell^r)$ may be identified with the complex interpolation space between $L^1(\mathbb{R}; \ell^{r_0})$ and $L^2(\mathbb{R}; \ell^2)$ for a suitable choice of exponent $1 \leq r_0 \leq \infty$ and therefore the result follows by complex interpolation (see Section 2.2). This state of affairs is often abbreviated by saying that ℓ^r has *Fourier type* p . We will take up this issue in more detail in Section 2.4.

A similar result holds for the Fourier transform on the circle; see Example 2.7.4 in the Notes at the end of the chapter.

Example 2.1.16 (Hilbert transform). It is a classical result of M. Riesz that the Hilbert transform, given by the principle value integral

$$Hf(x) = PV \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(y)}{x-y} dy, \quad x \in \mathbb{R},$$

defines a bounded operator on $L^p(\mathbb{R})$ for all $1 < p < \infty$. The importance of this operator comes from the following identity, valid for $f \in L^p(\mathbb{R}) \cap L^2(\mathbb{R})$:

$$\mathcal{F}(Hf)(\xi) = -i \operatorname{sgn}(\xi) \mathcal{F}f(\xi),$$

where $\mathcal{F}$ is the Fourier-Plancherel transform. Thus, H is the L^p -Fourier multiplier corresponding to the multiplier function $m(\xi) = -i \operatorname{sgn}(\xi)$. A detailed discussion of the Hilbert transform and Fourier multipliers will be undertaken in Chapter 5. Presently, we will check that H does not extend to a bounded operator on $L^p(\mathbb{R}; \ell^1)$ for any $1 \leq p < \infty$.

Let $1 \leq p < \infty$ be fixed and let $f \in L^p(\mathbb{R}; \ell^1)$ be given by $f_N = \sum_{n=1}^N \mathbf{1}_{(n-1, n)} \otimes e_n$, with e_n the n th unit vector in ℓ^1 . Clearly

$$\|f_N\|_{L^p(\mathbb{R}; \ell^1)} = N^{1/p}.$$

On the other hand, from the straightforward identity

$$H\mathbf{1}_{(a,b)}(x) = \frac{1}{\pi} \log \left| \frac{x-a}{x-b} \right| = \frac{1}{\pi} \log \left| 1 + \frac{b-a}{x-b} \right|, \quad x \in \mathbb{R} \setminus \{a, b\},$$

we obtain

$$(H \otimes I_{\ell^1})f_N(x) = \frac{1}{\pi} \sum_{n=1}^N \log \left| 1 + \frac{1}{x-n} \right| e_n, \quad x \in \mathbb{R} \setminus \{1, \dots, N\},$$

so that

$$\begin{aligned}
\|(H \otimes I_{\ell^1})f\|_{L^p(\mathbb{R};\ell^1)}^p &= \frac{1}{\pi^p} \int_{\mathbb{R}} \left(\sum_{n=1}^N \left| \log \left| 1 + \frac{1}{x-n} \right| \right| \right)^p dx \\
&\geq \frac{1}{\pi^p} \sum_{k=1}^N \int_k^{k+1} \left(\sum_{n=1}^k \log \left(1 + \frac{1}{x-n} \right) \right)^p dx.
\end{aligned}$$

For $x \in (k, k+1)$,

$$\begin{aligned}
\sum_{n=1}^k \log \left(1 + \frac{1}{x-n} \right) &\geq \sum_{n=1}^k \log \left(1 + \frac{1}{k-n+1} \right) \\
&\geq \frac{1}{2} \sum_{n=1}^k \frac{1}{k-n+1} \geq \frac{1}{2} \log(k),
\end{aligned}$$

where we used that $\log(1+y) \geq \frac{1}{2}y$ for $y \in [0, 1]$. Therefore

$$\begin{aligned}
\|(H \otimes I_{\ell^1})f\|_{L^p(\mathbb{R};\ell^1)} &\geq \frac{1}{2\pi} \left(\sum_{k=1}^N \log^p(k) \right)^{1/p} \\
&\geq \frac{1}{2\pi} \left(\sum_{k=\lceil N/2 \rceil}^N \log^p(k) \right)^{1/p} \geq \frac{1}{2\pi} (N/2)^{1/p} \log(N/2).
\end{aligned}$$

This implies that $\|H \otimes I_{\ell^1}\| \geq (2\pi)^{-1} 2^{-1/p} \log(N/2)$, and this lower bound tends to ∞ as $N \rightarrow \infty$.

Analogous non-extension results hold for the discrete versions of the Fourier transform and Hilbert transform; as in the continuous case these operators do not extend to bounded operators from $L^p(\mathbb{T}; \ell^r)$ to $L^{p'}(\mathbb{Z}; \ell^r)$, $r \in [1, p) \cup (p', \infty]$, and from $L^p(\mathbb{T}; \ell^1)$ to itself, respectively. The discrete Hilbert transform, which we shall denote by $\tilde{H}$, is defined as the L^p -Fourier multiplier corresponding to $m(n) = -i \operatorname{sgn}(n)$ (by convention we take $\operatorname{sgn}(0) = 0$) and is bounded on $L^p(\mathbb{T})$ for $1 < p < \infty$. The related multiplier

$$\tilde{m}(n) = \begin{cases} 0, & n < 0, \\ 1, & n \geq 0, \end{cases}$$

defines the Riesz projection

$$R : \sum_{n \in \mathbb{Z}} c_n e^{ink} \mapsto \sum_{n \geq 0} c_n e^{ink}.$$

Clearly, $R = \frac{1}{2}(i\tilde{H} + I + P)$, where P is the projection onto the constants. It follows that also the Riesz projection cannot be extended to $L^p(\mathbb{T}; \ell^1)$. This topic will be revisited in Chapter 5, where we will prove the boundedness of this projection in $L^p(\mathbb{T}; X)$ for UMD spaces X .

Example 2.1.17 (Wiener–Itô isometry). Let $(B(t))_{t \geq 0}$ be a standard Brownian motion on a probability space $(\Omega, \mathbb{P})$. The *Wiener–Itô isometry* is the isometry $W : L^2(\mathbb{R}_+) \rightarrow L^2(\Omega)$ which, on the dense subspace of step functions in $L^2(\mathbb{R}_+)$, is given by the stochastic integral

$$W\mathbf{1}_{(a,b)} := \int_0^\infty \mathbf{1}_{(a,b)} dB := B(b) - B(a), \quad 0 \leq a < b < \infty.$$

Let us show that for all $1 \leq p < 2$, $W \otimes I_{\ell^p}$ cannot be extended to a bounded operator from $L^2(\mathbb{R}_+; \ell^p)$ to $L^2(\Omega; \ell^p)$. Put $r_0 := 0$, $r_n := \sum_{k=1}^n k^{-\frac{1}{2} - \frac{1}{p}}$ for $n \geq 1$, and $r := \lim_{n \rightarrow \infty} r_n$. Let $I_n := [r_{n-1}, r_n)$. The union of these intervals is $[0, r)$. Define $f : \mathbb{R}_+ \rightarrow \ell^p$ by

$$f(t) := e_n, \quad t \in I_n, \quad n \geq 1,$$

with e_n the n th unit vector of ℓ^p , and set $f(t) = 0$ for $t \geq r$. Then $f \in L^\infty(0, r; \ell^p)$. Hence the step functions $f_n := \mathbf{1}_{[0, r_n)} f$ are uniformly bounded in $L^2(\mathbb{R}_+; \ell^p)$ and we have

$$Wf_n = \sum_{k=1}^n (B(r_k) - B(r_{k-1}))e_k.$$

Using (2.1) and Hölder's inequality, we obtain the inequality

$$\begin{aligned} \|Wf_n\|_{L^2(\Omega; \ell^p)}^p &\geq \|Wf_n\|_{L^p(\Omega; \ell^p)}^p = \mathbb{E} \sum_{k=1}^n |B(r_k) - B(r_{k-1})|^p \\ &\approx \sum_{k=1}^n (k^{-\frac{1}{2} - \frac{1}{p}})^{p/2} = \sum_{k=1}^n k^{-\frac{p}{4} - \frac{1}{2}} \end{aligned}$$

with constants only depending on p . Clearly the right-hand side diverges as $n \rightarrow \infty$.

The problem of determining the class of Banach spaces X for which the Fourier–Plancherel transform, the Hilbert transform, and the Wiener–Itô isometry do have bounded extensions has motivated a large body of profound investigations. As a preview to some of the main threads in this book we present the concise answers for these three operators.

Theorem 2.1.18 (Kwapień). *For a Banach space X the following assertions are equivalent:*

- (1) *the Fourier–Plancherel transform extends to a bounded operator on $L^2(\mathbb{R}; X)$;*
- (2) *X is isomorphic to a Hilbert space.*

Theorem 2.1.19 (Burkholder–Bourgain). *For a Banach space X the following assertions are equivalent:*

- (1) the Hilbert transform extends to a bounded operator on $L^p(\mathbb{R}; X)$ for some (equivalently, for all) $1 < p < \infty$;
- (2) X is a UMD space.

In contrast to these two results, whose proofs require considerable effort, the next one is fairly elementary.

Theorem 2.1.20. *For a Banach space X the following assertions are equivalent:*

- (1) the Wiener–Itô isometry extends to a bounded operator from $L^2(\mathbb{R}_+; X)$ to $L^2(\Omega; X)$;
- (2) X has type 2.

The class of Banach spaces with the UMD property spaces will be studied in Chapter 4. It contains the Hilbert spaces and the L^p -spaces with $1 < p < \infty$. The notion of type will be discussed in Volume II (see also Subsection 4.3.b). The spaces c_0 , ℓ^1 , and all Banach spaces containing these as closed subspaces, fail the UMD property and the type 2 property (and in fact have no non-trivial type).

In Volume II we shall see that bounded operators between L^2 -spaces always do have a certain ‘Gaussian’ X -valued extension.

2.2 Interpolation of Bochner spaces

On various occasions we will need interpolation results for the spaces $L^p(S; X)$ and operators defined on them. These will be collected in the present section. For generalities on interpolation of Banach spaces and operators defined thereon the reader is referred to Appendix C, whose terminology and results we will use freely here.

Let $(S, \mathcal{A}, \mu)$ be a measure space. The vector space of all equivalence classes of strongly measurable functions from S into a Banach space X , identifying functions which are equal almost everywhere, is denoted by $L^0(S; X)$. Extending Proposition A.2.4 to the vector-valued setting in the obvious way, if $(S, \mathcal{A}, \mu)$ is σ -finite, then $L^0(S; X)$ is a complete metric space.

2.2.a The Riesz–Thorin interpolation theorem

Perhaps the oldest theorem on interpolation of bounded operators defined on different L^p -spaces is due to M. Riesz; its modern formulation is due to Thorin.

Theorem 2.2.1 (Riesz–Thorin). *Let X and Y be complex Banach spaces, let $1 \leq p_0, p_1, q_0, q_1 \leq \infty$ and let $(S_0, \mathcal{A}_0, \mu_0)$ and $(S_1, \mathcal{A}_1, \mu_1)$ be measure spaces. Let $T : L^{p_0}(S_0; X) + L^{p_1}(S_0; X) \rightarrow L^0(S_1; Y)$ be a linear operator which maps $L^{p_0}(S_0; X)$ into $L^{q_0}(S_1; X)$ and $L^{p_1}(S_0; X)$ into $L^{q_1}(S_1; X)$. If*

$$\|Tf\|_{L^{q_j}(S_1;Y)} \leq A_j \|f\|_{L^{p_j}(S_0;X)} \quad \forall f \in L^{p_j}(S_0;X) \quad (j = 0, 1),$$

then for all $0 < \theta < 1$ the operator T maps $L^{p_\theta}(S_0;X)$ into $L^{q_\theta}(S_1;Y)$ and

$$\|Tf\|_{L^{q_\theta}(S_1;Y)} \leq A_0^{1-\theta} A_1^\theta \|f\|_{L^{p_\theta}(S_0;X)} \quad \forall f \in L^{p_\theta}(S_0;X),$$

where

$$\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad \frac{1}{q_\theta} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

The proof is based on the three lines lemma. Some background information about vector-valued holomorphic functions can be found in Section B.3.

Consider the open strip

$$\mathbb{S} := \{z \in \mathbb{C} : 0 < \Re z < 1\}.$$

Lemma 2.2.2 (Three lines lemma). *Suppose that $F : \overline{\mathbb{S}} \rightarrow X$ is a continuous function, holomorphic on $\mathbb{S}$, and that*

$$\sup_{v \in \mathbb{R}} \|F(iv)\| \leq A_0, \quad \sup_{v \in \mathbb{R}} \|F(1+iv)\| \leq A_1.$$

Then F is uniformly bounded on $\overline{\mathbb{S}}$ and for all $0 < \theta < 1$ we have

$$\sup_{v \in \mathbb{R}} \|F(\theta + iv)\| \leq A_0^{1-\theta} A_1^\theta.$$

Proof. For each $\varepsilon > 0$ the function $F_\varepsilon(z) = A_0^{z-1} A_1^{-z} \exp(\varepsilon z(z-1)) F(z)$ satisfies the assumptions of the lemma with $A_0 = A_1 = 1$. Moreover, $\lim_{v \rightarrow \infty} \|F_\varepsilon(u+iv)\| = 0$ uniformly with respect to $u \in [0, 1]$. Hence for large enough R we have $\|F_\varepsilon\| \leq 1$ on the boundary of the rectangle $\Re z \in [0, 1]$, $|\Im z| \leq R$. The maximum modulus principle implies that $\|F_\varepsilon\| \leq 1$ on this rectangle, and by letting $R \rightarrow \infty$ we find that $\|F_\varepsilon\| \leq 1$ on $\mathbb{S}$. The lemma now follows by letting $\varepsilon \downarrow 0$. $\square$

Proof of Theorem 2.2.1. If $p_0 = p_1 = \infty$, then for all $f \in L^\infty(S_0;X)$ we have $Tf \in L^{q_0}(S_1;Y) \cap L^{q_1}(S_1;Y)$ and therefore, by Hölder's inequality,

$$\|Tf\|_{q_\theta} \leq \|Tf\|_{q_0}^{1-\theta} \|Tf\|_{q_1}^\theta \leq A_0^{1-\theta} A_1^\theta \|f\|_\infty^{1-\theta} \|f\|_\infty^\theta = A_0^{1-\theta} A_1^\theta \|f\|_\infty.$$

This settles that case $p_0 = p_1 = \infty$. In the rest of the proof we may therefore assume that $\min\{p_0, p_1\} < \infty$. This assumption implies that $p_\theta < \infty$.

For $z \in \overline{\mathbb{S}}$ define $p_z, q_z \in \mathbb{C}$ by the relations

$$\frac{1}{p_z} = \frac{1-z}{p_0} + \frac{z}{p_1}, \quad \frac{1}{q_z} = \frac{1-z}{q'_0} + \frac{z}{q'_1},$$

where q'_0, q'_1, q'_z are the conjugate exponents of q_0, q_1, q_z . Let $a : S_0 \rightarrow X$ and $b : S_1 \rightarrow X^*$ be μ_0 - and μ_1 -simple functions and define, for each

$z \in \overline{\mathbb{S}}$, the μ_0 - and μ_1 -simple functions $f_z \in L^{p_0}(S_0; X) \cap L^{p_1}(S_0; X)$ and $g_z \in L^{q'_0}(S_1; Y^*) \cap L^{q'_1}(S_1; Y^*)$ by

$$\begin{aligned} f_z(s_0) &= \mathbf{1}_{\{a(s_0) \neq 0\}} \|a(s_0)\|^{p_\theta/p_z} \frac{a(s_0)}{\|a(s_0)\|}, \\ g_z(s_1) &= \mathbf{1}_{\{b(s_1) \neq 0\}} \|b(s_1)\|^{q'_\theta/q'_z} \frac{b(s_1)}{\|b(s_1)\|}. \end{aligned}$$

Then $Tf_z \in L^{q_0}(S_1; Y) \cap L^{q_1}(S_1; Y)$, and the function $F : \overline{\mathbb{S}} \rightarrow \mathbb{C}$ defined by

$$F(z) := \int_{S_1} \langle Tf_z, g_z \rangle d\mu_1$$

is continuous on $\overline{\mathbb{S}}$, holomorphic on $\mathbb{S}$, and for all $v \in \mathbb{R}$ we have

$$|F(iv)| \leq A_0 \|f_{iv}\|_{p_0} \|g_{iv}\|_{q'_0} \leq A_0 \|a\|_{p_\theta}^{p_\theta/p_0} \|b\|_{q'_\theta}^{q'_\theta/q'_0}$$

and similarly

$$|F(1 + iv)| \leq A_1 \|a\|_{p_\theta}^{p_\theta/p_1} \|b\|_{q'_\theta}^{q'_\theta/q'_1}.$$

Hence by the three lines lemma,

$$\begin{aligned} \left| \int_{S_1} \langle Ta, b \rangle d\mu_1 \right| &= |F(\theta)| \\ &\leq A_0^{1-\theta} A_1^\theta \|a\|_{p_\theta}^{(1-\theta)p_\theta/p_0} \|b\|_{q'_\theta}^{(1-\theta)q'_\theta/q'_0} \|a\|_{p_\theta}^{\theta p_\theta/p_1} \|b\|_{q'_\theta}^{\theta q'_\theta/q'_1} \\ &\leq A_0^{1-\theta} A_1^\theta \|a\|_{p_\theta} \|b\|_{q'_\theta}. \end{aligned}$$

Taking the supremum over all μ_1 -simple functions $b \in L^{q'_\theta}(S_1; Y^*)$, by Proposition 1.3.1 and Corollary 1.3.2 we obtain

$$\|Ta\|_{q_\theta} \leq A_0^{1-\theta} A_1^\theta \|a\|_{p_\theta}.$$

Since the μ_0 -simple functions are dense in $L^{p_\theta}(S_0; X)$ (here we use that $p_\theta < \infty$), this proves that the restriction of T to the μ_0 -simple functions has a unique extension to a bounded operator $\tilde{T}$ from $L^{p_\theta}(S_0; X)$ into $L^{q_\theta}(S_1; Y)$ of norm at most $A_0^{1-\theta} A_1^\theta$.

It remains to be checked that $\tilde{T}f = Tf$ for all $f \in L^{p_\theta}(S_0; X)$. To this end we may assume $p_0 \leq p_1$ and write $f = \mathbf{1}_{\{\|f\| > 1\}} f + \mathbf{1}_{\{\|f\| \leq 1\}} f =: f^0 + f^1$ and observe that $f^j \in L^{p_j}(S_0; X)$ ($j = 0, 1$). If $f_n \rightarrow f$ in $L^{p_\theta}(S_0; X)$ with each f_n μ_0 -simple, then, with obvious notations, $f_n^j \rightarrow f^j$ in $L^{p_j}(S_0; X)$ and therefore $\tilde{T}f^j = \lim_{n \rightarrow \infty} \tilde{T}f_n^j = \lim_{n \rightarrow \infty} Tf_n^j = Tf^j$ in $L^{q_j}(S_1; X)$. As a consequence $\tilde{T}f = \tilde{T}f^0 + \tilde{T}f^1 = Tf^0 + Tf^1 = Tf$. $\square$

2.2.b The Marcinkiewicz interpolation theorem

In certain applications the boundedness of operators on L^2 is readily proved (e.g., by Fourier methods). In order to obtain boundedness on L^p for $1 < p < 2$, a strategy could be to try to prove boundedness on L^1 and then to use the Riesz–Thorin interpolation theorem. It frequently happens, however, that L^2 -bounded operators fail to be L^1 -bounded; this is rather the rule than the exception. In such cases, the Marcinkiewicz interpolation theorem comes to rescue. It states that, rather than checking boundedness at the endpoint L^1 , it suffices to prove a weaker bound in terms of the $L^{1,\infty}$ -norm, often referred to as a *weak type* $(1,1)$ *bound*. At least as important is the fact that the Marcinkiewicz interpolation theorem also covers sub-linear operators, such as the Hardy–Littlewood maximal operator.

For a strongly μ -measurable function $f : S \rightarrow X$ and $p \in [1, \infty)$ consider the extended real number

$$\|f\|_{L^{p,\infty}(S;X)} := \sup_{r>0} r(\mu(\|f\| > r))^{1/p}.$$

The space $L^{p,\infty}(S;X)$ is defined as the subspace of functions for which $\|f\|_{L^{p,\infty}(S;X)}$ is finite. This space is sometimes referred to as *weak- L^p* ($S;X$). Note that $L^p(S;X) \subseteq L^{p,\infty}(S;X)$ and for all $f \in L^p(S;X)$,

$$\|f\|_{L^{p,\infty}(S;X)} \leq \|f\|_{L^p(S;X)}.$$

This is immediate from $r^p \mathbf{1}_{\{\|f\|>r\}} \leq \|f\|^p$.

Recall that $L^0(S;X)$ is the vector space of all (equivalence classes of) strongly μ -measurable functions $f : S \rightarrow X$. If $(S_0, \mathcal{A}_0, \mu_0)$ and $(S_1, \mathcal{A}_1, \mu_1)$ are measure spaces, a mapping T defined on a linear subspace E of $L^0(S_0;X)$ and taking values in $L^0(S_1;Y)$ is called *sub-linear* if for all $f \in E$ and $c \in \mathbb{K}$,

$$\|T(cf)\| = |c|\|T(f)\| \quad \text{almost everywhere,}$$

and for all $f, g \in E$,

$$\|T(f+g)\| \leq \|T(f)\| + \|T(g)\| \quad \text{almost everywhere.}$$

Theorem 2.2.3 (Marcinkiewicz). *Let X, Y be Banach spaces, let $0 < p_0 < p_1 \leq \infty$ and let $(S_0, \mathcal{A}_0, \mu_0)$ and $(S_1, \mathcal{A}_1, \mu_1)$ be measure spaces. Let $T : L^{p_0}(S_0;X) + L^{p_1}(S_0;X) \rightarrow L^0(S_1;Y)$ be a sub-linear operator that satisfies*

$$\|Tf\|_{L^{p_j,\infty}(S_1;Y)} \leq A_j \|f\|_{L^{p_j}(S_0;X)} \quad \forall f \in L^{p_j}(S_0;X) \quad (j = 0, 1),$$

where we interpret $L^{\infty,\infty} := L^\infty$. For all $\theta \in (0, 1)$, let $p_\theta \in (p_0, p_1)$ be defined by

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

Then we have the estimates

$$\begin{aligned} \|Tf\|_{L^{p_\theta}(S_1;Y)} &\leq c(\theta, p_0, p_1) \left(\frac{A_0}{1-\theta}\right)^{1-\theta} \left(\frac{A_1}{\theta}\right)^\theta \|f\|_{L^{p_\theta}(S_0;X)}, \\ \|Tf\|_{L^{p_\theta,\infty}(S_1;Y)} &\leq c(\theta, p_0, p_1) \left(\frac{A_0}{1-\theta}\right)^{1-\theta} \left(\frac{A_1}{\theta}\right)^\theta \|f\|_{L^{p_\theta,\infty}(S_0;X)}, \end{aligned} \quad (2.4)$$

where

$$c(\theta, p_0, p_1) = \left\{ \left(\frac{\Gamma(p_\theta - p_0 + 1)\Gamma(p_0 + 1)}{\Gamma(p_\theta)} \right)^{\frac{p_1 - p_\theta}{p_1 - p_0}} p_1^{\frac{p_\theta - p_0}{p_1 - p_0}} \frac{p_1 - p_0}{(p_1 - p_\theta)(p_\theta - p_0)} \right\}^{\frac{1}{p_\theta}}$$

and

$$c(\theta, p_0, \infty) = \left\{ \frac{\Gamma(p_\theta - p_0)\Gamma(p_0 + 1)}{\Gamma(p_\theta)} \right\}^{\frac{1}{p_\theta}}.$$

A case of primary interest for many applications is singled out in the following corollary:

Corollary 2.2.4 (Marcinkiewicz, the case $p_0 = 1, p_1 = \infty$). *Let X, Y be Banach spaces, let $(S_0, \mathcal{A}_0, \mu_0)$ and $(S_1, \mathcal{A}_1, \mu_1)$ be measure spaces and let $T : L^1(S_0; X) + L^\infty(S_0; X) \rightarrow L^0(S_1; Y)$ be a sub-linear operator that satisfies*

$$\|Tf\|_{L^{p,\infty}(S_1;Y)} \leq B_p \|f\|_{L^p(S_0;X)} \quad \forall f \in L^p(S_0;X) \quad (p = 1, \infty),$$

where we interpret $L^{\infty,\infty} := L^\infty$. For all $p \in (1, \infty)$, we then have the estimates

$$\begin{aligned} \|Tf\|_{L^p(S_1;Y)} &\leq p' B_1^{1/p} B_\infty^{1/p'} \|f\|_{L^p(S_0;X)}, \\ \|Tf\|_{L^{p,\infty}(S_1;Y)} &\leq p' B_1^{1/p} B_\infty^{1/p'} \|f\|_{L^{p,\infty}(S_0;X)}. \end{aligned}$$

The constants p' in these inequalities are sharp. This fact will be proved in the next chapter (see Proposition 3.2.4 and the observation preceding it).

Proof. This is the case $p_0 = 1, p_1 = \infty$ of Theorem 2.2.3. Thus $p = p_\theta$ with $1/p = (1 - \theta)/1 + \theta/\infty$, so that $\theta = 1/p'$. Then

$$\left(\frac{1}{1-\theta}\right)^{1-\theta} \left(\frac{1}{\theta}\right)^\theta = p^{1/p} (p')^{1/p'}$$

and

$$c(p, 1, \infty) = \left\{ \frac{\Gamma(p-1)\Gamma(2)}{\Gamma(p)} \right\}^{1/p} = \left(\frac{1}{p-1}\right)^{1/p}.$$

Hence

$$\begin{aligned} c(p, 1, \infty) \left(\frac{B_0}{1-\theta}\right)^{1-\theta} \left(\frac{B_1}{\theta}\right)^\theta &= \left(\frac{p}{p-1}\right)^{1/p} (p')^{1/p'} B_1^{1/p} B_\infty^{1/p'} \\ &= (p')^{1/p+1/p'} B_1^{1/p} B_\infty^{1/p'} = p' B_1^{1/p} B_\infty^{1/p'}, \end{aligned}$$

as claimed. $\square$

The proof of Theorem 2.2.3 rests on the following decomposition lemma:

Lemma 2.2.5. *For $f \in L^p(S; X)$ or $f \in L^{p,\infty}(S; X)$ and $\lambda > 0$, consider the splitting*

$$f = \tilde{f}^\lambda + \tilde{f}_\lambda,$$

where

$$\begin{aligned}\tilde{f}^\lambda &:= \left(f - \lambda \frac{f}{\|f\|}\right) \cdot \mathbf{1}_{\{\|f\| > \lambda\}}, \\ \tilde{f}_\lambda &:= f \cdot \mathbf{1}_{\{\|f\| \leq \lambda\}} + \lambda \frac{f}{\|f\|} \cdot \mathbf{1}_{\{\|f\| > \lambda\}}.\end{aligned}$$

For $0 < p_0 < p < p_1 < \infty$, these satisfy

$$\begin{aligned}\int_0^\infty p\lambda^{p-p_0-1} \|\tilde{f}^\lambda\|_{p_0}^{p_0} d\lambda &= \frac{\Gamma(p_0+1)\Gamma(p-p_0)}{\Gamma(p)} \|f\|_{L^p}^p, \\ \int_0^\infty p\lambda^{p-p_1-1} \|\tilde{f}_\lambda\|_{p_1}^{p_1} d\lambda &= \frac{p_1}{p_1-p} \|f\|_{L^p}^p.\end{aligned}\tag{2.5}$$

as well as

$$\begin{aligned}\|\tilde{f}^\lambda\|_{p_0}^{p_0} &\leq \frac{\Gamma(p_0+1)\Gamma(p-p_0)}{\Gamma(p)} \lambda^{p_0-p} \|f\|_{L^{p,\infty}}^p, \\ \|\tilde{f}_\lambda\|_{p_1}^{p_1} &\leq \frac{p_1}{p_1-p} \lambda^{p_1-p} \|f\|_{L^{p,\infty}}^p;\end{aligned}\tag{2.6}$$

moreover, we have $\|\tilde{f}_\lambda\|_\infty \leq \lambda$.

Observe that we have identities in the strong cases and estimates in the weak cases, but the constants on the right side are equal in both cases. The fractions involving the Γ function will arise from the following standard formulae:

$$\int_0^1 u^{\alpha-1} (1-u)^{\beta-1} du = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \int_0^\infty \frac{u^{\alpha-1}}{(1+u)^{\alpha+\beta}} du, \quad \alpha, \beta > 0.$$

If one is not interested in the specific quantitative form of the bounds in the lemma and then in Theorem 2.2.3, the simpler splitting

$$f = f^\lambda + f_\lambda, \quad f^\lambda = f \cdot \mathbf{1}_{\{\|f\| > \lambda\}}, \quad f_\lambda = f \cdot \mathbf{1}_{\{\|f\| \leq \lambda\}}$$

would work just as well.

Proof of Lemma 2.2.5. For the first identity in (2.5), we have

$$\begin{aligned}\int_0^\infty p\lambda^{p-p_0-1} \|\tilde{f}^\lambda\|_{p_0}^{p_0} d\lambda &= \int_0^\infty p\lambda^{p-p_0-1} \int_S (\|f\| - \lambda)_+^{p_0} d\mu d\lambda \\ &= \int_S \int_0^{\|f\|} p\lambda^{p-p_0-1} (\|f\| - \lambda)^{p_0} d\lambda d\mu \\ &= p\|f\|_p^p \int_0^1 u^{p-p_0-1} (1-u)^{p_0} du \\ &= p\|f\|_p^p \frac{\Gamma(p-p_0)\Gamma(p_0+1)}{\Gamma(p+1)},\end{aligned}$$

and the proof is completed by using $\Gamma(p+1) = p\Gamma(p)$. For the second identity in (2.5), we compute

$$\begin{aligned} \int_0^\infty p\lambda^{p-p_1-1} \|\tilde{f}_\lambda\|_{p_1}^{p_1} d\lambda &= \int_0^\infty p\lambda^{p-p_1-1} \int_S (\|f\| \wedge \lambda)^{p_1} d\mu d\lambda \\ &= \int_S \left(\int_0^{\|f\|} p\lambda^{p-1} d\lambda + \int_{\|f\|}^\infty p\lambda^{p-p_1-1} \|f\|^{p_1} d\lambda \right) d\mu \\ &= \|f\|_p^p \left(1 + \frac{p}{p_1 - p} \right) = \|f\|_p^p \frac{p_1}{p_1 - p}. \end{aligned}$$

Concerning the weak inequalities (2.6), we have

$$\begin{aligned} \|\tilde{f}^\lambda\|_{p_0}^{p_0} &= \int_S (\|f\| - \lambda)_+^{p_0} d\mu = \int_0^\infty p_0 t^{p_0-1} \mu(\|f\| > \lambda + t) dt \\ &\leq \int_0^\infty \frac{p_0 t^{p_0-1}}{(\lambda + t)^p} \|f\|_{L^{p,\infty}}^p dt = \int_0^\infty \frac{p_0 u^{p_0-1}}{(1+u)^p} du \cdot \lambda^{p_0-p} \|f\|_{L^{p,\infty}}^p \\ &= p_0 \frac{\Gamma(p_0)\Gamma(p-p_0)}{\Gamma(p)} \lambda^{p_0-p} \|f\|_{L^{p,\infty}}^p \\ &= \frac{\Gamma(p_0+1)\Gamma(p-p_0)}{\Gamma(p)} \lambda^{p_0-p} \|f\|_{L^{p,\infty}}^p, \end{aligned}$$

and

$$\begin{aligned} \|\tilde{f}_\lambda\|_{p_1}^{p_1} &= \int_S (\|f\| \wedge \lambda)^{p_1} d\mu = \int_0^\infty p_1 t^{p_1-1} \mu(\|f\| \wedge \lambda > t) d\mu \\ &= \int_0^\lambda p_1 t^{p_1-1} \mu(\|f\| > t) dt \leq \int_0^\lambda p_1 t^{p_1-p-1} \|f\|_{L^{p,\infty}}^p dt \\ &= \frac{p_1}{p_1 - p} \lambda^{p_1-p} \|f\|_{L^{p,\infty}}^p. \end{aligned}$$

The bound for $\|\tilde{f}_\lambda\|_\infty$ is obvious from the definition. $\square$

Now we are ready for:

Proof of the Marcinkiewicz Interpolation Theorem 2.2.3. Let $\alpha > 0$ be an auxiliary parameter to be chosen later and write $p = p_\theta$ for brevity. For $f \in L^p(S_0; X)$ or $f \in L^{p,\infty}(S_0; X)$ and every $\lambda > 0$, we make use of the splitting

$$f = \tilde{f}^{\alpha\lambda} + \tilde{f}_{\alpha\lambda}$$

from Lemma 2.2.5, but on the level $\alpha\lambda$ in place of λ . For the level sets $\{\|Tf\| > \lambda\}$, we then make the estimate

$$\mu_1(\|Tf\| > \lambda) \leq \mu_1(\|T\tilde{f}^{\alpha\lambda}\| > \theta_0\lambda) + \mu_1(\|T\tilde{f}_{\alpha\lambda}\| > \theta_1\lambda),$$

for positive numbers $\theta_0 + \theta_1 = 1$ also to be chosen. With this general strategy, we then consider several cases in turn.

Case $p_1 < \infty$

For the strong inequality, we have

$$\begin{aligned}
\|Tf\|_p^p &= \int_0^\infty p\lambda^p \mu_1(\|Tf\| > \lambda) \frac{d\lambda}{\lambda} \\
&\leq p \int_0^\infty \lambda^p [\mu_1(\|T\tilde{f}^{\alpha\lambda}\| > \theta_0\lambda) + \mu_1(\|T\tilde{f}_{\alpha\lambda}\| > \theta_1\lambda)] \frac{d\lambda}{\lambda} \\
&\leq p \int_0^\infty \lambda^p \left[\left(\frac{A_0}{\theta_0\lambda}\right)^{p_0} \|\tilde{f}^{\alpha\lambda}\|_{p_0}^{p_0} + \left(\frac{A_1}{\theta_1\lambda}\right)^{p_1} \|\tilde{f}_{\alpha\lambda}\|_{p_1}^{p_1} \right] \frac{d\lambda}{\lambda} \\
&= \left[\left(\frac{A_0}{\theta_0}\right)^{p_0} \alpha^{p_0-p} \frac{\Gamma(p-p_0)\Gamma(p_0+1)}{\Gamma(p)} + \left(\frac{A_1}{\theta_1}\right)^{p_1} \frac{\alpha^{p_1-p}p_1}{p_1-p} \right] \|f\|_p^p,
\end{aligned}$$

where the last step follows from (2.5) after simple scaling. Similarly, in the weak case with $p_1 < \infty$, we obtain

$$\begin{aligned}
\|Tf\|_{L^{p,\infty}}^p &= \sup_{\lambda>0} \lambda^p \mu_1(\|Tf\| > \lambda) \\
&\leq \sup_{\lambda>0} \lambda^p [\mu_1(\|T\tilde{f}^{\alpha\lambda}\| > \theta_0\lambda) + \mu_1(\|T\tilde{f}_{\alpha\lambda}\| > \theta_1\lambda)] \\
&\leq \sup_{\lambda>0} \lambda^p \left[\left(\frac{A_0}{\theta_0\lambda}\right)^{p_0} \|\tilde{f}^{\alpha\lambda}\|_{p_0}^{p_0} + \left(\frac{A_1}{\theta_1\lambda}\right)^{p_1} \|\tilde{f}_{\alpha\lambda}\|_{p_1}^{p_1} \right] \\
&\leq \left[\left(\frac{A_0}{\theta_0}\right)^{p_0} \alpha^{p_0-p} \frac{\Gamma(p-p_0)\Gamma(p_0+1)}{\Gamma(p)} + \left(\frac{A_1}{\theta_1}\right)^{p_1} \frac{\alpha^{p_1-p}p_1}{p_1-p} \right] \|f\|_{L^{p,\infty}}^p,
\end{aligned}$$

where the last step is a consequence of (2.6).

Optimising in $\alpha > 0$, we find the root of the derivative

$$\alpha = \left(\frac{\Gamma(p-p_0+1)\Gamma(p_0+1)}{p_1\Gamma(p)} \right)^{1/(p_1-p_0)} \left(\frac{A_0}{\theta_0} \right)^{p_0/(p_1-p_0)} \left(\frac{A_1}{\theta_1} \right)^{-p_1/(p_1-p_0)}$$

which, upon substitution to the preceding estimates, and making the optimal choice $\theta_0 = 1 - \theta$, $\theta_1 = \theta$, gives the claimed bounds (2.4) for $p_1 < \infty$.

Case $p_1 = \infty$

If $p_1 = \infty$, we choose $\alpha = \theta_1/A_1$. Then $\|\tilde{f}_{\alpha\lambda}\|_\infty \leq \alpha\lambda = \lambda\theta_1/A_1$, hence $\|T\tilde{f}_{\alpha\lambda}\|_\infty \leq A_1\|\tilde{f}_{\alpha\lambda}\|_\infty \leq \theta_1\lambda$, and thus the part involving $\mu_1(\|T\tilde{f}_{\alpha\lambda}\| > \theta_1\lambda) = 0$ is now absent from the computations. Running the same computations as for $p_1 < \infty$, we then arrive at

$$\begin{aligned}
\|Tf\|_p^p &\leq \left(\frac{A_0}{\theta_0}\right)^{p_0} \alpha^{p_0-p} \frac{\Gamma(p-p_0)\Gamma(p_0+1)}{\Gamma(p)} \|f\|_p^p \\
&\leq \left(\frac{A_0}{\theta_0}\right)^{p_0} \left(\frac{A_1}{\theta_1}\right)^{p-p_0} \frac{\Gamma(p-p_0)\Gamma(p_0+1)}{\Gamma(p)} \|f\|_p^p,
\end{aligned}$$

and exactly the same bound with all L^p norms replaced by $L^{p,\infty}$. Choosing $\theta_0 = 1 - \theta$, $\theta_1 = \theta$ as before, we arrive at the claim (2.4) for $p_1 = \infty$. $\square$

2.2.c Complex interpolation of the spaces $L^p(S; X)$

For interpolation couples (X_0, X_1) of complex Banach spaces, the complex interpolation method is introduced in Section C.2. For any measure space $(S, \mathcal{A}, \mu)$ and any choice of exponents $1 \leq p_0, p_1 \leq \infty$, the pair $(L^{p_0}(S; X_0), L^{p_1}(S; X_1))$ is an interpolation couple as well (both embed into the metric space $L^0(S; X_0 + X_1)$; see Example C.1.2), and we have the following representation of their complex interpolation spaces.

Theorem 2.2.6 (Complex interpolation of the spaces $L^p(S; X)$). *Let $1 \leq p_0 \leq p_1 < \infty$ or $1 \leq p_0 < p_1 = \infty$, and let $0 < \theta < 1$. For any interpolation couple (X_0, X_1) of complex Banach spaces and any measure space $(S, \mathcal{A}, \mu)$ we have*

$$[L^{p_0}(S; X_0), L^{p_1}(S; X_1)]_\theta = L^{p_\theta}(S; [X_0, X_1]_\theta)$$

isometrically, with $\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$.

The proof will be based on the following approximation result.

Lemma 2.2.7. *Let $\Sigma(X_0, X_1)$ denote the complex vector space of all μ -simple functions $f : S \rightarrow X_0 \cap X_1$. Under the assumptions of Theorem 2.2.6, the space $\Sigma(X_0, X_1)$ is dense in both $L^{p_\theta}(S; [X_0, X_1]_\theta)$ and $[L^{p_0}(S; X_0), L^{p_1}(S; X_1)]_\theta$.*

Proof. Density in the space $L^{p_\theta}(S; [X_0, X_1]_\theta)$ is a consequence of Corollary C.2.8, so it remains the result for $[L^{p_0}(S; X_0), L^{p_1}(S; X_1)]_\theta$.

We know that $L^{p_0}(S; X_0) \cap L^{p_1}(S; X_1)$ is dense in $[L^{p_0}(S; X_0), L^{p_1}(S; X_1)]_\theta$ by Corollary C.2.8. Thus, given an $f \in L^{p_0}(S; X_0) \cap L^{p_1}(S; X_1)$ and an $\varepsilon > 0$ it suffices to find $g \in \Sigma(X_0, X_1)$ such that

$$\|f - g\|_{[L^{p_0}(S; X_0), L^{p_1}(S; X_1)]_\theta} < \varepsilon.$$

We begin by noting that if $\mathbf{1}_{S_n} \uparrow \mathbf{1}$ μ -almost everywhere on the set $\{f \neq 0\}$, then by (C.1)

$$\begin{aligned} \|f - \mathbf{1}_{S_n} f\|_{[L^{p_0}(S; X_0), L^{p_1}(S; X_1)]_\theta} &\leq \|f - \mathbf{1}_{S_n} f\|_{L^{p_0}(S; X_0)}^{1-\theta} \|f - \mathbf{1}_{S_n} f\|_{L^{p_1}(S; X_1)}^\theta \\ &\leq \|f - \mathbf{1}_{S_n} f\|_{L^{p_0}(S; X_0)}^{1-\theta} \|f\|_{L^{p_1}(S; X_1)}^\theta \rightarrow 0, \end{aligned}$$

where we used that p_0 is finite. Taking $S_n := \{\|f\|_{X_0} \geq \frac{1}{n}\}$ it follows that in the rest of the proof we may assume that f has support in a set S' of finite measure. One can check that for such functions it holds that

$$\|f\|_{[L^{p_0}(S; X_0), L^{p_1}(S; X_1)]_\theta} = \|f\|_{[L^{p_0}(S'; X_0), L^{p_1}(S'; X_1)]_\theta}.$$

Almost everywhere on S' , f takes values in $X_0 \cap X_1$. We will show that f is strongly μ -measurable as a function with values in $X_0 \cap X_1$.

Indeed, it is immediate that $X_0^* + X_1^*$ is norming for $X_0 \cap X_1$, and for $x^* \in X_0^* + X_1^*$ the function $\langle f, x^* \rangle$ is μ -measurable. Also, by the strong μ -measurability of f in X_0 and X_1 there are separable subspaces $X'_0 \subseteq X_0$ and $X'_1 \subseteq X_1$ such that f takes values in $X'_0 \cap X'_1$ almost everywhere. This intersection is separable in $X_0 \cap X_1$: for if we cover X'_0 and X'_1 with countably many balls $B_{0,i}$ and $B_{1,j}$ respectively of prescribed radii $\delta > 0$, then the sets $B_{ij} := B_{0,i} \cap B_{1,j}$ cover $X'_0 \cap X'_1$ and are contained in $X_0 \cap X_1$ -open balls of radius 2δ . This allows us to pick, for any given $\delta > 0$, a countable 2δ -net in $X'_0 \cap X'_1$.

We are now in a position to apply the Pettis measurability theorem and conclude that f is strongly μ -measurable as a function with values in $X_0 \cap X_1$. Therefore we can find a countably-valued function $g : S' \rightarrow X_0 \cap X_1$ of the form $g(s) = \sum_{k \geq 1} \mathbf{1}_{B_k}(s)x_k$ with disjoint $B_k \subseteq S'$ such that

$$\|f - g\|_{L^\infty(S'; X_0 \cap X_1)} < \varepsilon.$$

Set $C_k := \bigcup_{j=1}^k B_k$. Then $\mathbf{1}_{C_k}g$ is μ -simple and therefore it belongs to $\Sigma(X_0, X_1)$. Using (C.1) we may estimate

$$\|f - \mathbf{1}_{C_k}g\|_{[L^{p_0}(S; X_0), L^{p_1}(S; X_1)]_\theta} \leq \|f - \mathbf{1}_{C_k}g\|_{L^{p_0}(S'; X_0)}^{1-\theta} \|f - \mathbf{1}_{C_k}g\|_{L^{p_1}(S'; X_1)}^\theta.$$

For k large enough,

$$\begin{aligned} \|f - \mathbf{1}_{C_k}g\|_{L^{p_0}(S'; X_0)} &\leq \|f - g\|_{L^{p_0}(S'; X_0)} + \|g - \mathbf{1}_{C_k}g\|_{L^{p_0}(S'; X_0)} \\ &\leq \mu(S')^{1/p_0} \|f - g\|_{L^\infty(S'; X_0)} + \varepsilon \leq \varepsilon(\mu(S')^{1/p_0} + 1). \end{aligned}$$

Also,

$$\begin{aligned} \|f - \mathbf{1}_{C_k}g\|_{L^{p_1}(S'; X_1)} &\leq \|f\|_{L^{p_1}(S'; X_1)} + \|g\|_{L^{p_1}(S'; X_1)} \\ &\leq 2\|f\|_{L^{p_1}(S'; X_1)} + \|g - f\|_{L^{p_1}(S'; X_1)} \\ &\leq 2\|f\|_{L^{p_1}(S'; X_1)} + \varepsilon\mu(S')^{1/p_1}. \end{aligned}$$

Putting together these three estimates, for large enough k we obtain the bound

$$\begin{aligned} \|f - \mathbf{1}_{C_k}g\|_{[L^{p_0}(S; X_0), L^{p_1}(S; X_1)]_\theta} &\leq \varepsilon^{1-\theta}(\mu(S')^{1/p_0} + 1)^{1-\theta} (2\|f\|_{L^{p_1}(S'; X_1)} + \varepsilon\mu(S')^{1/p_1})^\theta \end{aligned}$$

from which the density follows. $\square$

Proof of Theorem 2.2.6. We will write $p = p_\theta$ for simplicity. In view of Lemma 2.2.7, to complete the proof it suffices to prove the equality of norms

$$\|f\|_{[L^{p_0}(S; X_0), L^{p_1}(S; X_1)]_\theta} = \|f\|_{L^p(S; [X_0, X_1]_\theta)} \quad (2.7)$$

for functions $f \in \Sigma(X_0, X_1)$.

For the rest of the proof we fix a non-zero $f \in \Sigma(X_0, X_1)$, say $f = \sum_{n=1}^N \mathbf{1}_{A_n} \otimes x_n$ with $A_n \subseteq S$ measurable and disjoint and x_n in $X_0 \cap X_1$.

We first prove the inequality ‘ $\leq$ ’ in (2.7). Fix $\varepsilon > 0$ arbitrary. For $1 \leq n \leq N$ we pick functions $h_n \in \mathcal{H}(X_0, X_1)$ such that $h_n(\theta) = x_n$ and

$$\|h_n\|_{\mathcal{H}(X_0, X_1)} \leq (1 + \varepsilon)\|x_n\|_{[X_0, X_1]_\theta}.$$

Consider the μ -simple function $g : S \rightarrow \mathcal{H}(X_0, X_1)$, $g = \sum_{n=1}^N \mathbf{1}_{A_n} \otimes h_n$, and define the continuous function $F : \mathbb{S} \rightarrow L^{p_0}(S; X_0) + L^{p_1}(S; X_1)$ by

$$(F(z))(s) := (g(s))(z) \left[\frac{\|f(s)\|}{\|f\|_p} \right]^{p(\frac{1}{p_0} - \frac{1}{p_1})(\theta - z)}, \quad s \in S,$$

writing

$$\|f(s)\| := \|f(s)\|_{[X_0, X_1]_\theta}, \quad \|f\|_p := \|f\|_{L^p(S; [X_0, X_1]_\theta)}$$

for brevity. This function is holomorphic on $\mathbb{S}$ and satisfies $F(\theta) = f(\theta)$. Moreover, $v \mapsto F(iv)$ and $v \mapsto F(1 + iv)$ are continuous as functions with values $L^{p_0}(S; X_0)$ and $L^{p_1}(S; X_1)$, respectively. Noting that $\theta p p_0(\frac{1}{p_0} - \frac{1}{p_1}) = -p p_0(\frac{1}{p} - \frac{1}{p_0}) = p - p_0$ we obtain

$$\begin{aligned} \|F(iv)\|_{L^{p_0}(S; X_0)}^{p_0} &= \int_S \left\| (g(s))(iv) \left[\frac{\|f(s)\|}{\|f\|_p} \right]^{p(\frac{1}{p_0} - \frac{1}{p_1})(\theta - iv)} \right\|_{X_0}^{p_0} d\mu(s) \\ &= \sum_{n=1}^N \int_{A_n} \|h_n(iv)\|_{X_0}^{p_0} \frac{\|f(s)\|^{p-p_0}}{\|f\|_p^{p-p_0}} d\mu(s) \\ &\leq (1 + \varepsilon) \sum_{n=1}^N \|x_n\|_{[X_0, X_1]_\theta}^{p_0} \int_{A_n} \frac{\|f(s)\|^{p-p_0}}{\|f\|_p^{p-p_0}} d\mu(s) \\ &= (1 + \varepsilon) \int_S \|f(s)\|^{p_0} \frac{\|f(s)\|^{p-p_0}}{\|f\|_p^{p-p_0}} d\mu(s) \\ &= (1 + \varepsilon)^{p_0} \|f\|_p^{p_0}. \end{aligned}$$

Similarly,

$$\|F(1 + iv)\|_{L^{p_1}(S; X_1)} \leq (1 + \varepsilon)\|f\|_p.$$

From these estimates we infer that $F \in \mathcal{H}(L^{p_0}(S; X_0), L^{p_1}(S; X_1))$, so $f \in [L^{p_0}(S; X_0), L^{p_1}(S; X_1)]_\theta$, and

$$\|f\|_{[L^{p_0}(S; X_0), L^{p_1}(S; X_1)]_\theta} \leq \|F\|_{\mathcal{H}(L^{p_0}(S; X_0), L^{p_1}(S; X_1))} \leq (1 + \varepsilon)\|f\|_p.$$

For the proof of the inequality ‘ $\geq$ ’ in (2.7), let again $f \in \Sigma(X_0, X_1)$ be given and pick a function $F \in \mathcal{H}(L^{p_0}(S; X_0), L^{p_1}(S; X_1))$ such that $F(\theta) = f$. By Lemma C.2.10 and Hölder’s inequality (applied with $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$), Jensen’s inequality and (C.4),

$$\|f\|_p \leq \left(\int_S \left[\frac{1}{1-\theta} \int_{-\infty}^{\infty} \|(F(it))(s)\|_{X_0} P_0(\theta, t) dt \right]^{p_0} d\mu(s) \right)^{(1-\theta)/p_0}$$

$$\begin{aligned}
& \times \left(\int_S \left[\frac{1}{\theta} \int_{-\infty}^{\infty} \|(F(1+it))(s)\|_{X_1} P_1(\theta, t) dt \right]^{p_1} d\mu(s) \right)^{\theta/p_1} \\
& \leq \left(\int_S \frac{1}{1-\theta} \int_{-\infty}^{\infty} \|(F(it))(s)\|_{X_0}^{p_0} P_0(\theta, t) dt d\mu(s) \right)^{(1-\theta)/p_0} \\
& \quad \times \left(\int_S \frac{1}{\theta} \int_{-\infty}^{\infty} \|(F(1+it))(s)\|_{X_1}^{p_1} P_1(\theta, t) dt d\mu(s) \right)^{\theta/p_1} \\
& \leq \sup_{v \in \mathbb{R}} \|F(iv)\|_{L^p(S; X_0)}^{1-\theta} \cdot \sup_{v \in \mathbb{R}} \|F(1+iv)\|_{L^p(S; X_1)}^{\theta} \\
& \leq \max \left\{ \sup_{v \in \mathbb{R}} \|F(iv)\|_{(L^{p_0}(S; X_0))}, \sup_{v \in \mathbb{R}} \|F(1+iv)\|_{(L^{p_1}(S; X_1))} \right\} \\
& \leq \|F\|_{\mathcal{H}(L^{p_0}(S; X_0), L^{p_1}(S; X_1))}.
\end{aligned}$$

Taking the infimum over all admissible F it follows that

$$\|f\|_p \leq \|f\|_{[L^{p_0}(S; X_0), L^{p_1}(S; X_1)]_{\theta}}.$$

This concludes the proof of the theorem. $\square$

Example 2.2.8. As an application, let us prove that if T is a bounded operator on $L^p(S)$ for some $1 \leq p \leq 2$, then $T \otimes I_X$ has a bounded extension to $L^p(S; X)$ when $X = L^q(T)$ for all $1 \leq p \leq q \leq 2$. A slightly less general statement was asserted without proof in the remark following the proof of Proposition 2.1.2.

First of all, $T \otimes I_{L^p(T)}$ has a bounded extension to $L^p(S; L^p(T))$ by Proposition 2.1.2. Also, $T \otimes I_{L^2(T)}$ has a bounded extension to $L^p(S; L^2(T))$ by Theorem 2.1.9. For $p < q < 2$ these extensions interpolate, by complex interpolation, to a bounded extension of $T \otimes I_{L^q(T)}$ on $L^p(S; L^q(T)) = [L^p(S; L^p(T)), L^2(S; L^2(T))]_{\theta}$, where $\theta \in (0, 1)$ is determined by $1/q = (1-\theta)/p + \theta/2$.

As a second application we present a neat proof by interpolation of the Clarkson inequalities.

Corollary 2.2.9 (Clarkson inequalities). *Let $(S, \mathcal{A}, \mu)$ be a measure space.*

(1) *For all $1 \leq p \leq 2$ and $f, g \in L^p(S)$,*

$$(\|f + g\|_p^p + \|f - g\|_p^p)^{1/p} \leq 2^{1/p} (\|f\|_p^p + \|g\|_p^p)^{1/p}$$

and

$$(\|f + g\|_p^{p'} + \|f - g\|_p^{p'})^{1/p'} \leq 2^{1/p'} (\|f\|_p^p + \|g\|_p^p)^{1/p}.$$

(2) *For all $2 \leq p < \infty$ and $f, g \in L^p(S)$,*

$$(\|f + g\|_p^p + \|f - g\|_p^p)^{1/p} \leq 2^{1/p'} (\|f\|_p^p + \|g\|_p^p)^{1/p}$$

and

$$(\|f + g\|_p^p + \|f - g\|_p^p)^{1/p} \leq 2^{1/p} (\|f\|_p^{p'} + \|g\|_p^{p'})^{1/p'}.$$

Proof. We shall prove the two inequalities in (1); the inequalities in (2) may be proved similarly.

Let $1 \leq p \leq 2$. On the spaces $L^r(S) \oplus_r L^r(S) = \ell^r(\{1, 2\}; L^r(S))$ we consider the operator

$$T : (f, g) \mapsto (f + g, f - g).$$

Then $\|T\|_1 = 2$ and, by the parallelogram identity, $\|T\|_2 = \sqrt{2}$. Writing $\frac{1}{p} = \frac{1-\theta}{1} + \frac{\theta}{2} = 1 - \frac{\theta}{2}$ (so that $\theta = \frac{2}{p'}$), by Theorem 2.2.6 we obtain

$$\|T\|_p \leq 2^{1-2/p'} \cdot \sqrt{2}^{2/p'} = 2^{1-1/p'} = 2^{1/p}.$$

This is the first inequality in (1). To prove the second inequality we note that

$$\begin{aligned} \|T(f, g)\|_{L^1(S) \oplus_\infty L^1(S)} &= \max\{\|f + g\|_1, \|f - g\|_1\} \\ &\leq \|f\|_1 + \|g\|_1 = \|(f, g)\|_{L^1(S) \oplus_1 L^1(S)}. \end{aligned}$$

Note that $L^1(S) \oplus_r L^1(S) = \ell^r(\{1, 2\}; L^1(S))$ and similarly $L^2(S) \oplus_2 L^2(S) = \ell^2(\{1, 2\}; L^2(S))$. Writing $\frac{1}{p} = \frac{1-\theta}{1} + \frac{\theta}{2} = \frac{1-\theta}{2}$ (so that $\theta = 2(1 - 1/p)$), by Theorem 2.2.6 we obtain

$$\|T\|_{L^p(S) \oplus_p L^p(S) \rightarrow L^p(S) \oplus_{p'} L^p(S)} \leq 1^{1-\theta} \sqrt{2}^\theta = 2^{1-1/p} = 2^{1/p'}.$$

This is the second inequality in (1). □

Complex interpolation of real Banach spaces

There are instances where one would like to use the complex interpolation method in the context of real Banach spaces. This happens, in particular, in applications of interpolation in the theory of partial differential equations with real-valued coefficients, where complex interpolation enters in the identification of fractional powers of differential operators. In view of the fundamental result (Corollary C.2.8) that if (Y_0, Y_1) is an interpolation couple of complex Banach spaces, then the intersection $Y_0 \cap Y_1$ is dense in $[Y_0, Y_1]_\theta$, a natural definition can be given as follows.

Suppose that (X_0, X_1) is an interpolation couple of real Banach spaces, and let Y_i be the complexifications of X_i for $i = 0, 1$ (these are unique up to an equivalent norm). We may then define

$$[X_0, X_1]_\theta := \overline{X_0 \cap X_1}^{[Y_0, Y_1]_\theta}.$$

As a real-linear subspace of the complex Banach space $[Y_0, Y_1]_\theta$, this is a real Banach space in a natural way. For example, we have

$$[L^{p_0}(S; \mathbb{R}), L^{p_1}(S; \mathbb{R})]_\theta = L^{p_\theta}(S; \mathbb{R})$$

up to an equivalent norm, provided $1 \leq p_0, p_1 < \infty$ and $\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$ as usual; this identification is isometric if one complexifies $L^{p_\theta}(S; \mathbb{R})$ using the natural norm of $L^{p_\theta}(S; \mathbb{C})$.

2.2.d Real interpolation of the spaces $L^p(S; X)$

Having dealt with complex interpolation between $L^{p_0}(S; X_0)$ and $L^{p_1}(S; X_1)$, we now turn our attention to the real interpolation of these spaces. We recall from Section C.3 that for any interpolation couple (E_0, E_1) of Banach spaces one has

$$(E_0, E_1)_{\theta, p} = (E_0, E_1)_{\theta, p_0, p_1}$$

with equivalent norms, where $(E_0, E_1)_{\theta, p}$ is the usual real interpolation space (with $\frac{1}{p} := \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$) and $(E_0, E_1)_{\theta, p_0, p_1}$ is the interpolation space obtained by the so-called second mean method; see Theorem C.3.14. The use of the latter space turns the isomorphic identification in the next theorem into an isometric one.

Theorem 2.2.10 (Real interpolation of the spaces $L^p(S; X)$). *Let $1 \leq p_0 < \infty$ and $p_0 \leq p_1 \leq \infty$, let $0 < \theta < 1$, and set $\frac{1}{p_\theta} := \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$. For any interpolation couple (X_0, X_1) of Banach spaces and any measure space $(S, \mathcal{A}, \mu)$ we have a natural isometric isomorphism of Banach spaces*

$$(L^{p_0}(S; X_0), L^{p_1}(S; X_1))_{\theta, p_0, p_1} = L^{p_\theta}(S; (X_0, X_1)_{\theta, p_0, p_1}).$$

The proof of Theorem 2.2.10 is based on an analogue of Lemma 2.2.7, a proof of which is obtained by repeating the proof of this lemma *verbatim*, the only difference being that we use the log-convexity estimate of Lemma C.3.11 and the density result of Lemma C.3.12 instead of the log-convexity estimate (C.1).

Lemma 2.2.11. *Under the assumptions of Theorem 2.2.10, the space of all μ -simple functions $f : S \rightarrow X_0 \cap X_1$ is dense in $(L^{p_0}(S; X_0), L^{p_1}(S; X_1))_{\theta, p_0, p_1}$.*

Proof of Theorem 2.2.10. We set $p := p_\theta$ and $Y_j := L^{p_j}(S; X_j)$ for $j \in \{0, 1\}$. Observe that the μ -simple functions with values in $X_0 \cap X_1$ are dense in both $L^p(S; (X_0, X_1)_{\theta, p_0, p_1})$ (by Corollary C.3.15) and $(Y_0, Y_1)_{\theta, p}$ (by Lemma 2.2.11; the finiteness of p_0 is used here). Therefore, by a Cauchy sequence argument, one sees that the theorem follows as soon as we have proved the norm estimate for an arbitrary μ -simple function $f : S \rightarrow X_0 \cap X_1$.

Step 1 – In this step we prove the inclusion “ $\subseteq$ ”.

Referring to the definition of $(E_0, E_1)_{\theta, p_0, p_1}$ in Appendix C, choose a strongly measurable function $u : (0, \infty) \rightarrow Y_0 \cap Y_1$ with the following properties:

- (i) $t \mapsto t^{j-\theta} u(t)$ belongs to $L^{p_j}(\mathbb{R}_+, \frac{dt}{t}; Y_j)$ for $j \in \{0, 1\}$;
- (ii) $f = \int_0^\infty u(t) \frac{dt}{t}$ with convergence of the improper integral in the norm of $Y_0 + Y_1$.

By Proposition 1.2.25, for almost all $s \in S$ we have $f(s) = \int_0^\infty u(t, s) \frac{dt}{t}$ with convergence of the integral in $X_0 + X_1$, with $u(t, s) := (u(t))(s)$. By Lemma C.3.11 we obtain, for almost all $s \in S$,

$$\begin{aligned} & \|f(s)\|_{(X_0, X_1)_{\theta, p_0, p_1}} \\ & \leq \|t \mapsto t^{-\theta} u(t, s)\|_{L^{p_0}(\mathbb{R}_+, \frac{dt}{t}; X_0)}^{1-\theta} \cdot \|t \mapsto t^{1-\theta} u(t, s)\|_{L^{p_1}(\mathbb{R}_+, \frac{dt}{t}; X_1)}^{\theta}. \end{aligned}$$

Then by Hölder's inequality (with $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$) and Fubini's theorem,

$$\begin{aligned} \|f\|_{L^p(S; (X_0, X_1)_{\theta, p_0, p_1})} & \leq \left(\int_S \|t \mapsto t^{-\theta} u(t, s)\|_{L^{p_0}(\mathbb{R}_+, \frac{dt}{t}; X_0)}^{p_0} d\mu(s) \right)^{(1-\theta)/p_0} \\ & \quad \times \left(\int_S t \mapsto \|t^{1-\theta} u(t, s)\|_{L^{p_1}(\mathbb{R}_+, \frac{dt}{t}; X_1)}^{p_1} d\mu(s) \right)^{\theta/p_1} \\ & = \left(\int_0^\infty \int_S \|t \mapsto t^{-\theta} u(t, s)\|_{X_0}^{p_0} d\mu(s) \frac{dt}{t} \right)^{(1-\theta)/p_0} \\ & \quad \times \left(\int_0^\infty \int_S \|t \mapsto t^{1-\theta} u(t, s)\|_{X_1}^{p_1} d\mu(s) \frac{dt}{t} \right)^{\theta/p_1} \\ & = \|t \mapsto t^{-\theta} u(t, \cdot)\|_{L^{p_0}(\mathbb{R}_+, \frac{dt}{t}; Y_0)}^{1-\theta} \\ & \quad \times \|t \mapsto t^{1-\theta} u(t, \cdot)\|_{L^{p_1}(\mathbb{R}_+, \frac{dt}{t}; Y_1)}^{\theta}. \end{aligned}$$

Taking the infimum over all admissible u in the representation for f , another appeal to Lemma C.3.11 concludes this part of the proof.

Step 2 – In this step we prove the inclusion “ $\supseteq$ ” part. By assumption we have $f = \sum_{n=1}^N \mathbf{1}_{A_n} \otimes x_n$, with disjoint sets $A_n \in \mathcal{A}$ satisfying $\mu(A_n) < \infty$ and non-zero elements $x_n \in X_0 \cap X_1$. Fix $\varepsilon > 0$. By Definition C.3.10 for each n we may choose a strongly measurable function $u_n : (0, \infty) \rightarrow X_0 \cap X_1$ such that $x_n = \int_0^\infty u_n(t) \frac{dt}{t}$ with convergence of the integral in $X_0 + X_1$, and

$$\|t^{j-\theta} u_n(t)\|_{L^{p_j}(\mathbb{R}_+, \frac{dt}{t}; X_j)} \leq (1 + \varepsilon) \|x_n\|_{(X_0, X_1)_{\theta, p_0, p_1}}$$

for $j \in \{0, 1\}$. Set

$$v_n(t) := u_n(t) \|x_n\|_{(X_0, X_1)_{\theta, p_0, p_1}}^{\nu},$$

where the exponent ν will be chosen shortly. Then

$$\int_0^\infty v_n(t) \frac{dt}{t} = \int_0^\infty u_n(t) \frac{dt}{t} = x_n,$$

and, substituting $\tau = t \|x_n\|_{(X_0, X_1)_{\theta, p_0, p_1}}^{\nu}$, for $j \in \{0, 1\}$ we obtain

$$\begin{aligned} \|t \mapsto t^{j-\theta} v_n(t)\|_{L^{p_j}(\mathbb{R}_+, \frac{dt}{t}; X_j)} & = \left(\int_0^\infty t^{(j-\theta)p_j} \|v_n(t)\|_{X_j}^{p_j} \frac{dt}{t} \right)^{1/p_j} \\ & = \|x_n\|_{(X_0, X_1)_{\theta, p_0, p_1}}^{(\theta-j)\nu} \left(\int_0^\infty \tau^{(j-\theta)p_j} \|u_n(\tau)\|_{X_j}^{p_j} \frac{d\tau}{\tau} \right)^{1/p_j} \\ & \leq \|x_n\|_{(X_0, X_1)_{\theta, p_0, p_1}}^{(\theta-j)\nu} \times (1 + \varepsilon) \|x_n\|_{(X_0, X_1)_{\theta, p_0, p_1}} \\ & \leq (1 + \varepsilon) \|x_n\|_{(X_0, X_1)_{\theta, p_0, p_1}}^{(\theta-j)\nu+1}. \end{aligned}$$

We now take $\nu := p(\frac{1}{p_0} - \frac{1}{p_1})$; this choice gives $((\theta - j)\nu + 1)p_j = p$. Putting things together and using Lemma C.3.11 once more, we obtain

$$\begin{aligned}
\|f\|_{(Y_0, Y_1)_{\theta, p_0, p_1}} &= \left\| \sum_{n=1}^N \mathbf{1}_{A_n} x_n \right\|_{(Y_0, Y_1)_{\theta, p_0, p_1}} \\
&\leq \left\| t \mapsto t^{-\theta} \sum_{n=1}^N \mathbf{1}_{A_n} v_n(t) \right\|_{L^{p_0}(\mathbb{R}_+, \frac{dt}{t}; Y_0)}^{1-\theta} \\
&\quad \times \left\| t \mapsto t^{1-\theta} \sum_{n=1}^N \mathbf{1}_{A_n} v_n(t) \right\|_{L^{p_1}(\mathbb{R}_+, \frac{dt}{t}; Y_1)}^{\theta} \\
&= \left(\sum_{n=1}^N \mu(A_n) \|t \mapsto t^{-\theta} v_n(t)\|_{L^{p_0}(\mathbb{R}_+, \frac{dt}{t}; X_0)}^{p_0} \right)^{(1-\theta)/p_0} \\
&\quad \times \left(\sum_{n=1}^N \mu(A_n) \|t \mapsto t^{1-\theta} v_n(t)\|_{L^{p_1}(\mathbb{R}_+, \frac{dt}{t}; X_1)}^{p_1} \right)^{\theta/p_1} \\
&\leq (1 + \varepsilon) \left(\sum_{n=1}^N \mu(A_n) \|x_n\|_{(X_0, X_1)_{\theta, p_0, p_1}}^p \right)^{1/p} \\
&= (1 + \varepsilon) \|f\|_{L^p(S; (X_0, X_1)_{\theta, p_0, p_1})}.
\end{aligned}$$

□

2.3 The Hardy–Littlewood maximal operator

In the analysis of functions defined on the Euclidean space $\mathbb{R}^d$, the Hardy–Littlewood maximal operator plays a distinguished role. The maximal operator itself is sub-linear, but it is useful as a universal majorant for various classes of linear operators. The classical definition of the maximal operator immediately extends to the context of vector-valued functions simply by replacing the absolute values by the norms:

$$Mf(x) := \sup_{B \ni x} \int_B \|f(y)\| \, dy := \sup_{B \ni x} \frac{1}{|B|} \int_B \|f(y)\| \, dy, \quad (2.8)$$

where the supremum is taken over all Euclidean balls B that contain x , and the ‘average integral’ notation

$$\int_B := \frac{1}{|B|} \int_B$$

is handy to simplify writing. It is easy to check that an equivalent operator (one that differs at most by a multiplicative dimensional constant) is obtained by replacing balls by cubes, either arbitrary or ones parallel to coordinate axes.

Likewise, the condition that the balls or cubes contain x may be replaced by requiring that they be centred at x , changing the value of $Mf(x)$ at most by a multiplicative dimensional constant.

Since

$$Mf(x) = M(\|f\|_X)(x) \quad \text{ad} \quad \|f\|_{L^p(\mathbb{R}^d; X)} = \|\|f\|_X\|_{L^p(\mathbb{R}^d)},$$

all boundedness properties of M on the Bochner spaces $L^p(\mathbb{R}^d; X)$ can be immediately deduced from their scalar-valued analogues on $L^p(\mathbb{R}^d)$. Nevertheless, for completeness, we provide a self-contained approach here, although it essentially repeats the well-known arguments from the scalar case.

Lemma 2.3.1 (Vitali covering lemma). *Let $\mathcal{B}$ be a finite collection of balls in $\mathbb{R}^d$. Then there is a disjoint sub-collection $\mathcal{B}_0$ such that each $B \in \mathcal{B}$ is contained in some $3B'$, where $B' \in \mathcal{B}_0$; here $3B'$ is the ball with the same centre and 3 times the radius of B' .*

Proof. We proceed by induction on the number n of balls in the collection $\mathcal{B}$. If $n = 1$, the claim is trivial with $\mathcal{B}_0 = \mathcal{B}$. If the claim is verified for some n , let $\mathcal{B}$ be a collection of $n + 1$ balls, and let $\mathcal{B}' := \mathcal{B} \setminus \{B_0\}$, where B_0 is any ball in $\mathcal{B}$ of minimal radius. By the induction assumption, there is a disjoint sub-collection $\mathcal{B}'_0 \subseteq \mathcal{B}'$ such that each $B \in \mathcal{B}'$ is contained in $3B'$ for some $B' \in \mathcal{B}'_0$. There are two cases:

If B_0 is disjoint from each $B' \in \mathcal{B}'_0$, then $\mathcal{B}_0 := \mathcal{B}'_0 \cup \{B_0\}$ is a required sub-collection. Otherwise, B_0 intersects some $B' \in \mathcal{B}'_0$, which is at least as large as B_0 by construction. Then it follows that $B_0 \subseteq 3B'$, and hence $\mathcal{B}_0 := \mathcal{B}'_0$ is already a required sub-collection. $\square$

When $w \in L^0(\mathbb{R}^d)$ is a non-negative function, we shall write

$$L^p(\mathbb{R}^d, w; X) := L^p(\mathbb{R}^d, w(x) dx; X)$$

and, for Borel subsets $B \subseteq \mathbb{R}^d$,

$$w(B) := \int_B w(x) dx.$$

Theorem 2.3.2 (Hardy–Littlewood maximal theorem). *Let M be the Hardy–Littlewood maximal operator as defined in (2.8) and let $w \in L^0(\mathbb{R}^d)$ be a non-negative function.*

(1) *If $f \in L^1(\mathbb{R}^d, Mw; X)$, then for all $t > 0$,*

$$t w(\{Mf > t\}) = t \int_{\{Mf > t\}} w(y) dy \leq 3^d \|f\|_{L^1(\mathbb{R}^d, Mw; X)}.$$

(2) *If $f \in L^p(\mathbb{R}^d, Mw; X)$ with $1 < p \leq \infty$, then*

$$\|Mf\|_{L^p(\mathbb{R}^d, w)} \leq 3^{d/p} p' \|f\|_{L^p(\mathbb{R}^d, Mw; X)}, \quad \frac{1}{p} + \frac{1}{p'} = 1.$$

Proof. We first consider (1). By definition, for every $x \in \{Mf > t\}$, there exists a ball $B \ni x$ such that $\int_B \|f\| > t$. If $K \subseteq \{Mf > t\}$ is compact, it can be covered by a finite collection $\mathcal{B}$ of such balls. Let $\mathcal{B}_0$ be a disjoint sub-collection as provided by the Vitali covering lemma. Then

$$\begin{aligned}
 w(K) &\leq w\left(\bigcup_{B \in \mathcal{B}} B\right) \leq w\left(\bigcup_{B \in \mathcal{B}_0} 3B\right) \\
 &\leq \sum_{B \in \mathcal{B}_0} \frac{w(3B)}{|3B|} 3^d |B| \\
 &\stackrel{(*)}{\leq} \sum_{B \in \mathcal{B}_0} \inf_{z \in B} Mw(z) \frac{3^d}{t} \int_B \|f(y)\| \, dy \\
 &\leq \frac{3^d}{t} \sum_{B \in \mathcal{B}_0} \int_B \|f(y)\| Mw(y) \, dy \leq \frac{3^d}{t} \|f\|_{L^1(\mathbb{R}^d, Mw; X)},
 \end{aligned} \tag{2.9}$$

where $(*)$ follows from the choice of B along with the fact that for all $z \in B$ we have

$$Mw(z) = \sup_{B' \ni z} \frac{w(B')}{|B'|} \geq \frac{w(3B)}{|3B|}.$$

Since (2.9) holds for all compact subsets $K \subseteq \{Mf > t\}$, it also holds for $\{Mf > t\}$ in place of K .

We turn to (2) for $p = \infty$. If $w(x) = 0$ for Lebesgue-almost every $x \in \mathbb{R}^d$, the left side is zero and there is nothing to prove. Otherwise, it easily follows that $Mw(x) > 0$ at every x , simply by considering a ball $B \ni x$ large enough so as to include a part of positive measure of $\{w > 0\}$. It is immediate that $Mf(x) \leq \|f\|_{L^\infty(\mathbb{R}^d; X)}$ at every point $x \in \mathbb{R}^d$, and therefore

$$\|Mf\|_{L^\infty(\mathbb{R}^d, w; X)} \leq \|f\|_{L^\infty(\mathbb{R}^d; X)} \leq \|f\|_{L^\infty(\mathbb{R}^d, Mw; X)},$$

proving (2) for $p = \infty$.

Finally, the remaining case of (2) for $p \in (1, \infty)$ follows from the endpoint cases already considered via the Marcinkiewicz interpolation theorem as formulated in Corollary 2.2.4. $\square$

2.3.a Lebesgue points and differentiation

By $L^1_{\text{loc}}(\mathbb{R}^d; X)$ we denote the space (equivalence classes of) of functions $f : \mathbb{R}^d \rightarrow X$ that are *locally integrable*, i.e., integrable on every compact subset of $\mathbb{R}^d$. The following class of points related to a locally integrable function is useful, since various pointwise approximation properties hold true in these points:

Definition 2.3.3 (Lebesgue points). A point $x \in \mathbb{R}^d$ is called a Lebesgue point of the function $f \in L^1_{\text{loc}}(\mathbb{R}^d; X)$ if

$$\lim_{\substack{B \ni x \\ |B| \rightarrow 0}} \int_B \|f(y) - f(x)\| \, dy = 0,$$

where the limit is along all balls B containing x with $|B| \rightarrow 0$.

For any Lebesgue point $x \in \mathbb{R}^d$ of f one has

$$\lim_{\substack{B \ni x \\ |B| \rightarrow 0}} \int_B f(y) \, dy = f(x).$$

As a simple corollary to the Hardy–Littlewood maximal theorem we have the following:

Theorem 2.3.4 (Lebesgue differentiation theorem). *If $f \in L^1_{\text{loc}}(\mathbb{R}^d; X)$, then almost every $x \in \mathbb{R}^d$ is a Lebesgue point of f .*

Proof. Since the result is local in nature, it is easy to see that it suffices to prove the theorem under the stronger assumption that $f \in L^1(\mathbb{R}^d; X)$. Let

$$\Lambda f(x) := \limsup_{\substack{B \ni x \\ |B| \rightarrow 0}} \int_B \|f(y) - f(x)\| \, dy.$$

Then it is immediate that $\Lambda f(x) = 0$ if f is continuous at x , and $\Lambda f(x) \leq Mf(x) + \|f(x)\|$ at every point x . The theorem asserts that $\Lambda f(x) = 0$ almost everywhere. We will show that $\{\Lambda f > \varepsilon\}$ has measure zero for every $\varepsilon > 0$, which proves the claim. To this end, use density (see Lemma 1.2.31) to pick $g \in C_c(\mathbb{R}^d; X)$ such that $\|f - g\|_1 < \delta$, for a given $\delta > 0$. Now

$$\Lambda f = \Lambda(f - g + g) \leq \Lambda(f - g) + \Lambda g \leq M(f - g) + \|f - g\| + 0,$$

and hence

$$\begin{aligned} |\{\Lambda f > \varepsilon\}| &\leq |\{M(f - g) > \varepsilon/2\}| + |\{\|f - g\| > \varepsilon/2\}| \\ &\leq \frac{2 \cdot 3^d}{\varepsilon} \|f - g\|_1 + \frac{2}{\varepsilon} \|f - g\|_1 \leq \frac{2}{\varepsilon} (3^d + 1) \delta. \end{aligned}$$

Since $\delta > 0$ is arbitrary, we have shown that $|\{\Lambda f > \varepsilon\}| = 0$, as claimed. $\square$

In dimension one it is sometimes desirable to have the following one-sided version of the Lebesgue differentiation theorem:

Corollary 2.3.5. *If x is a Lebesgue point of a function $f \in L^1_{\text{loc}}(\mathbb{R}; X)$, then*

$$\lim_{h \downarrow 0} \frac{1}{h} \int_x^{x+h} \|f(y) - f(x)\| \, dy = \lim_{h \downarrow 0} \frac{1}{h} \int_{x-h}^x \|f(y) - f(x)\| \, dy = 0.$$

Proof. Considering $I = (x - h, x + h)$, so that $\frac{1}{h} = \frac{2}{|I|}$, the result for $\frac{1}{h} \int_x^{x+h}$ is immediate from

$$\frac{1}{h} \int_x^{x+h} \|f(y) - f(x)\| dy \leq \frac{2}{|I|} \int_I \|f(y) - f(x)\| dy.$$

The result for $\frac{1}{h} \int_x^{x+h}$ follows by the same argument. $\square$

As a variant of the same proof technique as in Theorem 2.3.4, we also provide the following general existence result for pointwise limits. We shall have an opportunity for its application in a concrete situation in our development of the theory of martingale transforms in Chapter 3.

Proposition 2.3.6. *Let (S, Σ, μ) be a measure space. Let $p \in [1, \infty)$ and suppose that $(T_n)_{n \geq 1}$ is a sequence of bounded linear operators from X to $L^{p, \infty}(S; Y)$ such that, for some constant $C \geq 0$ and all $x \in X$,*

$$\left\| \limsup_{n \rightarrow \infty} \|T_n x\|_Y \right\|_{L^{p, \infty}(S)} \leq C \|x\|.$$

Suppose furthermore that there is a dense linear subspace X_0 of X such that for $x_0 \in X_0$,

$$Tx_0(s) := \lim_{n \rightarrow \infty} T_n x_0(s) \text{ exists for almost all } s \in S. \quad (2.10)$$

Then T uniquely extends to a bounded linear operator from X to $L^{p, \infty}(S; Y)$ of norm $\leq C$ and (2.10) holds for all $x \in X$.

Remark 2.3.7. For later use we observe that the proposition remains true for doubly indexed sequences $(T_{m,n})_{m,n \geq 1}$ when passing to the limit $m, n \rightarrow \infty$. This can be seen by repeating the proof *mutatis mutandis*.

Proof. For $x \in X$ define the *oscillation* $Ox : S \rightarrow \mathbb{R}_+$ by

$$Ox(s) = \limsup_{n \rightarrow \infty} \limsup_{m \rightarrow \infty} \|T_n x(s) - T_m x(s)\|$$

and note that $Ox \leq 2 \sup_{n \geq 1} \|T_n x\|$ pointwise. We claim that for all $x \in X$, $Ox = 0$ almost everywhere. In order to prove this it is enough to show that $\mu(Ox > \varepsilon) = 0$ for all $\varepsilon > 0$. Since for all $x_0 \in X_0$, $Ox_0 = 0$ almost everywhere, it follows from the triangle inequality that

$$Ox \leq O(x - x_0) + Ox_0 \leq O(x - x_0) \leq 2 \sup_{n \geq 1} \|T_n(x - x_0)\|.$$

From this and the assumption we find

$$\mu(Ox > \varepsilon) \leq \mu(2 \sup_{n \geq 1} \|T_n(x - x_0)\| > \varepsilon) \leq \frac{2^p C^p}{\varepsilon^p} \|x - x_0\|^p.$$

Taking the infimum over all $x_0 \in X_0$, we find the required result.

From the claim we see that for each $x \in X$, $(T_n x)_{n \geq 1}$ is a Cauchy sequence in Y almost surely and hence converges to some function Tx almost everywhere. The operator T defined in this way is linear and its boundedness follows from the almost everywhere inequality $\|Tx\| \leq \limsup_{n \rightarrow \infty} \|T_n x\|$. $\square$

2.3.b Convolutions and approximation

As applications of the Lebesgue differentiation theorem, we shall develop some useful approximation results involving the convolutions

$$\phi_\varepsilon * f(x) := \int_{\mathbb{R}^d} \phi_\varepsilon(y) f(x-y) dy, \quad \phi_\varepsilon(y) := \frac{1}{\varepsilon^d} \phi\left(\frac{y}{\varepsilon}\right).$$

As a warm-up, we recall from Proposition 1.2.32 that if $f \in L^p(\mathbb{R}^d; X)$ for some $p \in [1, \infty)$, and $\phi \in L^1(\mathbb{R}^d)$, then $\phi_\varepsilon * f \rightarrow c_\phi f$ in $L^p(\mathbb{R}^d; X)$ as $\varepsilon \downarrow 0$, where $c_\phi := \int_{\mathbb{R}^d} \phi(y) dy$. Here we turn to the slightly more delicate pointwise convergence issue:

Theorem 2.3.8. *Suppose that $f \in L^1_{\text{loc}}(\mathbb{R}^d; X)$ and $\phi \in L^1(\mathbb{R}^d)$, and let x be a Lebesgue point of f . Then the convergence*

$$\lim_{\varepsilon \downarrow 0} \phi_\varepsilon * f(x) = c_\phi f(x), \quad c_\phi := \int_{\mathbb{R}^d} \phi(y) dy,$$

takes place under either of the following additional assumptions:

- (1) *ϕ is bounded and compactly supported, or*
- (2) *the least radially decreasing majorant of ϕ is integrable, namely,*

$$z \mapsto \Phi(z) := \operatorname{ess\,sup}_{|y| \geq |z|} |\phi(y)| \text{ belongs to } L^1(\mathbb{R}^d), \quad (2.11)$$

and in addition $Mf(x) < \infty$.

The condition $Mf(x) < \infty$ is automatically satisfied at Lebesgue points x for any $f \in L^p(\mathbb{R}^d; X)$, whatever the value of $p \in [1, \infty]$.

Of course, a sufficient condition for (2.11) is the existence of *any* radially decreasing majorant for ϕ in $L^1(\mathbb{R}^d)$; as a typical example, the estimate

$$|\phi(y)| \leq C(1 + |y|)^{-d-\varepsilon}$$

will do.

Proof. Recalling (1.4), in both cases we want to show that

$$F_\varepsilon := \int_{\mathbb{R}^d} |\phi(y)| \|f(x - \varepsilon y) - f(x)\| dy$$

converges to zero as $\varepsilon \downarrow 0$.

In case (1) this is immediate, since $|\phi| \leq C \mathbf{1}_{B_0}/|B_0|$ for some large enough ball B_0 centred at the origin, and hence

$$F_\varepsilon \leq C \int_{B_0} \|f(x - \varepsilon y) - f(x)\| dy = C \int_{x+\varepsilon B_0} \|f(z) - f(x)\| dz \rightarrow 0,$$

as $x + \varepsilon B_0$ is precisely a family of balls as in the Lebesgue differentiation theorem.

For case (2), we argue as follows. Since $\Phi(z)$ depends only on $|z|$ of which it is a decreasing function, we may express it as

$$\Phi(z) = \int_{[|z|, \infty)} d\mu(t)$$

for some positive measure μ on $\mathbb{R}_+$. Then

$$\begin{aligned} F_\varepsilon &\leq \int_{\mathbb{R}^d} \Phi(y) \|f(x - \varepsilon y) - f(x)\| dy \\ &= \int_{\mathbb{R}_+} \left(\int_{|y| \leq t} \|f(x - \varepsilon y) - f(x)\| dy \right) d\mu(t) \\ &= \int_{\mathbb{R}_+} |B(0, t)| \left(\int_{B(x, \varepsilon t)} \|f(z) - f(x)\| dy \right) d\mu(t). \end{aligned} \quad (2.12)$$

The term in parentheses is bounded by $Mf(x) + \|f(x)\|$ uniformly in ε and t , and it converges to zero as $\varepsilon \rightarrow 0$, for every fixed t . Thus $F_\varepsilon \rightarrow 0$ follows by dominated convergence, observing that

$$\begin{aligned} \int_{\mathbb{R}_+} |B(0, t)| d\mu(t) &= \int_{\mathbb{R}_+} \int_{|z| \leq t} dz d\mu(t) \\ &= \int_{\mathbb{R}^d} \int_{[|z|, \infty)} d\mu(t) dz = \int_{\mathbb{R}^d} \Phi(z) dz < \infty \end{aligned} \quad (2.13)$$

by assumption.

To prove the final claim of the theorem, let $f \in L^p(\mathbb{R}^d; X)$ for some $p \in [1, \infty]$, and x be a Lebesgue point of f . By the definition of a Lebesgue point, there exists an $\varepsilon > 0$ such that

$$\int_B \|f(y) - f(x)\| dy \leq 1$$

whenever $B \ni x$ satisfies $|B| < \varepsilon$, and hence $\int_B \|f\| dx \leq \|f(x)\| + 1$ for all such balls. On the other hand, if $|B| \geq \varepsilon$, then

$$\int_B \|f(y)\| dy \leq |B|^{-1/p} \|f\|_p \leq \varepsilon^{-1/p} \|f\|_p.$$

It follows that $Mf(x) \leq \max(\|f(x)\| + 1, \varepsilon^{-1/p} \|f\|_p) < \infty$, and the proof is complete. $\square$

From the previous proof we can also extract the following useful estimate:

Proposition 2.3.9. *For $f \in L^1_{\text{loc}}(\mathbb{R}^d; X)$ and $\phi \in L^1(\mathbb{R}^d)$ with decreasing radial majorant Φ as in (2.11), we have*

$$\sup_{\varepsilon > 0} \|\phi_\varepsilon * f(x)\| \leq \|\Phi\|_1 Mf(x) \quad \forall x \in \mathbb{R}^d.$$

Proof. Since we may redefine $f(x) = 0$ at the single point x without altering either the left or the right side of the claim, the bound is an immediate consequence of (2.12) and (2.13) (note that both estimates do not depend on x being a Lebesgue point or not). $\square$

2.4 The Fourier transform

This section takes up the study of the vector-valued Fourier transform. This is an entire subject in itself, and the present presentation is limited to those elements of the theory that are needed in later chapters. In particular, we limit ourselves to the case of $\mathbb{R}^d$; much of what we will have to say in the next three subsections has analogous formulations for the torus $\mathbb{T}^d$.

As the Fourier transform involves the multiplication of vectors in X with complex exponentials, throughout this section we shall assume that all Banach spaces are complex.

Definition 2.4.1 (Fourier transform). *The Fourier transform is the operator $\mathcal{F} : L^1(\mathbb{R}^d; X) \rightarrow L^\infty(\mathbb{R}^d; X)$, $f \mapsto \widehat{f}$, defined by*

$$\mathcal{F}f(\xi) := \widehat{f}(\xi) := \int_{\mathbb{R}^d} f(x) e^{-2\pi i x \cdot \xi} dx, \quad \xi \in \mathbb{R}^d.$$

The inverse Fourier transform is the operator $\mathcal{F}^{-1} : L^1(\mathbb{R}^d; X) \rightarrow L^\infty(\mathbb{R}^d; X)$, $f \mapsto \check{f}$, defined by

$$\mathcal{F}^{-1}f(x) := \check{f}(x) := \int_{\mathbb{R}^d} f(\xi) e^{2\pi i x \cdot \xi} d\xi, \quad x \in \mathbb{R}^d.$$

The latter is simply a name for the moment, but will receive justification shortly. Both the Fourier transform and the inverse Fourier transform are the tensor extension of the corresponding operators on $L^1(\mathbb{R}^d)$, which is seen from the fact that on the dense subspace $L^1(\mathbb{R}^d) \otimes X$ of $L^1(\mathbb{R}^d; X)$ one has

$$\left(\sum_{n=1}^N g_n \otimes x_n \right)^\wedge = \sum_{n=1}^N \widehat{g}_n \otimes x_n,$$

and similarly for the inverse Fourier transform.

Writing

$$\widetilde{f}(\cdot) := f(-\cdot)$$

for the *reflection* of a function f , it is immediate that

$$\check{f} = (\widehat{f})^\sim = (\widetilde{f})^\wedge, \quad \widehat{f} = (\check{f})^\sim = (\widetilde{\check{f}})^\wedge, \quad (2.14)$$

from which it follows that “any” reasonable property of the transform $f \mapsto \widehat{f}$ has a counterpart for $f \mapsto \check{f}$.

Remark 2.4.2. Observe the use of the factor 2π in the exponential functions in (2.14). This is the harmonic analyst's convention, which has the advantage that no powers of 2π appear as pre-factors in the inversion formula.

As we shall see, as a rule L^1 -results for the scalar Fourier transform extend to the vector-valued setting, but L^2 -results do not, unless X is a Hilbert space. An example of the former is the Riemann-Lebesgue lemma, which asserts that the Fourier transform maps $L^1(\mathbb{R}^d; X)$ into $C_0(\mathbb{R}^d; X)$, the space of continuous functions $f : \mathbb{R}^d \rightarrow X$ which satisfy $\lim_{|x| \rightarrow \infty} f(x) = 0$.

Lemma 2.4.3 (Riemann–Lebesgue). *Let X be any Banach space. For all $f \in L^1(\mathbb{R}^d; X)$ we have $\widehat{f} \in C_0(\mathbb{R}^d; X)$ and $\|\widehat{f}\|_\infty \leq \|f\|_1$.*

Proof. The estimate is immediate by moving the norm inside the integral. To see that $\widehat{f}$ is continuous, it suffices to observe that $\xi \mapsto e^{-2\pi i x \cdot \xi}$ is continuous for every x , and apply dominated convergence.

The fact that $\widehat{f}(\xi) \rightarrow 0$ as $|\xi| \rightarrow \infty$ requires a simple density argument. Since $\|\widehat{f}(\xi) - \widehat{g}(\xi)\| \leq \|f - g\|_1$, it is enough to show the convergence to zero for all g in a dense subspace $E \subseteq L^1(\mathbb{R}^d; X)$. A cheap choice would be to take $E = L^1(\mathbb{R}^d) \otimes X$, and resort to the classical Riemann–Lebesgue lemma for scalar-valued functions. Alternatively, we may repeat the classical proof, say with $E = C_c^1(\mathbb{R}^d; X)$. Writing $\partial_j := \partial/\partial x_j$ for the partial derivative with respect to the j th coordinate we have

$$2\pi i \xi_j \widehat{f}(\xi) = - \int_{\mathbb{R}^d} \partial_j e^{-2\pi i x \cdot \xi} f(x) dx = \int_{\mathbb{R}^d} e^{2\pi i x \cdot \xi} \partial_j f(x) dx = \widehat{\partial_j f}(\xi)$$

after an integration by parts, and thus

$$\|\widehat{f}(\xi)\| \leq \frac{1}{|\xi|} \sum_{j=1}^d \|\xi_j \widehat{f}(\xi)\| \leq \frac{1}{2\pi|\xi|} \sum_{j=1}^d \|\widehat{\partial_j f}\|_\infty \leq \frac{1}{2\pi|\xi|} \sum_{j=1}^d \|\partial_j f\|_1 \rightarrow 0.$$

□

2.4.a The inversion formula and Plancherel's theorem

To proceed any further we need the following classical identity.

Lemma 2.4.4. *The Gaussian function $g(x) = \exp(-\pi|x|^2)$ satisfies $\widehat{g} = g$.*

Proof. First let $d = 1$. Completing squares and using Cauchy's theorem to shift the path of integration, we find

$$\begin{aligned} \int_{\mathbb{R}} e^{-2\pi i x \xi} \exp(-\pi x^2) dx &= \int_{\mathbb{R}} \exp(-\pi[(x - i\xi)^2 + \xi^2]) dx \\ &= \exp(-\pi\xi^2) \int_{\mathbb{R} - i\xi} \exp(-\pi z^2) dz \end{aligned}$$

$$\begin{aligned}
&= \exp(-\pi\xi^2) \int_{\mathbb{R}} \exp(-\pi z^2) dz \\
&= \exp(-\pi\xi^2).
\end{aligned}$$

The general case follows from this via separation of variables. $\square$

Proposition 2.4.5 (Fourier inversion). *Suppose that the Fourier transform of a function $f \in L^1(\mathbb{R}^d; X)$ satisfies $\widehat{f} \in L^1(\mathbb{R}^d; X)$. Then $f = (\widehat{f})^\sim$ at all Lebesgue points of f .*

Proof. By Lemma 2.4.4, the Gaussian function $g(x) = e^{-\pi|x|^2}$ satisfies

$$g(x) = \int_{\mathbb{R}^d} g(\xi) e^{-2\pi i x \cdot \xi} d\xi = \int_{\mathbb{R}^d} g(\xi) e^{2\pi i x \cdot \xi} d\xi,$$

where the second identity uses that g is real-valued, so that taking complex conjugates leaves the expression unchanged. Substituting x/ε for x and changing integration variables, we obtain

$$g_\varepsilon(x) = \frac{1}{\varepsilon^d} g\left(\frac{x}{\varepsilon}\right) = \int_{\mathbb{R}^d} g(\varepsilon\xi) e^{2\pi i x \cdot \xi} d\xi.$$

Fix a Lebesgue point $x \in \mathbb{R}^d$ of f . By Theorem 2.3.8 we have $g_\varepsilon * f(x) \rightarrow f(x)$ as $\varepsilon \downarrow 0$. Using the above, it follows that

$$\begin{aligned}
f(x) &= \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^d} g_\varepsilon(y) f(x-y) dy \\
&= \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^d} \left[\int_{\mathbb{R}^d} g(\varepsilon\xi) e^{2\pi i y \cdot \xi} d\xi \right] f(x-y) dy \\
&= \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^d} g(\varepsilon\xi) \left[\int_{\mathbb{R}^d} e^{-2\pi i (x-y) \cdot \xi} f(x-y) dy \right] e^{2\pi i x \cdot \xi} d\xi \\
&= \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^d} g(\varepsilon\xi) \widehat{f}(\xi) e^{2\pi i x \cdot \xi} d\xi \\
&= \int_{\mathbb{R}^d} \widehat{f}(\xi) e^{2\pi i x \cdot \xi} d\xi = (\widehat{f})^\sim(x),
\end{aligned}$$

where the penultimate equality is justified by dominated convergence, which can be used here since $\widehat{f}$ is integrable and $g(\varepsilon\xi) \rightarrow g(0) = 1$ as $\varepsilon \downarrow 0$. We thus have shown that $f(x) = (\widehat{f})^\sim(x)$. $\square$

The weaker statement that $f = (\widehat{f})^\sim$ holds almost everywhere can be obtained without reference to Theorem 2.3.8. Instead, it suffices to apply Proposition 1.2.32 and pass to an almost everywhere convergent subsequence $g_{\varepsilon_j} * f$.

The Plancherel theorem

It will be useful to introduce the class

$$\check{L}^1(\mathbb{R}^d; X) := \{g \in L^\infty(\mathbb{R}^d; X) : g = \check{f} \text{ for some } f \in L^1(\mathbb{R}^d; X)\}$$

of all functions obtained as inverse Fourier transforms of X -valued L^1 -functions, the equality in the definition being immediate from (2.14).

The Fourier transform (or rather, its restriction to $L^1(\mathbb{R}^d; X) \cap \check{L}^1(\mathbb{R}^d; X)$) extends in a natural way to a bounded operator $\mathcal{F} : \check{L}^1(\mathbb{R}^d; X) \rightarrow L^1(\mathbb{R}^d; X)$ by defining, when $f = \check{g}$ for some $g \in L^1(\mathbb{R}^d; X)$,

$$\mathcal{F}f := \hat{f} := g.$$

Lemma 2.4.6. *The Fourier transform on $\check{L}^1(\mathbb{R}^d; X)$ is well defined, and for functions in $L^1(\mathbb{R}^d; X) \cap \check{L}^1(\mathbb{R}^d; X)$ it agrees with the Fourier transform on $L^1(\mathbb{R}^d; X)$.*

Proof. If $f = \check{h} = \check{g}$ with $h, g \in L^1(\mathbb{R}^d; X)$, then $k := g - h$ satisfies $k \in L^1(\mathbb{R}^d; X)$ and $\check{k} = \check{g} - \check{h} = 0 \in L^1(\mathbb{R}^d; X)$. By Proposition 2.4.5, $k = (\check{k})^\wedge = (0)^\wedge = 0$, so $g = h$.

If $f \in L^1(\mathbb{R}^d; X) \cap \check{L}^1(\mathbb{R}^d; X)$ satisfies $f = \check{h}$ with $h \in L^1(\mathbb{R}^d; X)$, then Proposition 2.4.5 guarantees that $h = (\check{h})^\wedge = \hat{f}$, so that the new definition agrees with the earlier one. $\square$

In questions of Fourier analysis in $L^p(\mathbb{R}^d; X)$ it is convenient to work with the space $L^p(\mathbb{R}^d; X) \cap \check{L}^1(\mathbb{R}^d; X)$, on which the Fourier transform is well defined. The next lemma provides the key density results needed to make this idea work.

Lemma 2.4.7. *For all $p \in [1, \infty)$,*

- (1) $L^p(\mathbb{R}; X) \cap \check{L}^1(\mathbb{R}; X)$ is dense in $L^p(\mathbb{R}; X)$.
- (2) $L^1(\mathbb{R}^d; X) \cap \check{L}^1(\mathbb{R}; X)$ is dense in $L^p(\mathbb{R}; X) \cap L^1(\mathbb{R}^d; X)$.

Proof. It suffices to prove the second assertion. The first assertion can be deduced from it, using that $L^p(\mathbb{R}^d; X) \cap L^1(\mathbb{R}^d; X)$ is dense in $L^p(\mathbb{R}^d; X)$.

To prove the second assertion, note that $\check{L}^1(\mathbb{R}^d; X) \subseteq L^\infty(\mathbb{R}^d; X)$ and therefore, $L^1(\mathbb{R}^d; X) \cap \check{L}^1(\mathbb{R}; X) \subseteq L^p(\mathbb{R}^d; X)$. Given $f \in L^p(\mathbb{R}^d; X) \cap L^1(\mathbb{R}^d; X)$, consider the convolutions $g_\varepsilon * f \in L^p(\mathbb{R}^d; X) \cap L^1(\mathbb{R}^d; X)$, where $g(x) = e^{-\pi|x|^2}$. By the second case of the next lemma, $(g_\varepsilon * f)^\wedge = g(\varepsilon \cdot) \hat{f}$ belongs to $L^1(\mathbb{R}^d; X)$ since it is a product of functions in $L^1(\mathbb{R}^d)$ and $L^\infty(\mathbb{R}^d; X)$, and therefore $g_\varepsilon * f \in \check{L}^1(\mathbb{R}^d; X)$. Since $g_\varepsilon * f \rightarrow f$ in $L^p(\mathbb{R}^d; X) \cap L^1(\mathbb{R}^d; X)$ as $\varepsilon \downarrow 0$ as a corollary to Proposition 1.2.32, we obtain the desired density. $\square$

Lemma 2.4.8. *Suppose that either one of the following two conditions is satisfied:*

- (i) $f \in \check{L}^1(\mathbb{R}^d; X)$ and $g \in L^1(\mathbb{R}^d)$;
- (ii) $f \in L^1(\mathbb{R}^d; X)$ and $g \in \check{L}^1(\mathbb{R}^d)$.

*Then $g * f \in \check{L}^1(\mathbb{R}^d; X)$ and $\widehat{g * f} = \widehat{g}\widehat{f}$.*

Proof. We prove the lemma assuming (i), the proof in case (ii) being similar. We have

$$\begin{aligned} g * f(x) &= \int_{\mathbb{R}^d} g(x-y) \int_{\mathbb{R}^d} \widehat{f}(\xi) e^{2\pi i y \xi} d\xi dy \\ &= \int_{\mathbb{R}^d} \left[\int_{\mathbb{R}^d} g(x-y) e^{2\pi i(y-x)\xi} dy \right] \widehat{f}(\xi) e^{2\pi i x \xi} d\xi \\ &= \int_{\mathbb{R}^d} \widehat{g}(\xi) \widehat{f}(\xi) e^{2\pi i x \xi} d\xi = (\widehat{g}\widehat{f})^\sim(x), \end{aligned}$$

where $\widehat{g}\widehat{f}$ belongs to $L^1(\mathbb{R}^d; X)$ since it is a product of functions in $L^\infty(\mathbb{R}^d)$ and $L^1(\mathbb{R}^d; X)$. It follows that $g * f$ belongs to $\check{L}^1(\mathbb{R}^d; X)$ and $\widehat{g * f} = \widehat{g}\widehat{f}$. $\square$

We are now ready to prove:

Theorem 2.4.9 (Plancherel theorem). *Let H be a Hilbert space. If $f \in L^2(\mathbb{R}^d; H) \cap L^1(\mathbb{R}^d; H)$, then $\widehat{f} \in L^2(\mathbb{R}^d; H)$ and*

$$\|\widehat{f}\|_2 = \|f\|_2. \quad (2.15)$$

In particular, the Fourier transform (or to be more precise, its restriction to $L^2(\mathbb{R}^d; H) \cap L^1(\mathbb{R}^d; H)$) extends to an isometry on $L^2(\mathbb{R}^d; H)$.

The resulting isometry $\mathcal{F}$ on $L^2(\mathbb{R}^d; H)$ is usually referred to as the *Fourier–Plancherel transform*.

Proof. We denote the inner product of H by $(\cdot|\cdot)$. By the Fubini theorem, for any two functions $f, g \in L^1(\mathbb{R}^d; H)$ we have

$$\int_{\mathbb{R}^d} (\widehat{f}(x)|g(x)) dx = \int_{\mathbb{R}^d} (f(x)|\check{g}(x)) dx.$$

For functions $f \in L^1(\mathbb{R}^d; H) \cap \check{L}^1(\mathbb{R}^d; H)$, (2.15) follows by taking $g = \widehat{f}$ in the above identity and applying the Fourier inversion theorem. Since $L^1(\mathbb{R}^d; H) \cap \check{L}^1(\mathbb{R}^d; H)$ is dense in $L^1(\mathbb{R}^d; H) \cap L^2(\mathbb{R}^d; H)$ by Lemma 2.4.7, an approximation argument completes the proof. $\square$

It has already been observed (in Theorem 2.1.18) that this theorem fails if one replaces H by an any non-Hilbertian Banach space. Let us also add the observation that the theorem could have alternatively been derived from its scalar-valued counterpart by invoking the extension theorem of Paley and Marcinkiewicz–Zygmund (Theorem 2.1.9).

By complex interpolation of the Fourier transform, viewed as a bounded operator on norm at most one from $L^1(\mathbb{R}^d; H)$ to $L^\infty(\mathbb{R}^d; H)$ and from $L^2(\mathbb{R}^d; H)$ to $L^2(\mathbb{R}^d; H)$ we obtain, via Theorem 2.2.1 or Theorems 2.2.6 and C.2.6:

Corollary 2.4.10 (Hausdorff–Young theorem). *For any $p \in [1, 2]$ the Fourier transform is bounded, of norm at most one, as an operator from $L^p(\mathbb{R}^d; H)$ to $L^{p'}(\mathbb{R}^d; H)$, $\frac{1}{p} + \frac{1}{p'} = 1$.*

2.4.b Fourier type

It has already been observed in Example 2.1.15 that the vector-valued Hausdorff–Young theorem fails for certain Banach spaces. Given an exponent $p \in (1, 2]$, this prompts the question for which Banach spaces X the Fourier transform, viewed as an operator from $L^p(\mathbb{R}^d)$ to $L^{p'}(\mathbb{R}^d)$, extends to a bounded operator from $L^p(\mathbb{R}^d; X)$ to $L^{p'}(\mathbb{R}^d; X)$. This question will be addressed presently. The results of this subsection will not be needed in the rest of this volume, but some applications will be presented in Volume II where we also discuss the relationship with the probabilistic notion of type.

We begin by showing that the question posed above is independent of the dimension d . As before we assume all Banach spaces to be complex.

Proposition 2.4.11. *Let X be a Banach space, fix $p \in [1, 2]$, and let $\frac{1}{p} + \frac{1}{p'} = 1$. The following assertions are equivalent:*

- (1) *for some $d \geq 1$, $\mathcal{F}$ extends to a bounded operator from $L^p(\mathbb{R}^d; X)$ to $L^{p'}(\mathbb{R}^d; X)$;*
- (2) *for all $d \geq 1$, $\mathcal{F}$ extends to a bounded operator from $L^p(\mathbb{R}^d; X)$ to $L^{p'}(\mathbb{R}^d; X)$.*

Denoting the norms of these extensions by $\varphi_{p,X}(\mathbb{R}^d)$, we have

$$\left(\frac{p^{1/p}}{(p)^{1/p'}} \right)^{(d-1)/2} \varphi_{p,X}(\mathbb{R}) \leq \varphi_{p,X}(\mathbb{R}^d) \leq (\varphi_{p,X}(\mathbb{R}))^d. \quad (2.16)$$

Proof. Suppose that $\mathcal{F}$ extends to bounded operator from $L^p(\mathbb{R}^d; X)$ to $L^{p'}(\mathbb{R}^d; X)$ for some $d \geq 1$.

By Lemma 2.4.4, the function $g : \mathbb{R}^{d-1} \rightarrow \mathbb{R}$ given by $g(s) := e^{-\pi|s|^2}$ satisfies $\mathcal{F}_{d-1}g = g$. Let $f \in L^p(\mathbb{R}; X)$ be arbitrary. Defining $F : \mathbb{R}^d \rightarrow X$ by $F(x, y) := f(x)g(y)$, where $x \in \mathbb{R}$ and $y \in \mathbb{R}^{d-1}$, we obtain

$$\|g\|_{L^{p'}(\mathbb{R}^{d-1})} \|\widehat{f}\|_{L^{p'}(\mathbb{R}; X)} = \|\widehat{g}\|_{L^{p'}(\mathbb{R}^{d-1})} \|\widehat{f}\|_{L^{p'}(\mathbb{R}; X)}$$

$$\begin{aligned}
&= \|\widehat{F}\|_{L^{p'}(\mathbb{R}^d; X)} \\
&\leq \varphi_{p, X}(\mathbb{R}^d) \|F\|_{L^p(\mathbb{R}^d; X)} \\
&= \varphi_{p, X}(\mathbb{R}^d) \|g\|_{L^p(\mathbb{R}^{d-1})} \|f\|_{L^p(\mathbb{R}; X)}.
\end{aligned}$$

Therefore, $\mathcal{F}$ is bounded from $L^p(\mathbb{R}; X)$ into $L^{p'}(\mathbb{R}; X)$. Moreover, since $\|g\|_{L^r(\mathbb{R}^{d-1})} = r^{-(d-1)/2r}$ for all $r \in [1, \infty)$, the left-hand side estimate in (2.16) follows.

To finish the proof, by iteration it suffices to show that $\varphi_{p, X}(\mathbb{R}^{d+1}) \leq \varphi_{p, X}(\mathbb{R}^d) \varphi_{p, X}(\mathbb{R})$. Fix $f \in L^p(\mathbb{R}^{d+1}; X)$. Let $\mathcal{F}_1$ denote the Fourier transform in the first coordinate and $\mathcal{F}_d$ the Fourier transform in the last d coordinates. By Minkowski's inequality (Proposition 1.2.22),

$$\begin{aligned}
\|\mathcal{F}f\|_{L^{p'}(\mathbb{R}^{d+1}; X)} &= \left(\int_{\mathbb{R}^d} \int_{\mathbb{R}} \|\mathcal{F}_1(\mathcal{F}_d(f)(\cdot, \eta))(\xi)\|^{p'} d\xi d\eta \right)^{1/p'} \\
&\leq \varphi_{p, X}(\mathbb{R}) \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}} \|\mathcal{F}_d(f)(x, \eta)\|^p dx \right)^{p'/p} d\eta \right)^{1/p'} \\
&\leq \varphi_{p, X}(\mathbb{R}) \left(\int_{\mathbb{R}} \left(\int_{\mathbb{R}^d} \|\mathcal{F}_d(f)(x, \eta)\|^{p'} d\eta \right)^{p/p'} dx \right)^{1/p} \\
&\leq \varphi_{p, X}(\mathbb{R}) \varphi_{p, X}(\mathbb{R}^d) \left(\int_{\mathbb{R}} \int_{\mathbb{R}^d} \|f(x, y)\|^p dy dx \right)^{1/p} \\
&= \varphi_{p, X}(\mathbb{R}) \varphi_{p, X}(\mathbb{R}^d) \|f\|_{L^p(\mathbb{R}^{d+1}; X)}.
\end{aligned}$$

This also proves the right-hand side estimate in (2.16). $\square$

Definition 2.4.12 (Fourier type). Let $p \in [1, 2]$. A Banach space X is said to have Fourier type p if the Fourier transform, as an operator from $L^p(\mathbb{R})$ to $L^{p'}(\mathbb{R})$, extends to a bounded operator from $L^p(\mathbb{R}; X)$ to $L^{p'}(\mathbb{R}; X)$.

The reason for considering exponents $p \in [1, 2]$ is that even in the scalar-valued case the Fourier transform fails to be bounded from $L^p(\mathbb{R})$ to $L^{p'}(\mathbb{R})$ for the remaining values $p \in (2, \infty]$. If X has Fourier type p , then by complex interpolation (see Theorem 2.2.6) X has Fourier type q for all $q \in [1, p]$, with $\varphi_{q, X}(\mathbb{R}^d) \leq (\varphi_{p, X}(\mathbb{R}^d))^\theta$, where $\theta = p'/q'$.

Example 2.4.13. Every Banach space X has Fourier type 1, with $\varphi_{1, X}(\mathbb{R}^d) = 1$. The Plancherel theorem implies that every Hilbert space H has Fourier type 2, with $\varphi_{2, H}(\mathbb{R}^d) = 1$, and in the converse direction any Banach space with Fourier type 2 is isomorphic to a Hilbert space by Kwapien's theorem (see Theorem 2.1.18).

Example 2.4.14. Let X be a Banach space, $(S, \mathcal{A}, \mu)$ a measure space, and let $r \in [1, \infty)$. If X has Fourier type p , then $L^r(S; X)$ has Fourier type $p \wedge r \wedge r'$, and

$$\varphi_{p \wedge r \wedge r', L^r(S; X)}(\mathbb{R}^d) \leq \varphi_{p \wedge r \wedge r', X}(\mathbb{R}^d).$$

To see this, first let $r \in [p, p']$. Then $r' \in [p, p']$ and $p \wedge r \wedge r' = p$. For any simple function $f : \mathbb{R}^d \rightarrow L^r(S; X)$, Minkowski's inequality (Proposition 1.2.22, applied first with $r \leq p'$ and then with $p \leq r$) implies

$$\begin{aligned} \|\widehat{f}\|_{L^{p'}(\mathbb{R}^d; L^r(S; X))} &\leq \|\widehat{f}\|_{L^r(S; L^{p'}(\mathbb{R}^d; X))} \\ &\leq \varphi_{p, X}(\mathbb{R}^d) \|f\|_{L^r(S; L^p(\mathbb{R}^d; X))} \leq \varphi_{p, X}(\mathbb{R}^d) \|f\|_{L^p(\mathbb{R}^d; L^r(S; X))} \end{aligned}$$

and the result follows.

Next consider the case $r \notin [p, p']$. In this situation $r \wedge r' < p$ and thus X has Fourier type $r \wedge r'$ (by the previous example). We can now apply the previous case with $p = r \wedge r'$.

Specialising to the scalar-valued setting we obtain:

Example 2.4.15. For all $r \in [1, \infty)$ the space $L^r(S)$ has Fourier type $r \wedge r'$, with constant $\varphi_{r \wedge r', L^r(S)}(\mathbb{R}^d) = \varphi_{r \wedge r', \mathbb{C}}(\mathbb{R}^d) \leq 1$.

The result of this example is optimal, in the sense that if $\dim L^r(S) = \infty$, then $L^r(S)$ does not have Fourier type q for any $q > r \wedge r'$; this is proved in the same way as in Example 2.1.15.

Proposition 2.4.16 (Duality). *Let X be a Banach space and let $p \in [1, 2]$. The space X has Fourier type p if and only if X^* has Fourier type p , and for all $d \geq 1$ we have*

$$\varphi_{p, X}(\mathbb{R}^d) = \varphi_{p, X^*}(\mathbb{R}^d).$$

Proof. Suppose that X has Fourier type p and let $f : \mathbb{R}^d \rightarrow X$ and $g : \mathbb{R}^d \rightarrow X^*$ be simple functions. By Fubini's theorem,

$$\int_{\mathbb{R}^d} \langle f(x), \widehat{g}(x) \rangle dx = \int_{\mathbb{R}^d} \langle \widehat{f}(\xi), g(\xi) \rangle d\xi.$$

Hence,

$$\begin{aligned} \left| \int_{\mathbb{R}^d} \langle f(x), \widehat{g}(x) \rangle dx \right| &\leq \|\widehat{f}\|_{L^{p'}(\mathbb{R}^d; X)} \|g\|_{L^p(\mathbb{R}^d; X^*)} \\ &\leq \varphi_{p, X}(\mathbb{R}^d) \|f\|_{L^p(\mathbb{R}^d; X)} \|g\|_{L^p(\mathbb{R}^d; X^*)}. \end{aligned}$$

Taking the supremum over all simple functions f with $\|f\|_{L^p(\mathbb{R}^d; X)} \leq 1$, it follows that

$$\|\widehat{g}\|_{L^{p'}(\mathbb{R}^d; X^*)} \leq \varphi_{p, X}(\mathbb{R}^d) \|g\|_{L^p(\mathbb{R}^d; X^*)}.$$

We conclude that X^* has Fourier type p and that $\varphi_{p, X^*}(\mathbb{R}^d) \leq \varphi_{p, X}(\mathbb{R}^d)$.

If X^* has Fourier type p , then by the above the bi-dual X^{**} has Fourier type p and $\varphi_{p, X^{**}}(\mathbb{R}^d) \leq \varphi_{p, X^*}(\mathbb{R}^d)$. Since X^{**} contains X isometrically, the result follows. $\square$

In order to be able to present some further examples, the following result is useful.

Proposition 2.4.17 (Interpolation). *Let (X_0, X_1) be an interpolation couple of Banach spaces and suppose that X_0 and X_1 have Fourier type $p_0, p_1 \in [1, 2]$, respectively. For all $\theta \in (0, 1)$ the complex interpolation space $X_\theta = [X_0, X_1]_\theta$ and the real interpolation space $(X_0, X_1)_{\theta, p_0, p_1}$ have Fourier type p_θ , where $\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$, and we have*

$$\begin{aligned}\varphi_{p, X_\theta}(\mathbb{R}^d) &\leq (\varphi_{p_0, X_0}(\mathbb{R}^d))^{1-\theta} (\varphi_{p_1, X_1}(\mathbb{R}^d))^\theta, \\ \varphi_{p, (X_0, X_1)_{\theta, p_0, p_1}} &\leq (\varphi_{p_0, X_0}(\mathbb{R}^d))^{1-\theta} (\varphi_{p_1, X_1}(\mathbb{R}^d))^\theta.\end{aligned}$$

Recall that $(X_0, X_1)_{\theta, p_0, p_1} = (X_0, X_1)_{\theta, p_\theta}$ with equivalent norms (see Theorem C.3.14).

Proof. For X_θ the result follows from Theorem 2.2.6:

$$\begin{aligned}[L^{p_0}(\mathbb{R}^d; X_0), L^{p_1}(\mathbb{R}^d; X_1)]_\theta &= L^p(\mathbb{R}^d; X_\theta) \\ [L^{p'_0}(\mathbb{R}^d; X_0), L^{p'_1}(\mathbb{R}^d; X_1)]_\theta &= L^{p'}(\mathbb{R}^d; X_\theta),\end{aligned}$$

where $\frac{1}{p'} = \frac{1-\theta}{p'_0} + \frac{\theta}{p'_1} = 1 - \frac{1}{p}$. The result for $(X_0, X_1)_{\theta, p_0, p_1}$ is proved in the same way, this time using Theorem 2.2.10. $\square$

Example 2.4.18. For $p \in (1, \infty)$, the Schatten class $\mathcal{C}^p$ has Fourier type $p \wedge p'$ with $\varphi_{p, \mathcal{C}^p}(\mathbb{R}^d) \leq 1$. Indeed, since $\mathcal{C}^2$ is a Hilbert space it has Fourier type 2, with constant 1. Moreover, $\mathcal{C}^1$ and $\mathcal{C}^\infty$ have Fourier type 1, with constant 1. Therefore, the result from Proposition 2.4.17 and the fact that $\mathcal{C}^p = [\mathcal{C}^1, \mathcal{C}^2]_{2/p'}$ if $p \in (1, 2)$ and $\mathcal{C}^p = [\mathcal{C}^\infty, \mathcal{C}^2]_{2/p}$ if $p \in (2, \infty)$ (see Appendix D).

Our final aim is to show that the notion of Fourier type could equivalently be defined in terms of the Fourier transform on the torus $\mathbb{T}^d = [0, 1]^d$,

$$\mathcal{F}f(n) := \int_{\mathbb{T}^d} f(t) e^{-2\pi i n \cdot t} dt, \quad n \in \mathbb{Z}^d.$$

As in the continuous case, $\mathcal{F}$ is bounded as an operator from $L^1(\mathbb{T}^d)$ to $\ell^\infty(\mathbb{Z}^d)$ and as an operator from $L^2(\mathbb{T}^d)$ to $\ell^2(\mathbb{Z}^d)$, and therefore by interpolation $\mathcal{F}$ is bounded as an operator from $L^p(\mathbb{T}^d)$ to $\ell^{p'}(\mathbb{Z}^d)$ for all $p \in [1, 2]$ and $\frac{1}{p} + \frac{1}{p'} = 1$. Our aim is to prove that this operator extends to a bounded operator from $L^p(\mathbb{T}^d; X)$ to $\ell^{p'}(\mathbb{Z}^d; X)$ if and only if X has Fourier type p and to provide two-sided bounds on the operator norm in that case. We begin with some preliminary observations.

For $2 \leq q < \infty$ we define the constant C_q as the global minimum of the 1-periodic function

$$x \mapsto \sum_{m \in \mathbb{Z}} \left| \frac{\sin(\pi(x+m))}{\pi(x+m)} \right|^q, \quad x \in \mathbb{R}. \quad (2.17)$$

It can be shown that this minimum is taken in the points $\frac{1}{2} + \mathbb{Z}$, so that

$$C_q = \frac{1}{\pi^q} \sum_{m \in \mathbb{Z}} \frac{1}{(\frac{1}{2} + m)^q},$$

but we are not aware of a simple direct proof of this fact. Here we content ourselves with the observation that

$$(2/\pi)^q \leq C_q \leq 1$$

and defer a further discussion to the Notes at the end of the chapter. To prove the bounds for C_q , first note that $C_q \leq C_2$ by monotonicity. Furthermore,

$$C_2 \leq \frac{1}{\pi^2} \sum_{m \in \mathbb{Z}} \frac{1}{(m + \frac{1}{2})^2} = 1$$

and, by 1-periodicity,

$$C_q = \min_{x \in [-\frac{1}{2}, \frac{1}{2}]} \sum_{m \in \mathbb{Z}} \left| \frac{\sin(\pi(x+m))}{\pi(x+m)} \right|^q \geq \min_{x \in [-\frac{1}{2}, \frac{1}{2}]} \left| \frac{\sin(\pi x)}{\pi x} \right|^q = \frac{1}{(\pi/2)^q}.$$

For $k \in \mathbb{Z}^d$ we consider the *trigonometric functions* $e_k : \mathbb{T}^d \rightarrow \mathbb{C}$,

$$e_k(t) := e^{2\pi i k \cdot t}, \quad t \in \mathbb{T}^d,$$

and write $|k| := \sum_{i=1}^d |k_i|$.

Lemma 2.4.19. *Let X be a Banach space and let $p \in (1, 2]$ and $p' \in [2, \infty)$ satisfy $\frac{1}{p} + \frac{1}{p'} = 1$. For fixed $a > 0$, $n \in \mathbb{N}$, and vectors $x_k \in X$, $|k| \leq n$, define $f : \mathbb{R}^d \rightarrow X$ by*

$$f(t) := a^{-d/p} \sum_{|k| \leq n} \mathbf{1}_{[0,1]^d - k}(a^{-1}t) x_k.$$

Then

$$C_{p'}^{d/p'} \left\| \sum_{|k| \leq n} e_k x_k \right\|_{L^{p'}(\mathbb{T}^d; X)} \leq \|\mathcal{F}f\|_{L^{p'}(\mathbb{R}^d; X)} \leq \left\| \sum_{|k| \leq n} e_k x_k \right\|_{L^{p'}(\mathbb{T}^d; X)}.$$

Proof. Writing $\text{sinc}(x) = \sin(x)/x$, we have

$$g(\xi) := \mathcal{F}(\mathbf{1}_{[0,1]^d})(\xi) = e^{-i(\pi\xi_1 + \dots + \pi\xi_d)} \prod_{n=1}^d \text{sinc}(\pi\xi_n).$$

Therefore $\mathcal{F}f(\xi) = a^{d/p'} g(a\xi) \sum_{|k| \leq n} e_k(a\xi) x_k$, and a change of variables gives

$$\begin{aligned}
\|\mathcal{F}f\|_{L^{p'}(\mathbb{R}^d; X)}^{p'} &= a^d \int_{\mathbb{R}^d} |g(a\xi)|^{p'} \left\| \sum_{|k| \leq n} e_k(a\xi) x_k \right\|^{p'} d\xi \\
&= \int_{\mathbb{R}^d} |g(\xi)|^{p'} \left\| \sum_{|k| \leq n} e_k(\xi) x_k \right\|^{p'} d\xi \\
&= \sum_{m \in \mathbb{Z}^d} \int_{[0,1]^d - m} |g(\xi)|^{p'} \left\| \sum_{|k| \leq n} e_k(\xi) x_k \right\|^{p'} d\xi \\
&= \int_{[0,1]^d} h_{p'}(\xi) \left\| \sum_{|k| \leq n} e_k(\xi) x_k \right\|^{p'} d\xi,
\end{aligned}$$

where, for every $r \geq 2$,

$$\begin{aligned}
h_r(\xi) &= \sum_{m \in \mathbb{Z}^d} |g(\xi + m)|^r = \sum_{m \in \mathbb{Z}^d} \prod_{n=1}^d |\operatorname{sinc}(\pi(\xi_n + m_n))|^r \\
&= \prod_{n=1}^d \sum_{m \in \mathbb{Z}} |\operatorname{sinc}(\pi(\xi_n + m_n))|^r = \prod_{n=1}^d h_{n,r}(\xi)
\end{aligned}$$

with $h_{n,r}(\xi) := \sum_{m \in \mathbb{Z}} |\operatorname{sinc}(\pi(\xi_n + m_n))|^r \geq C_r$. $\square$

Proposition 2.4.20. *Let X be a Banach space, fix $d \geq 1$ and $p \in (1, 2]$, and let $\frac{1}{p} + \frac{1}{p'} = 1$. The following assertions are equivalent:*

- (1) $\mathcal{F}$ extends to a bounded operator from $L^p(\mathbb{R}^d; X)$ into $L^{p'}(\mathbb{R}^d; X)$;
- (2) $\mathcal{F}$ extends to a bounded operator from $L^p(\mathbb{T}^d; X)$ into $\ell^{p'}(\mathbb{Z}^d; X)$.
- (3) $\mathcal{F}$ extends to a bounded operator from $\ell^p(\mathbb{Z}^d; X)$ into $L^p(\mathbb{T}^d; X)$.

Denoting the norms of these extensions by $\varphi_{p,X}(\mathbb{R}^d)$, $\varphi_{p,X}(\mathbb{T}^d)$ and $\varphi_{p,X}(\mathbb{Z}^d)$, we have

$$\varphi_{p,X}(\mathbb{R}^d) \leq \varphi_{p,X}(\mathbb{T}^d) \leq C_{p'}^{-d/p'} \varphi_{p,X}(\mathbb{R}^d)$$

and

$$\varphi_{p,X}(\mathbb{Z}^d) = \varphi_{p,X^*}(\mathbb{T}^d), \quad \varphi_{p,X}(\mathbb{T}^d) = \varphi_{p,X^*}(\mathbb{Z}^d).$$

Proof. (1) $\Rightarrow$ (3): Fix $n \geq 0$ and vectors $x_k \in X$, $k \in \mathbb{Z}^d$, $|k| \leq n$, and put $f(t) := \sum_{|k| \leq n} \mathbf{1}_{[0,1]}(t+k) x_k$. By Lemma 2.4.19,

$$\begin{aligned}
C_{p'}^{d/p'} \left\| \sum_{|k| \leq n} e_k x_k \right\|_{L^{p'}(\mathbb{T}^d; X)} &\leq \|\mathcal{F}f\|_{L^{p'}(\mathbb{R}^d; X)} \\
&\leq \varphi_{p,X}(\mathbb{R}^d) \|f\|_{L^p(\mathbb{R}^d; X)} \\
&= \varphi_{p,X}(\mathbb{R}^d) \left(\sum_{|k| \leq n} \|x_k\|^p \right)^{1/p}.
\end{aligned}$$

This also gives the inequality $\varphi_{p,X}(\mathbb{T}^d) \leq C_{p'}^{-d/p'} \varphi_{p,X}(\mathbb{R}^d)$.

(3) $\Rightarrow$ (1): By a density argument it suffices to consider the functions $f : \mathbb{R}^d \rightarrow X$ given by $f(t) = a^{-d/p} \sum_{|k| \leq n} \mathbf{1}_{[0,1]^d - k}(a^{-1}t) x_k$. By Lemma 2.4.19,

$$\begin{aligned} \|\mathcal{F}f\|_{L^{p'}(\mathbb{R}^d; X)} &\leq \left\| \sum_{|k| \leq n} e_k x_k \right\|_{L^{p'}(\mathbb{T}^d; X)} \\ &\leq \varphi_{p,X}(\mathbb{T}^d) \left(\sum_{|k| \leq n} \|x_k\|^p \right)^{1/p} = \varphi_{p,X}(\mathbb{T}^d) \|f\|_{L^p(\mathbb{R}^d; X)}. \end{aligned}$$

This also gives the inequality $\varphi_{p,X}(\mathbb{R}^d) \leq \varphi_{p,X}(\mathbb{T}^d)$.

(3) $\Rightarrow$ (2): By the equivalence (1) $\Leftrightarrow$ (3) X has Fourier type p , and therefore X^* has Fourier type p by Proposition 2.4.16. Hence, by the equivalence (1) $\Leftrightarrow$ (3) for X^* , we find that (3) holds in X^* .

Fix sequences $(x_k)_{|k| \leq n}$ in X and $(x_k^*)_{|k| \leq n}$ in X^* . Then,

$$\begin{aligned} \left| \sum_{|k| \leq n} \langle x_k, x_k^* \rangle \right| &= \left| \int_{\mathbb{T}^d} \left\langle \sum_{|j| \leq n} e_{-j} x_j, \sum_{|k| \leq n} e_k x_k^* \right\rangle dt \right| \\ &\leq \left\| \sum_{|j| \leq n} e_{-j} x_j \right\|_{L^{p'}(\mathbb{T}^d; X)} \left\| \sum_{|k| \leq n} e_k x_k^* \right\|_{L^p(\mathbb{T}^d; X^*)} \\ &\leq \varphi_{p,X^*}(\mathbb{Z}^d) \left\| \sum_{|j| \leq n} e_{-j} x_j \right\|_{L^{p'}(\mathbb{T}^d; X)} \left(\sum_{|k| \leq n} \|x_k^*\|^p \right)^{1/p}. \end{aligned}$$

Now (3) follows by taking the supremum over all sequences $(x_j^*)_{|j| \leq n}$ with $(\sum_{|j| \leq n} \|x_j^*\|^p)^{1/p} \leq 1$; this also gives the inequality $\varphi_{p,X}(\mathbb{T}^d) \leq \varphi_{p,X^*}(\mathbb{Z}^d)$.

(2) $\Rightarrow$ (3): This is proved by a similar duality argument, which also gives the inequality $\varphi_{p,X}(\mathbb{Z}^d) \leq \varphi_{p,X^*}(\mathbb{T}^d)$.

The estimates $\varphi_{p,X^*}(\mathbb{T}^d) \leq \varphi_{p,X}(\mathbb{Z}^d)$ and $\varphi_{p,X^*}(\mathbb{Z}^d) \leq \varphi_{p,X}(\mathbb{T}^d)$ can be proved in the same way. $\square$

2.4.c The Schwartz class $\mathcal{S}(\mathbb{R}^d; X)$

It will be useful to extend the classical theory of Schwartz functions and tempered distributions to the vector-valued setting. The results of this subsection and the next will only be used in our treatment of the Bessel potential spaces in Subsection 5.6.a, and even there, their use is motivated mostly by aesthetic considerations and could easily be avoided. Thus the reason for including this material in the present volume is mainly for pedagogical reasons, although in Volume II it will find important applications.

Let X be a Banach space.

Definition 2.4.21 (Schwartz class). *The Schwartz class of X -valued functions on $\mathbb{R}^d$ is the space*

$$\mathcal{S}(\mathbb{R}^d; X) := \left\{ f \in C^\infty(\mathbb{R}^d; X) : \right. \\ \left. \|f\|_{\alpha, \beta} := \|x \mapsto x^\beta \partial^\alpha f(x)\|_\infty < \infty \ \forall \alpha, \beta \in \mathbb{N}^d \right\}. \quad (2.18)$$

The countable collection of seminorms appearing in (2.18) defines a locally convex topology on $\mathcal{S}(\mathbb{R}^d; X)$. The distance function

$$d(f, g) := \sum_{\alpha \in \mathbb{N}^d} \sum_{\beta \in \mathbb{R}^d} 2^{-|\alpha| - |\beta|} \frac{\|f - g\|_{\alpha, \beta}}{1 + \|f - g\|_{\alpha, \beta}},$$

turns it into a complete metric space. It is easy to check that $\mathcal{S}(\mathbb{R}^d; X)$ is a linear subspace of $L^p(\mathbb{R}^d; X)$ for every $p \in [1, \infty]$ and that the inclusion mapping is continuous.

In the next two propositions, X is a complex Banach space.

Proposition 2.4.22. *The Fourier transform maps $\mathcal{S}(\mathbb{R}^d; X)$ continuously into itself.*

Proof. It is easy to check, by differentiation under the integral in the first case, and by integration by parts in the second, that

$$\partial^\alpha \widehat{f}(\xi) = [(-2\pi i x)^\alpha f]^\wedge(\xi), \quad (2\pi i)^\beta \xi^\beta \widehat{f}(\xi) = [\partial^\beta f]^\wedge.$$

Thus

$$\|\widehat{f}\|_{\alpha, \beta} = \|\xi \mapsto \xi^\beta \partial^\alpha \widehat{f}(\xi)\|_\infty = (2\pi)^{|\alpha| - |\beta|} \|[\partial^\beta (x \mapsto x^\alpha f(x))]^\wedge\|_\infty$$

and

$$\|[\partial^\beta (x^\alpha f)]^\wedge\|_\infty \leq \|\partial^\beta (x^\alpha f)\|_1 \leq \sum_{\gamma \leq \beta \wedge \alpha} \binom{\beta}{\gamma} \frac{\alpha!}{(\alpha - \gamma)!} \|x^{\alpha - \gamma} \partial^{\beta - \gamma} f\|_1,$$

where finally

$$\begin{aligned} \|x^{\alpha - \gamma} \partial^{\beta - \gamma} f\|_1 &\leq \|(1 + |x|)^{-d-1}\|_1 \|(1 + |x|)^{d+1} x^{\alpha - \gamma} \partial^{\beta - \gamma} f\|_\infty \\ &\leq c_d \sum_{|\theta| \leq d+1} \frac{(d+1)!}{\theta! (d+1 - |\theta|)!} \|x^{\theta + \alpha - \gamma} \partial^{\beta - \gamma} f\|_\infty, \end{aligned}$$

so that altogether

$$\|\widehat{f}\|_{\alpha, \beta} \leq c_{d, \alpha, \beta} \sum_{\substack{\gamma \leq \alpha \wedge \beta \\ |\theta| \leq d+1}} \|f\|_{\beta - \gamma, \theta + \alpha - \gamma},$$

from which the assertion follows. $\square$

Further test function classes suitable for different purposes are given by the smooth compactly supported functions

$$\mathcal{D}(\mathbb{R}^d; X) := C_c^\infty(\mathbb{R}^d; X), \quad \mathcal{D}(\mathbb{R}^d \setminus \{0\}; X) := C_c^\infty(\mathbb{R}^d \setminus \{0\}; X),$$

which are obviously subspaces of $\mathcal{S}(\mathbb{R}^d; X)$, and the spaces of inverse Fourier transforms of such functions,

$$\begin{aligned} \check{\mathcal{D}}(\mathbb{R}^d; X) &:= \{g \in \mathcal{S}(\mathbb{R}^d; X) : g = \check{f} \text{ for some } f \in \mathcal{D}(\mathbb{R}^d; X)\}, \\ \check{\mathcal{D}}(\mathbb{R}^d \setminus \{0\}; X) &:= \{g \in \mathcal{S}(\mathbb{R}^d; X) : g = \check{f} \text{ for some } f \in \mathcal{D}(\mathbb{R}^d \setminus \{0\}; X)\}. \end{aligned}$$

We record a basic density result:

Proposition 2.4.23. *The following spaces are dense in $L^p(\mathbb{R}^d; X) \cap \check{L}^1(\mathbb{R}^d; X)$ (and hence in $L^p(\mathbb{R}^d; X)$, by Lemma 2.4.7):*

- (i) $\mathcal{D}(\mathbb{R}^d \setminus \{0\}; X)$, $\mathcal{D}(\mathbb{R}^d; X)$, $\check{\mathcal{D}}(\mathbb{R}^d; X)$ and $\mathcal{S}(\mathbb{R}^d; X)$, for all $p \in [1, \infty)$.
- (ii) $\check{\mathcal{D}}(\mathbb{R}^d \setminus \{0\}; X)$, for all $p \in (1, \infty)$.

Proof. Let $\phi, \varphi \in \mathcal{S}(\mathbb{R}^d)$ with $\phi(0) = \int_{\mathbb{R}^d} \hat{\phi} = 1$ and $\int_{\mathbb{R}^d} \varphi = \hat{\varphi}(0) = 1$. As $\varepsilon, \delta \downarrow 0$ and $R \rightarrow \infty$, we have, with $\varphi_t(x) = t^{-d}\varphi(t^{-1}x)$,

- $\|\phi(\varepsilon \cdot)f - f\|_p \rightarrow 0$ and $\|[\phi(\varepsilon \cdot)f - f]^\wedge\|_1 = \|\hat{\phi}_\varepsilon * \hat{f} - \hat{f}\|_1 \rightarrow 0$.
- $\|\varphi_\delta * f - f\|_p \rightarrow 0$ and $\|[\varphi_\delta * f - f]^\wedge\|_1 = \|\hat{\varphi}(\delta \cdot)\hat{f} - \hat{f}\|_1 \rightarrow 0$.
- $\|\varphi_R * f\|_p \rightarrow 0$ if $p > 1$, and $\|[\varphi_R * f]^\wedge\|_1 = \|\hat{\varphi}(R \cdot)\hat{f}\|_1 \rightarrow 0$.

Now $\varphi_\delta * f$ is smooth and, taking ϕ compactly supported, $\phi(\varepsilon \cdot)(\varphi_\delta * f) \in \mathcal{D}(\mathbb{R}^d; X)$, and these functions approximate $f \in (L^p \cap \hat{L}^1)(\mathbb{R}^d; X)$ to any desired precision. Similarly, $\hat{\phi}_\varepsilon * \hat{f} = (\phi(\varepsilon \cdot)f)^\wedge$ is smooth and, taking $\hat{\varphi}$ compactly supported, $\hat{\varphi}(\delta \cdot)(\hat{\phi}_\varepsilon * \hat{f}) = [\varphi_\delta * (\phi(\varepsilon \cdot)f)]^\wedge \in \mathcal{D}(\mathbb{R}^d; X)$, so that $\varphi_\delta * (\phi(\varepsilon \cdot)f) \in \check{\mathcal{D}}(\mathbb{R}^d; X)$, and again these functions approximate $f \in (L^p \cap \hat{L}^1)(\mathbb{R}^d; X)$ to any desired precision. The density of $\mathcal{S}(\mathbb{R}^d; X)$ follows from either case, since it contains both $\mathcal{D}(\mathbb{R}^d; X)$ and $\check{\mathcal{D}}(\mathbb{R}^d; X)$.

If, in addition, we choose $\hat{\varphi} \equiv 1$ in a neighbourhood of the origin, then $(\hat{\varphi}(\delta \cdot) - \hat{\varphi}(R \cdot))(\hat{\phi}_\varepsilon * \hat{f}) \in \mathcal{D}(\mathbb{R}^d \setminus \{0\}; X)$, and thus the functions $(\varphi_\delta - \varphi_R) * (\phi(\varepsilon \cdot)f) \in \check{\mathcal{D}}(\mathbb{R}^d \setminus \{0\}; X)$ approximate $f \in (L^p \cap \hat{L}^1)(\mathbb{R}^d; X)$ if $p \in (1, \infty)$. $\square$

2.4.d The space of tempered distributions $\mathcal{S}'(\mathbb{R}^d; X)$

For locally convex spaces E and X , we denote by $\mathcal{L}(E, X)$ the space of continuous linear operators from E to X . When X a Banach space, the topology on $\mathcal{L}(E, X)$ generated by the family of seminorms $p_y(T) = \|T(y)\|$, $y \in E$, is locally convex. With respect to this topology one has $T_n \rightarrow T$ in $\mathcal{L}(E, X)$ if and only if $T_n(y) \rightarrow T(y)$ in X for all $y \in E$. In the special case when E and X are both Banach spaces, the topology coincides with strong operator topology of $\mathcal{L}(E, X)$ (cf. Appendix B).

Definition 2.4.24 (Tempered distributions). *The space*

$$\mathcal{S}'(\mathbb{R}^d; X) := \mathcal{L}(\mathcal{S}(\mathbb{R}^d), X)$$

is called the space of X -valued tempered distributions.

Note that $u_n \rightarrow u$ in $\mathcal{S}'(\mathbb{R}^d; X)$ if and only if $u_n(\phi) \rightarrow u(\phi)$ in X for all $\phi \in \mathcal{S}(\mathbb{R}^d)$.

Example 2.4.25. If $u \in \mathcal{S}'(\mathbb{R}^d)$ and $x \in X$, then $u \otimes x : \phi \mapsto u(\phi)x$ defines an element of $\mathcal{S}'(\mathbb{R}^d; X)$.

Example 2.4.26. Let $p \in [1, \infty]$. Every $f \in L^p(\mathbb{R}^d; X)$ defines a tempered distribution $u_f \in \mathcal{S}'(\mathbb{R}^d; X)$ by the prescription

$$u_f(\phi) := \int_{\mathbb{R}^d} f\phi(x) \, dx.$$

From $\|u_f(\phi)\|_X \leq \|f\|_{L^p(\mathbb{R}^d; X)} \|\phi\|_{p'}$ and $\|\phi\|_{p'} \leq \|(1+|x|)^{-k}\|_{p'} \|(1+|x|)^k \phi\|_\infty$ with large enough k , it easily follows that the mapping $f \mapsto u_f$ is continuous. As a consequence of Proposition 2.5.2, this mapping is injective, and therefore it defines a continuous embedding of $L^p(\mathbb{R}^d; X)$ into $\mathcal{S}'(\mathbb{R}^d; X)$. This identification will be used without further comment from now on.

Similarly, if $f : \mathbb{R}^d \rightarrow X$ is a strongly measurable function of polynomial growth, i.e., $\|f(x)\| \leq C(1+|x|)^k$, the above prescription defines a tempered distribution u_f which satisfies $\|u_f(\phi)\|_X \leq C\|(1+|\cdot|)^k \phi\|_1$.

Continuous operators on $\mathcal{S}(\mathbb{R}^d)$ can be extended continuously to $\mathcal{S}'(\mathbb{R}^d; X)$ by a duality argument. Indeed, if $T \in \mathcal{L}(\mathcal{S}(\mathbb{R}^d))$ is continuous, then its adjoint $T' : \mathcal{S}'(\mathbb{R}^d) \rightarrow \mathcal{S}'(\mathbb{R}^d)$ is continuous as well and we obtain a vector-valued extension of T' to a continuous operator on $\mathcal{S}'(\mathbb{R}^d; X)$ by setting

$$(T'(u))(\phi) = u(T(\phi)).$$

For $v = u \otimes x$ with $u \in \mathcal{S}'(\mathbb{R}^d)$ one has $T'(v) = (T'(u)) \otimes x$.

This construction will be applied frequently to the standard operators of harmonic analysis such as dilations, translations, convolutions, differentiations, reflections, and Fourier transforms. Sometimes we perform similar constructions on the level of, say, $L^p(\mathbb{R}^d; X)$, and in such cases we need to verify that the extensions coincide. This will usually be easy, and often it is a consequence of the density of $\mathcal{S}(\mathbb{R}^d) \otimes X$ in $L^p(\mathbb{R}^d; X)$.

Example 2.4.27 (Distributional derivatives). Borrowing the multi-index notation introduced in the next section, for a multi-index $\alpha \in \mathbb{N}^d$ we define the distributional derivative ∂^α on $\mathcal{S}'(\mathbb{R}^d; X)$ by

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

Example 2.4.28 (Pointwise multiplication). Pointwise function with a function $\zeta \in C^\infty(\mathbb{R}^d)$ of polynomial growth defines a continuous linear mapping on $\mathcal{S}(\mathbb{R}^d)$. This mapping may be extended continuously to $\mathcal{S}'(\mathbb{R}^d; X)$ by setting

$$(\zeta u)(\phi) := u(\zeta \phi).$$

On $L^p(\mathbb{R}^d; X)$ this definition agrees with the pointwise definition of multipliers.

Example 2.4.29 (Fourier transform). The Fourier transform $\phi \mapsto \widehat{\phi}$ is continuous on $\mathcal{S}(\mathbb{R}^d)$ by Proposition 2.4.22 and may be extended continuously to $\mathcal{S}'(\mathbb{R}^d; X)$ by setting

$$\widehat{u}(\phi) := u(\widehat{\phi}).$$

For functions in $f \in L^1(\mathbb{R}^d; X) \cup \widetilde{L}^1(\mathbb{R}^d; X)$ and $\phi \in \mathcal{S}(\mathbb{R}^d)$, Fubini's theorem implies

$$\int_{\mathbb{R}^d} \widehat{f} \phi \, dx = \int_{\mathbb{R}^d} f \widehat{\phi} \, dx,$$

and therefore on $L^1(\mathbb{R}^d; X) \cup \widetilde{L}^1(\mathbb{R}^d; X)$ the two definitions agree.

Example 2.4.30 (Fourier multipliers). Combining the previous two examples, for functions $m \in C^\infty(\mathbb{R}^d)$ of polynomial growth we may define a Fourier multiplier $T_m \in \mathcal{L}(\mathcal{S}'(\mathbb{R}^d; X))$ by setting

$$T_m u := (m \widehat{u})^\vee.$$

Alternatively, we could have defined this mapping by the abstract construction using adjoints outlined above. Indeed, we then first define T_m on $\mathcal{S}(\mathbb{R}^d)$ by $T_m \phi := (m \widehat{\phi})^\vee$, and extend this operator to $\mathcal{S}'(\mathbb{R}^d; X)$ by putting $(T_m u)(\phi) := u(T_m \phi)$. The two definitions of T_m as a continuous operator on $\mathcal{S}'(\mathbb{R}^d; X)$ obviously agree.

In Definition 5.3.1 below, another definition of a Fourier multiplier will be given. There, no smoothness of m will be imposed but m instead it is required that m be pointwise bounded. In the present setting, m is required to be smooth but may have polynomial growth.

Example 2.4.31 (Convolution with a Schwartz function). Given a function $g \in \mathcal{S}(\mathbb{R}^d)$, we define a continuous mapping $u \mapsto g * u$ on $\mathcal{S}'(\mathbb{R}^d; X)$ by

$$(g * u)(\phi) := u(\widetilde{g} * \phi),$$

recalling the notation $\widetilde{g}(x) := g(-x)$. This is consistent with the usual definition of the convolution $f \mapsto g * f$ on $L^p(\mathbb{R}^d; X)$, since one has

$$\int_{\mathbb{R}^d} g * f(x) \phi(x) \, dx = \int_{\mathbb{R}^d} f(y) \widetilde{g} * \phi(y) \, dy.$$

More can be said about the tempered distribution $g * u$:

Proposition 2.4.32. *If $u \in \mathcal{S}'(\mathbb{R}^d; X)$ and $g \in \mathcal{S}(\mathbb{R}^d)$, then $g * u$ belongs to $C^\infty(\mathbb{R}^d; X)$, and both $g * u$ and all its derivatives are of polynomial growth.*

Proof. For all $\phi \in \mathcal{S}(\mathbb{R}^d)$ we have

$$(g * u)(\phi) = \int_{\mathbb{R}^d} u(\tau^x(\tilde{g}))\phi(x) dx,$$

recalling the notation $(\tau^h\psi)(x) := \psi(x - h)$. This identity shows that $g * u$ is given by integration against the function $x \mapsto u(\tau^x(\tilde{g}))$.

We claim that $g * u$ is C^∞ and $\partial^\alpha(g * u)(x) = u(\tau^x(\partial^\alpha\tilde{g}))$ for all multi-indices α . In order to prove this, by iteration it suffices to consider the case $\alpha = e_j$. Passing to the limit $t \rightarrow 0$, we have

$$\frac{1}{t}((g * u)(x + te_j) - (g * u)(x)) = \frac{1}{t}(u(\tau^{x+te_j}(\tilde{g})) - u(\tau^x(\tilde{g}))) \rightarrow u(\tau^x\partial^\alpha\tilde{g}),$$

where the latter follows from the elementary fact that for all $\phi \in \mathcal{S}(\mathbb{R}^d)$, $(\tau^{te_j}(\phi) - \phi)/t \rightarrow \partial^\alpha\phi$ in $\mathcal{S}(\mathbb{R}^d)$. This completes the proof of the claim.

We shall use this identity to prove that $\partial^\alpha(g * u)$ has polynomial growth. Indeed, since u is a tempered distribution, there exists an integer N and a constant $C \geq 0$ such that, with the notation introduced in (2.18),

$$\begin{aligned} |\partial^\alpha(g * u)(x)| &= |u(\tau^x(\partial^\alpha\tilde{g}))| \leq C \sum_{|\beta|, |\gamma| \leq N} \|\tau^x(\partial^\alpha\tilde{g})\|_{\beta, \gamma} \\ &= C \sum_{|\beta|, |\gamma| \leq N} \sup_{y \in \mathbb{R}^d} |x + y|^\gamma |\partial^{\alpha+\beta}\tilde{g}(y)|. \end{aligned}$$

Clearly, the right-hand side grows polynomially in the variable x . $\square$

We conclude with a density result that will be useful in our treatment of Bessel potential spaces in Chapter 5.

Proposition 2.4.33. *$C_c^\infty(\mathbb{R}^d; X)$ is sequentially dense in $\mathcal{S}'(\mathbb{R}^d; X)$.*

Proof. Fix $u \in \mathcal{S}'(\mathbb{R}^d; X)$. We must construct $f_n \in C_c^\infty(\mathbb{R}^d; X)$ such that $u_{f_n}(\phi) \rightarrow u(\phi)$ in X for all $\phi \in \mathcal{S}(\mathbb{R}^d)$. Pick a test function $\zeta \in C_c^\infty(\mathbb{R}^d)$ such that $\int_{\mathbb{R}^d} \zeta(x) dx = 1$ and let $\zeta_n(x) := n^d \zeta(nx)$. Let $\theta \in C_c^\infty(\mathbb{R}^d)$ be such that $0 \leq \theta \leq 1$ pointwise and $\theta \equiv 1$ on $\{|x| \leq 1\}$, and put $\theta_n(x) := \theta(x/n)$.

Define $f_n : \mathbb{R}^d \rightarrow X$ by $f_n(x) = \theta_n \cdot (\zeta_n * u)$. Then, by Proposition 2.4.32, $f_n \in C_c^\infty(\mathbb{R}^d; X)$. Moreover, for each $\phi \in \mathcal{S}(\mathbb{R}^d)$,

$$u_{f_n}(\phi) = u_f(\zeta_n * (\theta_n\phi)) \rightarrow u_f(\phi)$$

as $n \rightarrow \infty$, using the elementary fact that $\zeta_n * (\theta_n\phi) \rightarrow \phi$ in $\mathcal{S}(\mathbb{R}^d)$. $\square$

2.5 Sobolev spaces and differentiability

In this section we apply some of the techniques developed so far to a detailed study of differentiability properties of functions taking values in a Banach space X . We introduce the notion of weak derivative and investigate its relationship with almost everywhere differentiability. Along the way, the Sobolev spaces $W^{k,p}(\mathbb{R}^d; X)$ are introduced and some of their properties are developed. We also characterize the real interpolation spaces between $L^p(\mathbb{R}^d; X)$ and $W^{1,p}(\mathbb{R}^d; X)$.

2.5.a Weak derivatives

The following multi-index notation will be used. For $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$ we write

$$\partial^\alpha := \partial_1^{\alpha_1} \circ \dots \circ \partial_d^{\alpha_d}$$

with $\partial_j = \partial/\partial x_j$ the partial derivative with respect to the j th coordinate. The *order* of a multi-index α is the number $|\alpha| := \alpha_1 + \dots + \alpha_d$. Finally, for $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ we write

$$x^\alpha := x_1^{\alpha_1} \cdot \dots \cdot x_d^{\alpha_d}.$$

Let $D \subseteq \mathbb{R}^d$ be an open set. A function $f : D \rightarrow X$ is said to be *locally integrable*, notation $f \in L^1_{\text{loc}}(D; X)$, if it is Bochner integrable on every compact subset of D . We denote by $C^\infty_c(D)$ the space of C^∞ -functions compact support in D .

Definition 2.5.1 (Weak derivatives). *Let $D \subseteq \mathbb{R}^d$ be an open set and let $\alpha \in \mathbb{N}^d$ be a multi-index. A function $g \in L^1_{\text{loc}}(D; X)$ is said to be a weak derivative of order α of the function $f \in L^1_{\text{loc}}(D; X)$ if*

$$\int_D f(x) \partial^\alpha \phi(x) \, dx = (-1)^{|\alpha|} \int_D g(x) \phi(x) \, dx \quad \text{for all } \phi \in C^\infty_c(D).$$

Notions such as the *weak gradient* and *weak Laplacian* are defined similarly.

For functions $f \in L^p(\mathbb{R}^d; X)$ the above definition is consistent with the definition of the distributional derivative introduced in the previous section if we identify such a function with a tempered distribution.

The next proposition settles the uniqueness problem for weak derivatives: if a weak derivative exists, it is necessarily unique. Accordingly, when g is a weak derivative f of order α , it is justified to speak about *the* weak derivative of f and write

$$\partial^\alpha f := g.$$

Proposition 2.5.2. *If a function $f \in L^1_{\text{loc}}(D; X)$ satisfies*

$$\int_D f(x) \phi(x) \, dx = 0 \quad \forall \phi \in C^\infty_c(D),$$

then $f = 0$ almost everywhere.

Proof. Step 1 – It is enough to show that $f = 0$ almost everywhere on every ball B whose closure is contained in D . Fix such a ball B , choose a larger ball $\tilde{B} \subseteq D$ with the same centre as B , and pick $\psi \in C_c^\infty(D)$ satisfying $\psi = 1$ on B and $\psi = 0$ outside $\tilde{B}$. Since $\psi f = f$ on B , it suffices to show that $\psi f = 0$; here ψf is viewed as a function defined on all of $\mathbb{R}^d$.

Step 2 – By Step 1 it suffices to consider the case $D = \mathbb{R}^d$ and $f \in L^1(\mathbb{R}^d; X)$. Let $\phi \in C_c^\infty(\mathbb{R}^d)$ satisfy $\phi(x) = 0$ for $|x| \geq 1$ and $\int_{\mathbb{R}^d} \phi(x) dx = 1$, and for $t > 0$ set $\phi_t(x) := t^{-d} \phi(t^{-1}x)$. For all $t > 0$ and $x \in \mathbb{R}^d$, by the assumption we have

$$\phi_t * f(x) = \int_{\mathbb{R}^d} \phi_t(x-y) f(y) dy = 0.$$

Letting $t \downarrow 0$, it follows from Proposition 1.2.32 that $\phi_t * f \rightarrow f$ in $L^1(\mathbb{R}^d; X)$. Hence $0 = \phi_t * f \rightarrow f$ in $L^1(\mathbb{R}^d; X)$. $\square$

For later use we also record the following fact.

Proposition 2.5.3. *If $D \subseteq \mathbb{R}^d$ is open and connected and $f \in L_{\text{loc}}^1(\mathbb{R}^d; X)$ has weak gradient $\nabla f = 0$ on D , then f equals a constant almost everywhere.*

Proof. Step 1 – For all $\phi, \psi \in C_c^\infty(D)$ we have

$$\int_D (\phi(x) \nabla \psi(x) + \nabla \phi(x) \psi(x)) f(x) dx = - \int_D \phi(x) \psi(x) \nabla f(x) dx,$$

from which it follows that ψf has weak gradient $\nabla(\psi f) = \nabla \psi f + \psi \nabla f$. In particular, if $\psi = 1$ on some closed ball $B \subseteq D$, then $\nabla(\psi f) = 0$ on B .

Step 2 – By Step 1 it suffices to consider the case $D = \mathbb{R}^d$ and $f \in L^1(\mathbb{R}^d; X)$. Pick a non-negative function $\zeta \in C_c^1(\mathbb{R}^d)$ satisfying $\int_{\mathbb{R}^d} \zeta(x) dx = 1$ and set $\zeta_n(x) := n^d \zeta(nx)$. Then $\zeta_n * f \in C^1(\mathbb{R}^d; X)$ and one readily checks that

$$\nabla(\zeta_n * f) = \zeta_n * (\nabla f) = 0.$$

Therefore $\zeta_n * f$ is constant, say equal to C_n . As in the preceding proof, the result now follows from Proposition 1.2.32. $\square$

2.5.b The Sobolev spaces $W^{k,p}(D; X)$

Let X be a Banach space.

Definition 2.5.4 (The Sobolev spaces $W^{k,p}(D; X)$). *Let $D \subseteq \mathbb{R}^d$ be a open set. For $k \in \mathbb{N}$ and $p \in [1, \infty]$, the Sobolev space $W^{k,p}(D; X)$ is the space of all $f \in L^p(D; X)$ whose weak derivatives of all orders $|\alpha| \leq k$ exist and belong to $L^p(D; X)$.*

Endowed with the norm

$$\|f\|_{W^{k,p}(D;X)} := \sum_{|\alpha| \leq k} \|\partial^\alpha f\|_{L^p(D;X)},$$

$W^{k,p}(D;X)$ is a Banach space. The routine proof proceeds by noting that if $(f_n)_{n \geq 1}$ is a Cauchy sequence in $W^{k,p}(D;X)$, then for all $|\alpha| \leq k$ the sequence of weak derivatives $(\partial^\alpha f_n)_{n \geq 1}$ is a Cauchy sequence in $L^p(D;X)$ and hence convergent to some f^α . Set $f := f^0$. It is readily checked that the weak derivatives satisfy $D^\alpha f = f^\alpha$ and that $f_n \rightarrow f$ in $W^{k,p}(D;X)$.

In some situations it is useful to observe that for $1 \leq p < \infty$, an equivalent norm is given by $(\sum_{|\alpha| \leq k} \|\partial^\alpha f\|_{L^p(D)}^p)^{1/p}$. For $p = 2$, this norm turns $W^{2,p}(D)$ into a Hilbert space.

For later use we finish this subsection with a simple approximation result for the case $D = \mathbb{R}^d$. Let us agree on the terminology *differential operator with constant coefficients* for any formal linear combination $A = \sum_{m=1}^M c_m \partial^{\alpha_m}$ where the α_m are multi-indices and $c_m \in \mathbb{K}$ constants. The *weak A-derivative* of a function $f \in L^p(\mathbb{R}^d; X)$ can then be defined in the obvious manner.

Lemma 2.5.5. *Let A be a differential operator with constant coefficients. There exist functions $f_j \in C_c^\infty(\mathbb{R}^d; X)$ such that*

$$f_j \rightarrow f, \quad Af_j \rightarrow Af \quad \text{in } L^p(\mathbb{R}^d; X).$$

Proof. Fix a test function $\theta \in C_c^\infty(\mathbb{R}^d)$ satisfying $\theta \equiv 1$ on $\{|x| \leq 1\}$ and put $\theta_m(x) = \theta(x/m)$. Then it is a simple matter to check that if $f \in L^p(\mathbb{R}^d)$ has a weak A -derivative, then the functions $\theta_m f$ have weak A -derivatives and that $\theta_m f \rightarrow f$ and $Af_m \rightarrow Af$ in $L^p(\mathbb{R}^d; X)$ as $m \rightarrow \infty$. Thus we may assume that f has compact support.

Let us fix another test function $\zeta \in C_c^\infty(\mathbb{R}^d)$, this time with the property that $\int_{\mathbb{R}^d} \zeta(x) dx = 1$, and set $\zeta_n(x) := n^d \zeta(nx)$. It is routine to check that $\zeta_n * f \in C_c^\infty(\mathbb{R}^d)$ and that it has a weak A -derivative given by $A(\zeta_n * f) = \zeta_n * Af$. By Proposition 1.2.32, $\zeta_n * f \rightarrow f$ and $A(\zeta_n * f) = \zeta_n * Af \rightarrow Af$ in $L^p(\mathbb{R}^d; X)$. $\square$

Corollary 2.5.6. *For all $k \in \mathbb{N}$ and $p \in [1, \infty)$, $W^{k,p}(\mathbb{R}^d; X)$ contains $C_c^\infty(\mathbb{R}^d; X)$ as a dense subspace.*

2.5.c Almost everywhere differentiability

The Sobolev spaces $W^{k,p}(D;X)$ have been introduced in terms of the existence of weak derivatives. We now turn to the problem of almost everywhere differentiability of functions belonging to these spaces. Since this is a local question, henceforth we will take $D = \mathbb{R}^d$.

We start with a characterisation of $W^{1,p}(\mathbb{R}^d; X)$ in terms of difference quotients. For $h \in \mathbb{R}^d$ we set

$$\delta_j^t f(x) := \frac{1}{t}(f(x + te_j) - f(x)), \quad 1 \leq j \leq d, \quad t \in \mathbb{R} \setminus \{0\}.$$

Here e_j is the j th standard unit vector of $\mathbb{R}^d$. Later on we shall only need the easy first part of the next result.

Proposition 2.5.7. *Let X be a Banach space.*

(1) *Let $1 \leq p \leq \infty$. If $f \in W^{1,p}(\mathbb{R}^d; X)$, then*

$$\|\delta_j^t f\|_p \leq \|\partial_j f\|_p, \quad 1 \leq j \leq d, \quad t \in \mathbb{R} \setminus \{0\},$$

where $\partial_j f$ denotes the j th partial derivative of f .

(2) *Let $1 < p < \infty$. If X has the Radon–Nikodým property and for all $f \in L^p(\mathbb{R}^d; X)$ there exists a constant $C \geq 0$ such that*

$$\|\delta_j^t f\|_p \leq C, \quad 1 \leq j \leq d, \quad t \in \mathbb{R} \setminus \{0\},$$

then $f \in W^{1,p}(\mathbb{R}^d; X)$ and

$$\|\partial_j f\|_p \leq C, \quad 1 \leq j \leq d.$$

Part (2) of the above result actually characterises the Radon–Nikodým property; this will be an easy corollary to Theorem 2.5.12. Note that (2) is wrong for $p = 1$, even for $X = \mathbb{K}$ and $d = 1$: consider the indicator function $f = \mathbf{1}_{(0,1)}$.

Proof. (1): By density it suffices to consider $f \in C_c^1(\mathbb{R}^d)$. For such f we have

$$f(x+h) - f(x) = \int_0^1 \frac{d}{ds} f(x+sh) dt = \int_0^1 \nabla f(x+sh) \cdot h ds. \quad (2.19)$$

Taking $h = te_j$ we find

$$\|\delta_j^t f\|_p \leq \int_0^1 \|\partial_j f(\cdot + se_j)\|_p ds = \|\partial_j f\|_p.$$

(2): Fix $f \in L^p(\mathbb{R}^d; X)$ and let C be the constant in the statement of the result. For all $\phi \in C_c^\infty(\mathbb{R}^d)$ we have

$$\int_{\mathbb{R}^d} (f(x+h) - f(x))\phi(x) dx = \int_{\mathbb{R}^d} f(x)(\phi(x-h) - \phi(x)) dx.$$

By the assumption, for all $t > 0$ and $h := e_j$ this gives

$$\left\| \int_{\mathbb{R}^d} f(x) \frac{\phi(x - te_j) - \phi(x)}{t} dx \right\| \leq C \|\phi\|_{p'}, \quad 1 \leq j \leq d.$$

Passing to the limit $t \downarrow 0$ we obtain

$$\left\| \int_{\mathbb{R}^d} f(x) \partial_j \phi(x) \, dx \right\| \leq C \|\phi\|_{p'}. \quad (2.20)$$

Thus $\phi \mapsto \int_{\mathbb{R}^d} f(x) \partial_j \phi(x) \, dx$ defines a bounded operator $T_j : L^{p'}(\mathbb{R}^d) \rightarrow X$.

Fix an increasing sequence of balls $(B_n)_{n \geq 1}$ in $\mathbb{R}^d$ such that $\bigcup_{n \geq 1} B_n = \mathbb{R}^d$. Define $G_{j,n} : \mathcal{B}(\mathbb{R}^d) \rightarrow X$ by $G_n(A) := T_j(\mathbf{1}_{A \cap B_n})$. As in the proof of Theorem 1.3.10, (1) $\Rightarrow$ (2), this defines an X -valued vector measure of bounded variation (it is here that we need $p > 1$). Since X has the RNP, $G_{j,n}$ is represented by a function $g_{j,n} \in L^1(\mathbb{R}^d; X)$ which vanishes outside B_n :

$$T_j(\mathbf{1}_{A \cap B_n}) = \int_A g_{j,n}(x) \, dx.$$

As in the proof of Theorem 1.3.10, the functions $g_{j,n}$ agree on the intersections of their supports, and therefore they can be pieced together to a globally defined function $g_j \in L^1_{\text{loc}}(\mathbb{R}^d; X)$.

By linearity, for simple functions ϕ this gives

$$T_j \phi = \int_{\mathbb{R}^d} g_j(x) \phi(x) \, dx.$$

By approximation, this identity extends to functions $\phi \in C_c(\mathbb{R}^d)$. Combining the resulting identity with the definition of T , it takes the form

$$\int_{\mathbb{R}^d} f(x) \partial_j \phi(x) \, dx = T_j \phi = \int_{\mathbb{R}^d} g_j(x) \phi(x) \, dx.$$

This shows that $g_j = \partial_j f$ in the sense of weak derivatives.

Next we show that $\partial_j f \in L^p(\mathbb{R}^d; X)$. Arguing as in (2.20) one sees that for all $\phi \in C_c^\infty(\mathbb{R}^d; X^*)$

$$\left| \int_{\mathbb{R}^d} \langle f(x), \partial_j \phi(x) \rangle \, dx \right| \leq C \|\phi\|_{p'}.$$

Combining this with the definition of the weak derivative yields that for all $\phi \in C_c^\infty(\mathbb{R}^d) \otimes X^*$,

$$\left| \int_{\mathbb{R}^d} \langle \partial_j f(x), \phi(x) \rangle \, dx \right| \leq C \|\phi\|_{p'}.$$

Since $C_c^\infty(\mathbb{R}^d) \otimes X^*$, as a subspace of $L^{p'}(\mathbb{R}^d; X^*)$, is norming for $L^p(\mathbb{R}^d; X)$ (see Propositions 1.3.1 and 2.4.23) we conclude that $\partial_j f \in L^p(\mathbb{R}^d; X)$. $\square$

As an application of the Lebesgue differentiation theorem, we shall prove next that every Lipschitz continuous function $f : \mathbb{R} \rightarrow X$ is almost everywhere differentiable if and only if X has the Radon–Nikodým property. In Proposition 2.5.7 we have already seen that the RNP can be used to obtain the existence of weak derivatives for functions $f : \mathbb{R}^d \rightarrow X$ satisfying a certain Lipschitz

estimate. We will now focus on the one-dimensional case and obtain equivalences between weak differentiability and almost everywhere differentiability. Some of the results do extend to higher dimensions, but this requires a more thorough development of Sobolev spaces.

We begin by stating a version of the fundamental theorem of calculus for weak and almost everywhere derivatives. In order to distinguish between them, we shall write f' for the almost everywhere derivative and ∂f for the weak derivative whenever they exist.

It will be convenient to put

$$\int_a^t g(s) \, ds := - \int_t^a g(s) \, ds \quad \text{for } t < a.$$

Lemma 2.5.8. *Let $g \in L^1_{\text{loc}}(\mathbb{R}; X)$ and $a \in \mathbb{R}$, and define $f : \mathbb{R} \rightarrow X$ by*

$$f(t) := \int_a^t g(s) \, ds.$$

Then the weak derivative ∂f and almost everywhere derivative f' of f both exist in $L^1_{\text{loc}}(\mathbb{R}; X)$ and are given by $\partial f = f' = g$.

Proof. For any test function $\phi \in C_c^\infty(\mathbb{R}^d)$ we have

$$\begin{aligned} \int_{\mathbb{R}} f(t) \phi'(t) \, dt &= \int_a^\infty \int_a^t g(s) \phi'(t) \, ds \, dt - \int_{-\infty}^a \int_t^a g(s) \phi'(t) \, ds \, dt \\ &= \int_a^\infty \int_s^\infty g(s) \phi'(t) \, dt \, ds - \int_{-\infty}^a \int_{-\infty}^s g(s) \phi'(t) \, dt \, ds \\ &= - \int_a^\infty \phi(s) g(s) \, ds - \int_{-\infty}^a \phi(s) g(s) \, ds. \end{aligned}$$

This shows that the weak derivative of f exists and equals g .

We will prove next that f is almost everywhere differentiable on the interval (a, ∞) , with derivative g . The corresponding result on $(-\infty, a)$ is proved in the same way.

Fix an arbitrary $b > a$ and let $g_n : (a, b) \rightarrow X$, $n \geq 1$, be bounded continuous functions such that $\lim_{n \rightarrow \infty} g_n = g$ almost everywhere. Such functions exist by the density of the step functions of the form $\sum_{n=1}^N \mathbf{1}_{(t_{n-1}, t_n)} \otimes x_n$ in $L^1(a, b; X)$ (see Remark 1.2.20) and the fact that each indicator function $\mathbf{1}_{(t_{n-1}, t_n)}$ can be approximated in $L^1(a, b)$ by bounded continuous functions. By the Lebesgue differentiation theorem (Corollary 2.3.5), applied to the non-negative functions $s \mapsto \|g(s) - g_n(s)\|$, there exists a null set N such that for all $t \in (a, b) \setminus N$,

$$\limsup_{h \downarrow 0} \left\| \frac{1}{h} (f(t+h) - f(t)) - g(t) \right\|$$

$$\begin{aligned}
&\leq \limsup_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} \|g(s) - g(t)\| \, ds \\
&\leq \limsup_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} \|g(s) - g_n(s)\| \, ds \\
&\quad + \limsup_{h \downarrow 0} \frac{1}{h} \int_t^{t+h} \|g_n(s) - g_n(t)\| \, ds + \|g_n(t) - g(t)\| \\
&= 2\|g_n(t) - g(t)\|.
\end{aligned}$$

The almost everywhere right differentiability of f follows from this by letting $n \rightarrow \infty$. The almost everywhere left differentiability of f is proved in the same way. $\square$

The anti-derivative f of g in Lemma 2.5.8 is *absolutely continuous* on every bounded interval $I \subseteq \mathbb{R}$, i.e., for $\varepsilon > 0$ there exists a $\delta > 0$ such that if (a_n, b_n) is a sequence of disjoint intervals in I satisfying $\sum_{n \geq 1} (b_n - a_n) < \delta$, then

$$\sum_{n \geq 1} \|f(b_n) - f(a_n)\| < \varepsilon.$$

This raises the question whether, conversely, every (locally) absolutely continuous function $f : \mathbb{R} \rightarrow X$ is the anti-derivative of some (locally) Bochner integrable function $g : \mathbb{R} \rightarrow X$. Theorem 2.5.12 provides the answer. Before we prove it we further investigate the relation between almost everywhere and weak derivatives.

A function $f : \mathbb{R} \rightarrow X$ is said to belong to $W_{\text{loc}}^{1,p}(\mathbb{R}; X)$ if f belongs to $L_{\text{loc}}^p(\mathbb{R}; X)$ and has a weak derivative belonging to $L_{\text{loc}}^p(\mathbb{R}; X)$.

Proposition 2.5.9. *Let $p \in [1, \infty]$. For a locally absolutely continuous function $f : \mathbb{R} \rightarrow X$ the following are equivalent:*

- (1) $f \in W_{\text{loc}}^{1,p}(\mathbb{R}; X)$;
- (2) *there exists a function $g \in L_{\text{loc}}^p(\mathbb{R}; X)$ such that*

$$f(t) - f(s) = \int_s^t g(r) \, dr \quad \text{for all } -\infty < s \leq t < \infty;$$

- (3) *f is almost everywhere differentiable and $f' \in L_{\text{loc}}^p(\mathbb{R}; X)$.*

In this case $\partial f = g = f'$ almost everywhere.

For the proof that (3) implies (2) we need a covering lemma. A cover of a set E by a family $\mathcal{B}$ of balls is called a *Vitali cover* if for all $x \in E$ and all $r > 0$ there is a ball $B \in \mathcal{B}$ of radius less than r containing x .

Lemma 2.5.10. *Let $\mathcal{B}$ be a Vitali cover of a measurable set $E \subseteq (a, b)$ of full measure. Then for all $\delta > 0$ there exist disjoint $B_1, \dots, B_N \in \mathcal{B}$ such that*

$$|E \setminus \bigcup_{n=1}^N B_n| \leq \delta.$$

Proof. We may assume that $\delta < b - a$. Set $E_1 := E$. Pick a compact set $K_1 \subseteq E_1$ of measure $> \delta$ and let $B_1, \dots, B_{N_1}$ be a finite cover of K by balls in $\mathcal{B}$. By the Vitali covering lemma (Lemma 2.3.1) it is possible to select pairwise disjoint balls $B_{n_1}, \dots, B_{n_{k_1}}$ from this cover in such a way that

$$\sum_{j=1}^{k_1} |B_{n_j}| \geq \frac{1}{3} |K_1| > \frac{1}{3} \delta.$$

If $|E \setminus \bigcup_{j=1}^{k_1} B_{n_j}| \leq \delta$ we are finished; otherwise $E_2 := E_1 \setminus \sum_{j=1}^{k_1} \overline{B_{n_j}}$ has measure $> \delta$. We can then choose a compact $K_2 \subseteq E_2$ of measure $> \delta$. We cover K_2 by balls from $\mathcal{B}$, all of which may be taken disjoint from the balls already chosen (here we use that $\mathcal{B}$ is a Vitali cover: this permits us to pick balls for the cover of small enough radii), and use the Vitali covering lemma to find pairwise disjoint balls $B_{n_{k_1+1}}, \dots, B_{n_{k_2}}$ from this cover such that

$$\sum_{j=k_1+1}^{k_2} |B_{n_j}| \geq \frac{1}{3} |K_2| > \frac{1}{3} \delta.$$

If $|E \setminus \bigcup_{j=1}^{k_2} B_{n_j}| \leq \delta$ we are finished; otherwise we continue.

This procedure must come to a halt after finitely many steps, for

$$|E_n| < |E_{n-1}| - \frac{1}{3} \delta < \dots < |E_1| - \frac{1}{3} (n-1) \delta |E_1| = (b-a) - \frac{1}{3} (n-1) \delta,$$

so eventually $|E_n| \leq \delta$. □

We use this lemma to prove:

Lemma 2.5.11. *Let $f : [0, 1] \rightarrow X$ be locally absolutely continuous and differentiable almost everywhere. If $f' = 0$ almost everywhere, then f is constant.*

Proof. Fix $\varepsilon > 0$ and choose $\delta > 0$ as in the definition of absolute continuity. Fix $0 \leq a < b \leq 1$ and let $E \subseteq (a, b)$ be a measurable set of full measure on which f is differentiable with derivative equal to 0. Since for each $x \in E$ we have $\lim_{h \rightarrow 0} \frac{1}{h} (f(x+h) - f(x)) = 0$, for each $r > 0$ there is an open interval $(a_x, b_x) \subseteq (a, b)$ containing x of diameter $< r$ such that

$$\|f(b_x) - f(a_x)\| < \varepsilon (b_x - a_x).$$

These intervals form a Vitali cover for E . Using the previous lemma we select finitely many intervals from this cover, say $I_1, \dots, I_N$, which are pairwise disjoint and satisfy $\sum_{n=1}^N |I_n| \geq (b-a) - \delta$. Writing $I_n = (a_n, b_n)$, we have

$$\sum_{n=1}^N \|f(b_n) - f(a_n)\| < \varepsilon \sum_{n=1}^N (b_n - a_n) < \varepsilon(b - a).$$

The complement of these intervals consists of at most $N + 1$ disjoint intervals $J_m = [c_m, d_m]$ of total length $< \delta$. Therefore, by the absolute continuity of f ,

$$\sum_{J_m} \|f(d_m) - f(c_m)\| < \varepsilon.$$

Putting together these estimates, it follows that $\|f(b) - f(a)\| < \varepsilon(b - a) + \varepsilon$. Since $\varepsilon > 0$ was arbitrary, this proves that $f(a) = f(b)$. $\square$

Proof of Proposition 2.5.9. (1) $\Rightarrow$ (2): Fix real numbers $s < s'$ and define $h : [s, s'] \rightarrow X$ by $h(t) := \int_a^t g(r) dr$, where $g = \partial f$ is the weak derivative of f . By Lemma 2.5.8, h has weak derivative $\partial h = \partial f$. By Proposition 2.5.3, $f - h \equiv x$ is constant, and therefore $f(t) = x + \int_s^t g(r) dr$ for all $t \in [s, s']$. Passing to the limit $t \downarrow s$ we see that $x = f(s)$ and (2) follows.

(2) $\Rightarrow$ (1) and (2) $\Rightarrow$ (3) both follow from Lemma 2.5.8.

(3) $\Rightarrow$ (2): Again fix real numbers $s < s'$ and set $h(t) := \int_s^t f'(r) dr$. By Lemma 2.5.8 this function is absolutely continuous and almost everywhere differentiable with derivative f' . It follows that $f - h$ is absolutely continuous and almost everywhere differentiable on $[s, s']$, with derivative 0. By Lemma 2.5.11, $f - h$ is constant. The proof can now be completed as in the same way as above. $\square$

Theorem 2.5.12. *For any Banach space X , the following assertions are equivalent:*

- (1) X has the Radon–Nikodým property;
- (2) every locally absolutely continuous function $f : \mathbb{R} \rightarrow X$ is differentiable almost everywhere;
- (3) every locally Lipschitz continuous function $f : \mathbb{R} \rightarrow X$ is differentiable almost everywhere;
- (4) every locally absolutely continuous function $f : \mathbb{R} \rightarrow X$ belongs to $W_{\text{loc}}^{1,1}(\mathbb{R}; X)$;
- (5) every locally Lipschitz continuous function $f : \mathbb{R} \rightarrow X$ belongs to $W_{\text{loc}}^{1,\infty}(\mathbb{R}; X)$.

Proof. (1) $\Rightarrow$ (3): It is possible to deduce this from Proposition 2.5.7, but we prefer to give a different argument here. It suffices to prove the almost everywhere differentiability on a finite interval $I = (a, b)$. For $a \leq s < t \leq b$ define

$$T\mathbf{1}_{(s,t)} := f(t) - f(s)$$

and extend this definition to step functions by linearity. For such functions $g = \sum_{n=1}^N c_n \mathbf{1}_{(t_{n-1}, t_n)}$,

$$\|Tg\| \leq L \sum_{n=1}^N |c_n|(t_n - t_{n-1}) = L\|g\|_1,$$

where L is the Lipschitz constant of f . Therefore T extends to a bounded operator from $L^1(I)$ to X of norm $\|T\| \leq L$. By Theorem 1.3.10, T is representable by a function $\phi \in L^\infty(I; X)$. For all $t \in I$ this gives

$$f(t) - f(a) = T\mathbf{1}_{(a,t)} = \int_a^t \phi(r) \, dr,$$

and Lemma 2.5.8 gives the result.

(3) $\Leftrightarrow$ (5): This follows from Proposition 2.5.9, noting that the almost everywhere derivative of a Lipschitz function is bounded.

(5) $\Rightarrow$ (1): Let $T : L^1(0, 1) \rightarrow X$ be a bounded operator. Define $f : [0, 1] \rightarrow X$ by

$$f(t) := T\mathbf{1}_{(0,t)}.$$

Then for all $0 \leq s < t \leq 1$ we have

$$\|f(t) - f(s)\| \leq \|T\| \|\mathbf{1}_{(s,t)}\|_1 = (t - s)\|T\|,$$

so f is Lipschitz continuous. Now extend f to a Lipschitz function on $\mathbb{R}$ by setting $f(t) = f(0)$ for $t < 0$ and $f(t) = f(1)$ for $t > 1$. Let f' be the almost everywhere defined derivative of f . Clearly, $f' \in L^\infty(\mathbb{R}; X)$ with $\|f'\|_\infty \leq L$, the Lipschitz constant of f . Using Proposition 2.5.9, for all $0 \leq a < b \leq 1$ we obtain

$$T\mathbf{1}_{(a,b)} = f(b) - f(a) = \int_a^b f'(t) \, dt = \int_0^1 \mathbf{1}_{(a,b)}(t) f'(t) \, dt.$$

It follows that $T\phi = \int_0^1 \phi(t) f'(t) \, dt$ holds for all step function ϕ . Since these are dense in $L^1(0, 1)$ and T is bounded, this identity extends to arbitrary $\phi \in L^1(0, 1)$. This proves that T is representable, and therefore X has the RNP with respect to $[0, 1]$ by Theorem 1.3.10.

(5) $\Rightarrow$ (4): Let $f : \mathbb{R} \rightarrow X$ be absolutely continuous. It suffices to prove that f is almost everywhere differentiable with integrable derivative on finite intervals. By a scaling and translation argument it suffices to consider $I = (0, 1)$. Define the function $V_f : [0, 1] \rightarrow \mathbb{R}$ by

$$V_f(t) := \sup \sum_{n=1}^N \|f(t_n) - f(t_{n-1})\|,$$

where the supremum is taken over all finite partitions $0 = t_0 < \dots < t_N = 1$. Note that $V_f(t)$ is finite, because f is absolutely continuous, and as a function of t , V_f is non-decreasing. The function

$$h(t) := \frac{V_f(t) + t}{V_f(1) + 1}$$

is absolutely continuous, strictly increasing, and satisfies $h(0) = 0$, $h(1) = 1$. Moreover, for all $0 \leq s \leq t \leq 1$,

$$\|f(t) - f(s)\| \leq V_f(t) - V_f(s) \leq (V_f(1) + 1)(h(t) - h(s)).$$

Hence $f \circ h^{-1}$ is Lipschitz continuous on $[0, 1]$, with Lipschitz constant $\leq V_f(1) + 1$. By assumption, this implies that $f \circ h^{-1}$ is differentiable almost everywhere, say outside some null set N_0 , and clearly its derivative is bounded by $V_f(1) + 1$ almost everywhere, say outside some null set N_1 . Also, h is differentiable almost everywhere as well, say outside some null set N_2 , and the derivative h' is integrable; the proof of this fact will be given in Lemma 2.5.13 below. From $|h(t) - h(s)| \geq |t - s|/(V_f(1) + 1)$ it follows that h^{-1} maps null sets to null sets. Therefore, by the chain rule, $f = (f \circ h^{-1}) \circ h$ is differentiable on the complement of the null set $N := h^{-1}(N_0 \cup N_1) \cup N_2$ and there it satisfies $f'(t) = (f \circ h^{-1})'(h(t)) \cdot h'(t)$ and

$$\|f'(t)\| \leq (V_f(1) + 1)|h'(t)|.$$

This shows that f' is Bochner integrable.

(4) $\Rightarrow$ (2): This is a consequence of Proposition 2.5.9.

(2) $\Rightarrow$ (3): Every Lipschitz function is absolutely continuous. $\square$

It remains to prove the following fact, which was used in the proof of the implication (5) $\Rightarrow$ (2):

Lemma 2.5.13. *If $f : \mathbb{R} \rightarrow [0, \infty)$ is non-decreasing and absolutely continuous, then f is differentiable almost everywhere and f' is locally integrable.*

Proof. For $a < b$ set $\mu([a, b)) := f(b) - f(a)$ and extend this definition to finite unions of such half-open intervals. By the Carathéodory extension theorem, μ extends to a Borel measure μ on $\mathbb{R}$ which is finite on bounded intervals. We will show that μ is absolutely continuous with respect to the Lebesgue measure. To this end fix $\varepsilon > 0$ and let δ be as in the definition of absolute continuity.

If B is a Borel set of Lebesgue measure $|B| < \frac{1}{2}\delta$, invoking Lemma A.1.2 we find sets $I_n = [a_n, b_n)$, $n = 1, \dots, N$, such that both $|B \Delta \bigcup_{n=1}^N I_n| < \frac{1}{2}\delta$ and $\mu(B \Delta \bigcup_{n=1}^N I_n) < \varepsilon$. The former implies that

$$\sum_{n=1}^n (b_n - a_n) = \bigcup_{n=1}^N |I_n| < \delta.$$

Hence by the absolute continuity of f ,

$$\mu\left(\bigcup_{n=1}^N I_n\right) = \sum_{n=1}^N (f(b_n) - f(a_n)) < \varepsilon.$$

Therefore $\mu(B) < 2\varepsilon$. This proves the absolute continuity of μ . By the Radon–Nikodým theorem, applied to the intervals $(-r, r)$, μ has a density $g \in L^1_{\text{loc}}(\mathbb{R})$. In particular,

$$f(b) - f(a) = \int_a^b g(s) \, ds$$

for all $a < b$. Now the result, with $f' = g$, follows from Lemma 2.5.8. $\square$

Remark 2.5.14. The implication (1) $\Rightarrow$ (2) of the theorem admits the following direct and more transparent proof, although its rigorous implementation is somewhat laborious. The idea is that the X -valued Stieltjes integral

$$G(A) := \int_0^1 \mathbf{1}_A \, df$$

defines an X -valued measure of bounded variation which is absolutely continuous with respect to the Lebesgue measure. If $\phi \in L^1(0, 1; X)$ denotes its density, we have

$$f(t) = \int_0^t \phi(t) \, dt$$

and the result follows from Lemma 2.5.8.

Corollary 2.5.15. *Let X be reflexive.*

- (1) *Every locally absolutely continuous function $f : \mathbb{R} \rightarrow X$ is differentiable almost everywhere, and $f \in W^{1,1}_{\text{loc}}(\mathbb{R}; X)$.*
- (2) *Every locally Lipschitz continuous function $f : \mathbb{R} \rightarrow X$ is differentiable almost everywhere, and $f \in W^{1,\infty}_{\text{loc}}(\mathbb{R}; X)$.*

In both situations, the almost everywhere derivative and the weak derivative agree.

2.5.d The fractional Sobolev spaces $W^{s,p}(\mathbb{R}^d; X)$

In this subsection we introduce the scale of Sobolev spaces $W^{s,p}(\mathbb{R}^d; X)$ and identify them, for $0 < s < 1$, as real interpolation spaces between $L^p(\mathbb{R}^d; X)$ and $W^{1,p}(\mathbb{R}^d; X)$. The results of this subsection will not be needed elsewhere in this volume. We will revisit Sobolev spaces in Chapter 5.

Definition 2.5.16 (Sobolev–Slobodetskii spaces). *Let $D \subseteq \mathbb{R}^d$ be an open set. For $p \in [1, \infty)$ and $0 < s < 1$ we define $W^{s,p}(D; X)$ as the space of all $f \in L^p(D; X)$ for which*

$$[f]_{W^{s,p}(D; X)} := \left(\int_D \int_D \frac{\|f(x) - f(y)\|^p}{|x - y|^{sp+d}} \, dx \, dy \right)^{1/p} < \infty.$$

Endowed with the norm

$$\|f\|_{W^{s,p}(D;X)} := \|f\|_p + [f]_{W^{s,p}(D;X)},$$

$W^{s,p}(\mathbb{R}^d; X)$ is a Banach space.

In this subsection we will take $D = \mathbb{R}^d$ and note that

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

with continuous embeddings. For the first embedding note that

$$[f]_{W^{s,p}(\mathbb{R}^d; X)} = \left(\int_{\mathbb{R}^d} \frac{\|f(x+h) - f(x)\|_p^p}{|h|^{sp+d}} dx dh \right)^{1/p}.$$

Now for $|h| > 1$ one may use $\|f(\cdot + h) - f(\cdot)\|_p \leq 2\|f\|_p$, and for $|h| \leq 1$ one may use that $\|f(\cdot + h) - f(\cdot)\|_p \leq \|f\|_{W^{1,p}(\mathbb{R}^d; X)}|h|$, which follows from (2.19).

We are now ready to state and prove the main result of this section:

Theorem 2.5.17 (Real interpolation). *Let X be a Banach space. For all $p \in [1, \infty)$ and $0 < s < 1$ we have*

$$(L^p(\mathbb{R}^d; X), W^{1,p}(\mathbb{R}^d; X))_{s,p} = W^{s,p}(\mathbb{R}^d; X)$$

with equivalent norms.

Proof. Below we write $K(t, f) := K(t, f, L^p, W^{1,p})$ for the K -functional as introduced in Section C.3.

‘ $\subseteq$ ’: Fix $f \in (L^p(\mathbb{R}^d; X), W^{1,p}(\mathbb{R}^d; X))_{s,p}$. For fixed $h \in \mathbb{R}^d$ choose $a = a_h \in L^p(\mathbb{R}^d; X)$ and $b = b_h \in W^{1,p}(\mathbb{R}^d; X)$ such that $f = a + b$ and

$$\|a\|_p + |h|\|b\|_{W^{1,p}} \leq 2K(|h|, f).$$

Using the triangle inequality and Proposition 2.5.7 we find

$$\begin{aligned} \|f(\cdot + h) - f\|_p &\leq \|a(\cdot + h) - a\|_p + \|b(\cdot + h) - b\|_p \\ &\leq 2\|a\|_p + |h|\|\nabla b\|_p \leq 2\|a\|_p + |h|\|b\|_{W^{1,p}} \leq 4K(|h|, f). \end{aligned}$$

Integrating over h and turning to polar coordinates,

$$[f]_{W^{s,p}(\mathbb{R}^d; X)}^p \leq 4^p \int_{\mathbb{R}^d} \frac{(K(|h|, f))^p}{|h|^{sp+d}} dh \quad (2.21)$$

$$= C_{d,p} \int_0^\infty t^{-sp} (K(t, f))^p \frac{dt}{t} = C_{d,p} \|f\|_{(L^p(\mathbb{R}^d; X), W^{1,p}(\mathbb{R}^d; X))_{s,p}}^p \quad (2.22)$$

with $C_{d,p} = 4^p \times$ the area of the unit sphere of $\mathbb{R}^d$. Secondly, since $W^{1,p}(\mathbb{R}^d; X)$ embeds into $L^p(\mathbb{R}^d; X)$ continuously, we also have $\min\{1, t\}\|f\|_p \lesssim_{d,p} K(t, f)$

(see (C.7)) for all $t > 0$. Multiplying both sides with t^{-s} and integrating, we obtain

$$\|f\|_p \leq C_{p,s} \|f\|_{(L^p(\mathbb{R}^d; X), W^{1,p}(\mathbb{R}^d; X))_{s,p}}. \quad (2.23)$$

Combining the estimates (2.21) and (2.23), the desired embedding is obtained.

' $\supseteq$ ': For $t \geq 1$ the trivial estimate $K(t, f) \leq \|f\|_p$ already gives, for $f \in W^{s,p}(\mathbb{R}^d; X)$,

$$\int_1^\infty t^{-sp} (K(t, f))^p \frac{dt}{t} \leq \int_1^\infty t^{-sp} \|f\|_p^p \frac{dt}{t} = C_{s,p} \|f\|_p^p \leq C_{s,p} \|f\|_{s,p}^p.$$

To estimate the integral over $0 < t < 1$, let $\phi \in C_c^1(\mathbb{R}^d)$ be a non-negative test function supported in the unit ball of $\mathbb{R}^d$ satisfying $\int_{\mathbb{R}^d} \phi(x) dx = 1$. Set $\phi_t(x) := t^{-d} \phi(t^{-1}x)$ and define

$$b_t := \int_{\mathbb{R}^d} \phi_t(x - y) f(y) dy, \quad a_t := f - b_t.$$

In order to prove the required embedding it remains to be shown that the right-hand side terms of

$$\begin{aligned} & \left(\int_0^1 t^{-sp} (K(t, f))^p \frac{dt}{t} \right)^{1/p} \\ & \leq \left(\int_0^1 t^{-sp} (\|a_t\|_p + t \|b_t\|_{W^{1,p}})^p \frac{dt}{t} \right)^{1/p} \\ & \leq \left(\int_0^1 t^{-sp} \|a_t\|_p^p \frac{dt}{t} \right)^{1/p} + \left(\int_0^1 t^{(1-s)} \left(\sum_{|\alpha| \leq 1} \|\partial^\alpha b_t\|_p \right)^p \frac{dt}{t} \right)^{1/p} \end{aligned}$$

can be estimated by $C_{s,d,p} \|f\|_{s,p}$.

From $f(x) = \int_{\mathbb{R}^d} \phi_t(x - y) f(y) dy$ we find

$$\begin{aligned} \|a_t\|_p^p &= \int_{\mathbb{R}^d} \left\| \int_{\mathbb{R}^d} \phi_t(x - y) [f(x) - f(y)] dy \right\|^p dx \\ &\leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \phi_t(x - y) \|f(y) - f(x)\|^p dy dx, \end{aligned}$$

applying Jensen's inequality to the probability measures $\phi_t(x - y) dy$. Hence,

$$\begin{aligned} \int_0^\infty t^{-sp} \|a_t\|_p^p \frac{dt}{t} &\leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left(\int_0^\infty t^{-sp} \phi_t(x - y) \frac{dt}{t} \right) \|f(y) - f(x)\|^p dy dx \\ &\stackrel{(*)}{=} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left(\int_{|x-y|}^\infty t^{-sp-d} \phi\left(\frac{x-y}{t}\right) \frac{dt}{t} \right) \|f(y) - f(x)\|^p dy dx \\ &\leq \frac{\|\phi\|_\infty}{sp+d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{\|f(y) - f(x)\|^p}{|x-y|^{sp+d}} dy dx \end{aligned}$$

$$= \frac{\|\phi\|_\infty}{sp+d} [f]_{W^{s,p}(\mathbb{R}^d; X)}^p,$$

where $(*)$ uses the definition of ϕ_t and the assumption that ϕ be supported in the unit ball.

Clearly $b_t \in W^{1,p}(\mathbb{R}^d)$. Since $\|b_t\|_p \leq \|f\|_p$ by Young's inequality,

$$\int_0^1 t^{(1-s)p} \|b_t\|_p^p \frac{dt}{t} \leq \int_0^1 t^{(1-s)p} \|f\|_p^p \frac{dt}{t} = C_{p,s} \|f\|_p^p.$$

Next,

$$\begin{aligned} \partial_j b_t(x) &= \frac{1}{t^{d+1}} \int_{\mathbb{R}^d} \partial_j \phi\left(\frac{x-y}{t}\right) f(y) dy \\ &= \frac{1}{t^{d+1}} \int_{\mathbb{R}^d} \partial_j \phi\left(\frac{x-y}{t}\right) [f(y) - f(x)] dy \\ &= \frac{1}{t} \int_{\mathbb{R}^d} (\partial_j \phi)_t(x-y) [f(y) - f(x)] dy, \end{aligned}$$

since $\int_{\mathbb{R}^d} \partial_j \phi\left(\frac{x-y}{t}\right) dy = 0$ by an integration by parts. Arguing along similar lines as above, applying Jensen's inequality after normalising the measures $|(\partial_j \phi)_t(x-y)| dy$, we obtain

$$\begin{aligned} &\int_0^\infty t^{(1-s)p} \|\partial_j b_t\|_p^p \frac{dt}{t} \\ &= \int_0^\infty t^{(1-s)p} \int_{\mathbb{R}^d} \left\| \frac{1}{t} \int_{\mathbb{R}^d} (\partial_j \phi)_t(x-y) [f(y) - f(x)] dy \right\|^p dx \frac{dt}{t} \\ &\lesssim_{\phi,j} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left(\int_0^\infty t^{-sp} |(\partial_j \phi)_t(x-y)| \frac{dt}{t} \right) \|f(y) - f(x)\|^p dy dx \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left(\int_{|x-y|}^\infty t^{-sp-d} \left| \partial_j \phi\left(\frac{x-y}{t}\right) \right| \frac{dt}{t} \right) \|f(y) - f(x)\|^p dy dx \\ &\leq \frac{\|\partial_j \phi\|_\infty}{sp+d} [f]_{W^{s,p}(\mathbb{R}^d; X)}^p. \end{aligned}$$

□

Remark 2.5.18. We deliberately refrain from discussing the more general problem of characterising the interpolation spaces

$$(L^p(\mathbb{R}^d; X), W^{1,p}(\mathbb{R}^d; X))_{s,q}$$

for $p \neq q$ as this would take us into the realm of the Besov spaces. For similar reasons, we restrict ourselves to the Sobolev spaces over $\mathbb{R}^d$. A comprehensive treatment of Sobolev spaces over open domains $D \subseteq \mathbb{R}^d$ can be developed as well, but again their treatment requires the use of extension techniques that would take us too far afield. All these topics will be covered in Volume III.

Remark 2.5.19. The identification of the complex interpolation spaces

$$[L^p(\mathbb{R}^d; X), W^{1,p}(\mathbb{R}^d; X)]_s$$

for UMD spaces X will be presented in Chapter 5 as an application of the Mihlin multiplier theorem.

2.6 Conditional expectations

We now introduce and study an important class of operators, the conditional expectations, which will then accompany us for the rest of this volume. In particular, they are used to define the fundamental concept of a martingale in the following chapter.

In harmonic analysis, conditional expectations are used through the ubiquitous averaging procedures. In this context we are usually working over $\mathbb{R}^d$ endowed with the Lebesgue measure, and it is for this reason that we have chosen to give a unified presentation of the topic in the setting of σ -finite measure spaces.

In probability theory, it is customary to define conditional expectations only for integrable functions. This automatically implies that the conditional expectation of an L^p -function is well defined. In the σ -finite context this is no longer the case and for this reason we choose to work with strongly measurable functions which enjoy a minimal ‘local’ integrability condition needed to make the definition of a conditional expectation a meaningful one. Even in the case of probability spaces this leads to a conveniently general set-up.

Throughout this section we fix a measure space $(S, \mathcal{A}, \mu)$ and a Banach space X . We recall that $L^0(S; X)$ denotes the vector space of all strongly μ -measurable functions $f : S \rightarrow X$, identifying functions that are equal almost everywhere. When we wish to emphasise the underlying σ -algebra we shall write $L^0(S, \mathcal{A}; X)$. In particular, when $\mathcal{F} \subseteq \mathcal{A}$ is a sub- σ -algebra, the space $L^0(S, \mathcal{F}; X)$ is defined with reference to the measure space $(S, \mathcal{F}, \mu|_{\mathcal{F}})$. It has already been observed that this space coincides with the subspace of $L^0(S; X)$ of functions having a strongly $\mathcal{F}$ -measurable representative. The elements of $L^0(S, \mathcal{F}; X)$ will be referred to as the $\mathcal{F}$ -measurable elements of $L^0(S; X)$.

Given a sub- σ -algebra $\mathcal{F} \subseteq \mathcal{A}$, for a function $f \in L^0(S; X)$ we denote

$$\mathcal{F}_f := \{F \in \mathcal{F} : \mathbf{1}_F f \in L^1(S; X)\}.$$

Definition 2.6.1. A function $f \in L^0(S; X)$ is called σ -integrable over $\mathcal{F}$ if S can be covered by at most countably many sets in $\mathcal{F}_f$. Any covering sequence in $\mathcal{F}_f$ will be called an exhausting sequence for f in $\mathcal{F}$.

The sets in an exhausting sequence for f can be taken disjoint: if $(F_n)_{n \geq 1}$ is exhausting, then so is the sequence $(\tilde{F}_n)_{n \geq 1}$, where $\tilde{F}_1 = F_1$ and $\tilde{F}_{n+1} = F_{n+1} \setminus \bigcup_{j=1}^n F_j$, $n \geq 1$.

The σ -integrability of f with respect to $\mathcal{F}$ allows us to split the function f in integrable parts which respect the σ -algebra $\mathcal{F}$. This will enable us to build a conditional expectation without assuming global integrability on S .

The next two propositions provide easy examples of σ -integrability.

Proposition 2.6.2. *Every function $f \in L^0(S; X)$ is σ -integrable over $\mathcal{A}$.*

Proof. By Proposition 1.1.15 we have a disjoint decomposition $S = F_0 \cup F_1$ with $F_0, F_1 \in \mathcal{A}$ such that $f \equiv 0$ almost everywhere on F_0 and μ is σ -finite on F_1 . Let $(F^{(n)})_{n \geq 1}$ be an exhausting sequence for F_1 . For each $j \geq 1$ let $F_j = \{\|f\| \leq j\}$. Then f is integrable on F_0 and on each of the sets $F_{jn} := F^{(n)} \cap F_j$, $j, n \geq 1$; together, these sets cover S and therefore provide an exhausting sequence for f in $\mathcal{A}$. $\square$

Although every function $f \in L^0(S; X)$ is σ -integrable over $\mathcal{A}$, such functions need not to be σ -integrable over a sub- σ -algebra of $\mathcal{A}$. For example, this happens when $S = \mathbb{R}$ and $\mathcal{A} = \mathcal{B}(\mathbb{R})$ is the Borel σ -algebra; the constant function $\mathbf{1}$ is σ -integrable over $\mathcal{A}$ but not over the sub- σ -algebra $\mathcal{F} = \{\emptyset, \mathbb{R}\}$. This defect cannot be remedied by assuming that μ be σ -finite on $\mathcal{F}$: if f is a function which fails to be integrable on every interval $[n, n+1)$, then f fails to be σ -integrable over the σ -finite sub- σ -algebra generated by these intervals. If we add a global integrability assumption such problems do not arise:

Proposition 2.6.3. *If μ is σ -finite on the sub- σ -algebra $\mathcal{F}$ of $\mathcal{A}$ and if $1 \leq p \leq \infty$, then every function $f \in L^p(S; X)$ is σ -integrable over $\mathcal{F}$.*

Proof. Let $(S_n)_{n \geq 1}$ is an exhaustion of S by sets in $\mathcal{F}$ of finite μ -measure. The functions $\mathbf{1}_{S_n} f$ belong to $L^1(S; X)$ by Hölder's inequality, and therefore $(S_n)_{n \geq 1}$ is an exhausting sequence for f in $\mathcal{F}$. $\square$

Definition 2.6.4 (Conditional expectation). *Let $\mathcal{F} \subseteq \mathcal{A}$ be a sub- σ -algebra. A function $g \in L^0(S; X)$ is said to be a conditional expectation with respect to $\mathcal{F}$ of the function $f \in L^0(S; X)$ if*

$$\int_F g \, d\mu = \int_F f \, d\mu, \quad \text{for all } F \in \mathcal{F}_f \cap \mathcal{F}_g. \quad (2.24)$$

Without a σ -integrability assumption on f it may happen that $\mathcal{F}_f = \{\emptyset\}$, in which case the above definition is vacuous. Accordingly, in what follows we will usually assume f to be σ -integrable over $\mathcal{F}$. Among other things this assumption implies that a conditional expectation, if one exists, is unique (Theorem 2.6.18). In this situation we moreover have $\mathcal{F}_f \subseteq \mathcal{F}_g$ and therefore (2.24) holds for all $F \in \mathcal{F}_f$. If we furthermore assume that μ is σ -finite on $\mathcal{F}$, then f admits a unique conditional expectation g with respect to $\mathcal{F}$ (Theorem 2.6.20).

It will be convenient to have a criterion which enables us to verify the defining condition (2.24) by testing it on an appropriate sub-collection $\mathcal{C}$ of $\mathcal{F}_f \cap \mathcal{F}_g$. For this purpose we introduce the following definition.

Definition 2.6.5. Let $\mathcal{F} \subseteq \mathcal{A}$ be a sub- σ -algebra. A family $\mathcal{C} \subseteq \mathcal{F}$ will be called an *exhausting ideal* for $\mathcal{F}$ if it has the following two properties:

- (i) $C \cap F \in \mathcal{C}$ for all $C \in \mathcal{C}$ and $F \in \mathcal{F}$;
- (ii) S can be covered by at most countably many sets from $\mathcal{C}$.

Example 2.6.6. In a σ -finite measure space $(S, \mathcal{A}, \mu)$, the sets of finite μ -measure form an exhausting ideal for $\mathcal{A}$.

Any exhausting ideal contains a disjoint sequence covering S . Indeed, by (ii) there is a sequence $(C_n)_{n \geq 1}$ in $\mathcal{C}$ covering S . Set $C'_1 = C_1$ and $C'_{n+1} = C_{n+1} \setminus \bigcup_{j=1}^n C_j$. In view of $C'_{n+1} = C_{n+1} \cap \mathbb{C}(\bigcup_{j=1}^n C_j)$ and $\mathbb{C}(\bigcup_{j=1}^n C_j) \in \mathcal{F}$, (i) implies that $C'_{n+1} \in \mathcal{C}$. The sets C'_n are disjoint and cover S .

Exhausting ideals are closed under taking intersections:

Lemma 2.6.7. If $\mathcal{C}_1$ and $\mathcal{C}_2$ are exhausting ideals for $\mathcal{F}$, then $\mathcal{C}_1 \cap \mathcal{C}_2$ is a exhausting ideal for $\mathcal{F}$ and

$$\mathcal{C}_1 \cap \mathcal{C}_2 = \{C_1 \cap C_2 : C_1 \in \mathcal{C}_1, C_2 \in \mathcal{C}_2\}.$$

We leave the simple proof to the reader.

Example 2.6.8. The constant function $\mathbf{1}$ is σ -integrable over $\mathcal{F}$ if and only if μ is σ -finite on $\mathcal{F}$, if and only if $\mathcal{F}_1$ (i.e., $\mathcal{F}_h$ for the constant function $h = \mathbf{1}$) is an exhausting ideal for $\mathcal{F}$. More generally, a function $f \in L^0(S; X)$ is σ -integrable over $\mathcal{F}$ if and only if the measure $d\nu := \|f\| d\mu$ is σ -finite on $\mathcal{F}$, if and only if $\mathcal{F}_f$ is an exhausting ideal for $\mathcal{F}$. In particular, if $f \in L^0(S, \mathcal{F}; X)$, then $\mathcal{F}_f$ is an exhausting ideal for $\mathcal{F}$ by Proposition 2.6.2.

The next result simplifies the task of verifying condition (2.24) considerably.

Proposition 2.6.9. Let $\mathcal{F} \subseteq \mathcal{A}$ be a sub- σ -algebra, let $f \in L^0(S; X)$ be σ -integrable over $\mathcal{F}$, and let $g \in L^0(S, \mathcal{F}; X)$ be given. The following assertions are equivalent:

- (1) g is a conditional expectation of f with respect to $\mathcal{F}$;
- (2) there exists an exhausting ideal $\mathcal{C} \subseteq \mathcal{F}_f \cap \mathcal{F}_g$ for $\mathcal{F}$ such that

$$\int_C g d\mu = \int_C f d\mu \quad \text{for all } C \in \mathcal{C}. \quad (2.25)$$

Proof. (1) $\Rightarrow$ (2): It follows from Example 2.6.8 that both $\mathcal{F}_f$ and $\mathcal{F}_g$ are exhausting ideals for $\mathcal{F}$. By Lemma 2.6.7, $\mathcal{C} = \mathcal{F}_f \cap \mathcal{F}_g$ is an exhausting ideal for $\mathcal{F}$, and (2.25) holds by the definition of a conditional expectation.

(2) $\Rightarrow$ (1): Choose a cover $(C_n)_{n \geq 1}$ of S by pairwise disjoint sets and fix a set $F \in \mathcal{F}_f \cap \mathcal{F}_g$. It follows from (2.25) and the fact that $F \cap C_n \in \mathcal{C}$ that

$$\int_{C_n \cap F} g d\mu = \int_{C_n \cap F} f d\mu.$$

Summing over n , the identity (2.24) follows by dominated convergence. $\square$

Example 2.6.10. If f_1 and f_2 are σ -integrable and have conditional expectations g_1 and g_2 , then for all scalars c_1 and c_2 the function $c_1g_1 + c_2g_2$ is a conditional expectation for $c_1f_1 + c_2f_2$. This easily follows from Proposition 2.6.9, noting that $\mathcal{C} := \mathcal{F}_{f_1} \cap \mathcal{F}_{f_2} \cap \mathcal{F}_{g_1} \cap \mathcal{F}_{g_2}$ is an exhausting ideal for $\mathcal{F}$ by Lemma 2.6.7. This example clearly shows the flexibility of the notion of an exhausting ideal. Working with the ‘maximal’ exhausting ideal $\mathcal{F}_{c_1f_1+c_2f_2} \cap \mathcal{F}_{c_1g_1+c_2g_2}$ would make the exercise much harder!

Example 2.6.11 (Restriction to sets in $\mathcal{F}$). If $f \in L^0(S; X)$ admits a conditional expectation g with respect to $\mathcal{F}$, then for every $A \in \mathcal{F}$, $\mathbf{1}_A g$ is a conditional expectation of $\mathbf{1}_A f$ with respect to $\mathcal{F}$. Indeed, for every $F \in \mathcal{F}_f \cap \mathcal{F}_g$ also $F \cap A \in \mathcal{F}_f \cap \mathcal{F}_g$ and

$$\int_F \mathbf{1}_A f \, d\mu = \int_{F \cap A} f \, d\mu = \int_{F \cap A} g \, d\mu = \int_F \mathbf{1}_A g \, d\mu.$$

Example 2.6.12 (Applying bounded operators). If $f \in L^0(S; X)$ admits a conditional expectation g with respect to $\mathcal{F}$, and if $T : X \rightarrow Y$ is a bounded operator, then Tg is a conditional expectation of Tf with respect to $\mathcal{F}$. The proof is similar to the previous one; this time one uses that $\mathcal{F}_f \cap \mathcal{F}_g$ is contained in $\mathcal{F}_{Tf} \cap \mathcal{F}_{Tg}$.

In particular it follows that for every $x^* \in X^*$ the function $\langle g, x^* \rangle$ is a conditional expectation of $\langle f, x^* \rangle$.

Example 2.6.13 (Averaging). For each $n \in \mathbb{Z}$, let $\mathcal{F}_n$ be the σ -algebra in $\mathbb{R}^d$ generated by the family of dyadic cubes $Q_k^n = 2^{-n}(k + [0, 1)^d)$, $k \in \mathbb{Z}^d$. A function $f \in L^0(\mathbb{R}^d; X)$ is σ -integrable with respect to $\mathcal{F}_n$ if and only if $f \in L^1_{\text{loc}}(\mathbb{R}^d; X)$. In this case, a conditional expectation of f with respect to $\mathcal{F}_n$ exists and is given by

$$\sum_{k \in \mathbb{Z}^d} \left(\frac{1}{|Q_k^n|} \int_{Q_k^n} f(x) \, dx \right) \mathbf{1}_{Q_k^n}.$$

Example 2.6.14 (Even part of a function). Let $f \in L^0(\mathbb{R}; X)$, where on $\mathbb{R}$ we take the Borel σ -algebra $\mathcal{B}(\mathbb{R})$ and the Lebesgue measure. Let $\mathcal{F} = \{B \in \mathcal{B}(\mathbb{R}) : -B = B\}$. Then $g : \mathbb{R} \rightarrow X$ defined by $g(t) = \frac{1}{2}(f(t) + f(-t))$ is a conditional expectation of f given $\mathcal{F}$. Indeed, $g \in L^0(S, \mathcal{F}; X)$, $\mathcal{F}_f \subseteq \mathcal{F}_g$, and for any set in $B \in \mathcal{F}_f = \mathcal{F}_f \cap \mathcal{F}_g$ we have

$$\int_B f(t) \, dt = \int_B f(-t) \, dt = \int_B g(t) \, dt.$$

2.6.a Uniqueness

Our task in this section is to prove the uniqueness of conditional expectations. We begin with some lemmas concerning scalar-valued functions.

We fix a measure space $(S, \mathcal{A}, \mu)$ and let $\mathcal{F} \subseteq \mathcal{A}$ be sub- σ -algebra.

Lemma 2.6.15. *Let $f \in L^0(S, \mathcal{F})$ be given and suppose that $\mathcal{C} \subseteq \mathcal{F}_f$ is an exhausting ideal for $\mathcal{F}$. If $\int_C f \, d\mu \geq 0$ (respectively, $= 0$, ≤ 0) for all $C \in \mathcal{C}$, then $f \geq 0$ (respectively, $= 0$, ≤ 0) almost everywhere.*

Proof. For any $\mathcal{F}$ -measurable representative of f we have $F_k := \{f < -1/k\} \in \mathcal{F}$ for all $k \geq 1$. Fix $C \in \mathcal{C}$. Then $C \cap F_k \in \mathcal{C}$, the integrability of f on this set implies $\mu(C \cap F_k) < \infty$, and

$$0 \leq \int_{C \cap F_k} f \, d\mu \leq -\frac{1}{k} \mu(C \cap F_k) \leq 0.$$

Hence $\mu(C \cap F_k) = 0$. Since $\{f < 0\} = \bigcup_{k \geq 1} F_k$, we see that $\mu(C \cap \{f < 0\}) = 0$. Applying this to the sets C in a countable cover of S , we see that $\mu(\{f < 0\}) = 0$, which is the same as saying that $f \geq 0$ almost everywhere.

The assertion concerning ' ≤ 0 ' follows by applying the previous case to $-f$, and the assertion concerning ' $= 0$ ' follows by combining the two previous cases. $\square$

Lemma 2.6.16. *Let $f \in L^0(S)$ be σ -integrable over $\mathcal{F}$. If $g \in L^0(S)$ is a conditional expectation of f with respect to $\mathcal{F}$, then $\mathcal{F}_f \subseteq \mathcal{F}_g$ and*

$$\int_F f \, d\mu = \int_F g \, d\mu \quad \forall F \in \mathcal{F}_f. \quad (2.26)$$

Proof. In the complex case we have $\mathcal{F}_f = \mathcal{F}_{\Re f} \cap \mathcal{F}_{\Im f}$ and $\mathcal{F}_g = \mathcal{F}_{\Re g} \cap \mathcal{F}_{\Im g}$, from which it is easy to see that $\Re g$ and $\Im g$ are conditional expectations of $\Re f$ and $\Im f$, respectively. It suffices therefore to prove the lemma over the real scalars.

By Proposition 2.6.9, $\mathcal{C} = \mathcal{F}_f \cap \mathcal{F}_g$ is an exhausting ideal for $\mathcal{F}$ and therefore we can choose a sequence $(C_n)_{n \geq 1}$ in $\mathcal{C}$ covering S . By the observation made after Definition 2.6.5, we may assume that the sets C_n are pairwise disjoint.

Fix a set $F \in \mathcal{F}_f$. Choosing an $\mathcal{F}$ -measurable representative of g , we have $\{g \geq 0\} \in \mathcal{F}$, so that $F \cap C_n \cap \{g \geq 0\} \in \mathcal{C}$ and

$$\int_{F \cap C_n \cap \{g \geq 0\}} g \, d\mu = \int_{F \cap C_n \cap \{g \geq 0\}} f \, d\mu.$$

Summing over n , we obtain (by monotone convergence on the left-hand side and dominated convergence on the right-hand side)

$$\int_{F \cap \{g \geq 0\}} g \, d\mu = \int_{F \cap \{g \geq 0\}} f \, d\mu.$$

In the same way we prove that $\int_{F \cap \{g < 0\}} g \, d\mu = \int_{F \cap \{g < 0\}} f \, d\mu$. Therefore,

$$\int_F |g| \, d\mu = \int_{F \cap \{g \geq 0\}} g \, d\mu - \int_{F \cap \{g < 0\}} g \, d\mu$$

$$= \int_{F \cap \{g \geq 0\}} f \, d\mu - \int_{F \cap \{g < 0\}} f \, d\mu \leq \int_F |f| \, d\mu < \infty$$

and

$$\int_F g \, d\mu = \int_{F \cap \{g \geq 0\}} g \, d\mu + \int_{F \cap \{g < 0\}} g \, d\mu = \int_F f \, d\mu.$$

This first relation shows that $F \in \mathcal{F}_g$ and the second proves (2.26). $\square$

Before turning to the proof of the uniqueness of conditional expectations we record a useful consequence of the previous two lemmas.

Lemma 2.6.17. *Let $f_1, f_2 \in L^0(S)$ be σ -integrable over $\mathcal{F}$ and $g_1, g_2 \in L^0(S)$ be conditional expectations of these functions with respect to $\mathcal{F}$. If $f_1 \leq f_2$ almost everywhere, then $g_1 \leq g_2$ almost everywhere.*

Proof. Under the stated assumptions, $g := g_2 - g_1$ is a conditional expectation of the non-negative function $f := f_2 - f_1 \geq 0$, and it suffices to prove that g is non-negative.

By Lemma 2.6.16, for all $F \in \mathcal{F}_f$ we have $F \in \mathcal{F}_g$ and

$$\int_F g \, d\mu = \int_F f \, d\mu \geq 0.$$

We now apply Lemma 2.6.15 (to g and $\mathcal{C} = \mathcal{F}_f$) to conclude that $g \geq 0$ almost everywhere. $\square$

Theorem 2.6.18 (Uniqueness of conditional expectations). *Suppose that $f \in L^0(S; X)$ is σ -integrable with respect to $\mathcal{F}$. If $g \in L^0(S; X)$ and $\tilde{g} \in L^0(S; X)$ are conditional expectations of f with respect to $\mathcal{F}$, then $g = \tilde{g}$ almost everywhere.*

Proof. For any $x^* \in X^*$, both $\langle g, x^* \rangle$ and $\langle \tilde{g}, x^* \rangle$ are conditional expectations of $\langle f, x^* \rangle$. Moreover, for all $F \in \mathcal{F}_{\langle f, x^* \rangle}$ we have $F \in \mathcal{F}_{\langle g, x^* \rangle} \cap \mathcal{F}_{\langle \tilde{g}, x^* \rangle}$ by Lemma 2.6.16, and therefore

$$\int_F \langle g, x^* \rangle \, d\mu = \int_F \langle f, x^* \rangle \, d\mu = \int_F \langle \tilde{g}, x^* \rangle \, d\mu.$$

Hence $\langle g, x^* \rangle = \langle \tilde{g}, x^* \rangle$ almost everywhere by Lemma 2.6.15 (applied to $\langle g - \tilde{g}, x^* \rangle$ and $\mathcal{C} = \mathcal{F}_{\langle f, x^* \rangle}$). An appeal to Corollary 1.1.25 concludes the proof. $\square$

Thus a conditional expectation of a σ -integrable function $f \in L^0(S; X)$, if it exists, is unique as an element of $L^0(S, \mathcal{F}; X)$. This allows us to speak of *the* conditional expectation of f with respect to $\mathcal{F}$.

Notation. The conditional expectation of a σ -integrable function $f \in L^0(S; X)$ with respect to $\mathcal{F}$, if it exists, will be denoted by $\mathbb{E}(f|\mathcal{F})$.

For later use we state the following simple lemma.

Lemma 2.6.19. *Let $f \in L^0(S; X)$ be σ -integrable over $\mathcal{F} \subseteq \mathcal{A}$. If the conditional expectations of f and $\|f\|$ with respect to $\mathcal{F}$ exist, then almost everywhere we have*

$$\|\mathbb{E}(f|\mathcal{F})\| \leq \mathbb{E}(\|f\||\mathcal{F}).$$

Proof. By forgetting the complex scalar multiplication we may assume that $\mathbb{K} = \mathbb{R}$. Since a strongly measurable function takes its values in a separable subspace almost everywhere, we may assume that X is separable. Pick a norming sequence $(x_n^*)_{n \geq 1}$ of unit vectors in X^* .

By Example 2.6.12, the conditional expectation of $\langle f, x_n^* \rangle$ exists for every $n \geq 1$, and

$$\|\mathbb{E}(f|\mathcal{F})\| = \sup_{n \geq 1} \langle \mathbb{E}(f|\mathcal{F}), x_n^* \rangle = \sup_{n \geq 1} \mathbb{E}(\langle f, x_n^* \rangle | \mathcal{F}) \leq \mathbb{E}(\|f\||\mathcal{F})$$

almost everywhere, where the last step follows from Lemma 2.6.17 since $\langle f, x_n^* \rangle \leq \|f\|$ almost everywhere. $\square$

2.6.b Existence

We now turn to the existence problem for conditional expectations. As always we fix a measure space $(S, \mathcal{A}, \mu)$ and let $\mathcal{F} \subseteq \mathcal{A}$ be sub- σ -algebra.

Theorem 2.6.20 (Existence of conditional expectations). *If the function $f \in L^0(S; X)$ is σ -integrable over $\mathcal{F}$ and μ is σ -finite on the sub- σ -algebra $\mathcal{F}$, then f admits a conditional expectation with respect to $\mathcal{F}$. In this case, the conditional expectation is uniquely defined as an element of $L^0(S, \mathcal{F}; X)$. Denoting it by $\mathbb{E}(f|\mathcal{F})$, we have $\mathcal{F}_f \subseteq \mathcal{F}_{\mathbb{E}(f|\mathcal{F})}$ and*

$$\int_F f \, d\mu = \int_F \mathbb{E}(f|\mathcal{F}) \, d\mu \quad \forall F \in \mathcal{F}_f. \quad (2.27)$$

Furthermore, if $(F_n)_{n \geq 1}$ is an exhausting sequence for f in $\mathcal{F}$ consisting of disjoint sets, then

$$\sum_{n \geq 1} \mathbf{1}_{F_n} \mathbb{E}(\mathbf{1}_{F_n} f | \mathcal{F}). \quad (2.28)$$

As we have already noted in the previous subsection, the σ -integrability over $\mathcal{F}$ of the constant function $\mathbf{1}$ is equivalent to the σ -finiteness of μ on $\mathcal{F}$. It is for this reason that in the sequel we only consider conditional expectations with respect to σ -finite σ -algebras. We will also provide an example of the anomalies that arise in the absence of this assumption.

The proof of Theorem 2.6.20 will be accomplished in several steps. The starting point is provided by the next lemma.

Lemma 2.6.21 (Existence in $L^2(S)$). *If μ is σ -finite on the sub- σ -algebra $\mathcal{F}$, then every $f \in L^2(S)$ has a unique conditional expectation with respect to $\mathcal{F}$. It belongs to $L^2(S)$ and is given by Pf , where P be the orthogonal projection of $L^2(S)$ onto $L^2(S, \mathcal{F})$.*

Proof. Since μ is assumed to be σ -finite on $\mathcal{F}$, the family $\mathcal{F}_1$ (i.e., $\mathcal{F}_h$ for the constant function $h = \mathbf{1}$) is an exhausting ideal for $\mathcal{F}$. Moreover, $\mathcal{F}_1 \subseteq \mathcal{F}_f \cap \mathcal{F}_{Pf}$ by the fact that $f, Pf \in L^2(S)$ and the Cauchy–Schwarz inequality. Finally, for $F \in \mathcal{F}_1$, using that orthogonal projections are self-adjoint, we have

$$\int_F Pf \, d\mu = \int_S \mathbf{1}_F(Pf) \, d\mu = \int_S (P\mathbf{1}_F)f \, d\mu = \int_S \mathbf{1}_F f \, d\mu = \int_F f \, d\mu.$$

This completes the proof. $\square$

Lemma 2.6.22 (Existence in $L^1(S)$). *If μ is σ -finite on the sub- σ -algebra $\mathcal{F}$, then every $f \in L^1(S)$ has a unique conditional expectation with respect to $\mathcal{F}$. It belongs to $L^1(S)$ and satisfies*

$$\|\mathbb{E}(f|\mathcal{F})\|_1 \leq \|f\|_1$$

and

$$\int_F \mathbb{E}(f|\mathcal{F}) \, d\mu = \int_F f \, d\mu \quad \forall F \in \mathcal{F}.$$

Proof. First let $f \in L^1(S) \cap L^\infty(S)$. Then $f, |f| \in L^2(S)$, so their conditional expectations exist by Lemma 2.6.21, and Lemma 2.6.19 shows that

$$\int_S |\mathbb{E}(f|\mathcal{F})| \, d\mu \leq \int_S \mathbb{E}(|f|) \, d\mu = \int_S |f| \, d\mu,$$

where the last step follows from (2.26) and the fact that $S \in \mathcal{F}_{|f|}$ for $f \in L^1(S)$.

This proves that the conditional expectation is contractive with respect to the L^1 -norm on the dense subspace $L^1(S) \cap L^\infty(S)$ of $L^1(S)$. As a consequence, this operator admits a unique extension to a contraction, denoted by P , on $L^1(S)$ which takes its values in the closed subspace $L^1(S, \mathcal{F})$ of $L^1(S)$. To see that for any $f \in L^1(S)$, Pf is the conditional expectation of f , we set $f_n = \mathbf{1}_{\{|f| \leq n\}} f$. Then $f_n \in L^1(S) \cap L^\infty(S)$, $f_n \rightarrow f$ in $L^1(S)$, and for all $F \in \mathcal{F}$ we have

$$\begin{aligned} \int_F Pf \, d\mu &= \lim_{n \rightarrow \infty} \int_F Pf_n \, d\mu = \lim_{n \rightarrow \infty} \int_F \mathbb{E}(f_n|\mathcal{F}) \, d\mu \\ &= \lim_{n \rightarrow \infty} \int_F f_n \, d\mu = \int_F f \, d\mu, \end{aligned}$$

where the second-to-last step used (2.26) and the fact that $\mathcal{F}_{f_n} = \mathcal{F}$. $\square$

After these scalar-valued auxiliary results, we return to the realm of vector-valued functions:

Theorem 2.6.23 (Existence in $L^1(S; X)$). *If μ is σ -finite on the sub- σ -algebra $\mathcal{F}$, then every $f \in L^1(S; X)$ admits a unique conditional expectation with respect to $\mathcal{F}$. It belongs to $L^1(S; X)$ and satisfies*

$$\|\mathbb{E}(f|\mathcal{F})\|_1 \leq \|f\|_1$$

and

$$\int_F \mathbb{E}(f|\mathcal{F}) \, d\mu = \int_F f \, d\mu \quad \forall F \in \mathcal{F}.$$

Proof. By Lemma 2.6.22, the operator $\mathbb{E}(\cdot|\mathcal{F})$ is a well defined contraction on $L^1(S)$, which is positive by Lemma 2.6.17. By Theorem 2.1.3, it extends to a contraction on $L^1(S; X)$. To see that this extension is indeed the conditional expectation on $L^1(S; X)$ we note that for μ -simple functions this follows from the scalar case (it is immediate from the definition of conditional expectations that $\mathbb{E}(g \otimes x|\mathcal{F}) = \mathbb{E}(g|\mathcal{F}) \otimes x$), and for the general case we approximate f by a sequence of μ -simple functions f_n and argue as in the proof of Lemma 2.6.22.

The final identity follows from the scalar case by a similar approximation, noting that it is true for μ -simple functions since it is true for scalar functions by Lemma 2.6.22. $\square$

Proof of Theorem 2.6.20. First we prove that if μ is σ -finite on the sub- σ -algebra $\mathcal{F}$, then every $f \in L^0(S; X)$ that is σ -integrable on $\mathcal{F}$ admits a conditional expectation with respect to $\mathcal{F}$.

Fix an arbitrary exhausting sequence $(F_n)_{n \geq 1}$ for f in $\mathcal{F}$ consisting of disjoint sets. Since integrable functions have a conditional expectation by Theorem 2.6.23, the function

$$\sum_{n \geq 1} \mathbf{1}_{F_n} \mathbb{E}(\mathbf{1}_{F_n} f|\mathcal{F})$$

is well defined. By Proposition 2.6.9 $\mathcal{C}_n = \mathcal{F}|_{F_n} \cap \mathcal{F}|_{\mathbb{E}(\mathbf{1}_{F_n} f|\mathcal{F})}$ is an exhausting ideal in $\mathcal{F}$. It is easily checked that $\mathcal{C} := \bigcup_{n \geq 1} \mathcal{C}_n|_{F_n}$ is an exhausting ideal in $\mathcal{F}$, where $\mathcal{C}_n|_{F_n} := \{C \cap F_n : C \in \mathcal{C}_n\}$. Proposition 2.6.9 then implies that that the function just defined is a conditional expectation of f .

The proof of Theorem 2.6.20 is completed by checking that the assertion (2.27) holds true. Let $F \in \mathcal{F}_f$, so that $\mathbf{1}_F f \in L^1(S; X)$ by definition. Theorem 2.6.23 guarantees that this function has a conditional expectation $h \in L^1(S; X)$ and

$$\int_S h \, d\mu = \int_S \mathbf{1}_F f \, d\mu = \int_F f \, d\mu.$$

On the other hand, by Example 2.6.11, the conditional expectation of $\mathbf{1}_F f$, which we denoted by h , is given by $\mathbf{1}_F \mathbb{E}(f|\mathcal{F})$. So in fact we checked that $\mathbf{1}_F \mathbb{E}(f|\mathcal{F}) \in L^1(S; X)$ and

$$\int_F \mathbb{E}(f|\mathcal{F}) \, d\mu = \int_S h \, d\mu = \int_F f \, d\mu,$$

as we wanted to prove. $\square$

Examples of non-existence

We now give an example to demonstrate the failure of existence of conditional expectations, even for functions $f \in L^1(S)$ (which is σ -integrable over any sub- σ -algebra), when μ is not assumed to be σ -finite on $\mathcal{F}$.

Example 2.6.24. Let S be a set with a distinct point s_1 and equipped with a measure μ such that $\mu\{s_1\} = 1$ and $\mu(S \setminus \{s_1\}) = \infty$. The function $f(s_1) = a$, $f(s) = 0$ for $s \neq s_1$ is σ -integrable over $\mathcal{F} = \{\emptyset, S\}$ but has a conditional expectation with respect to $\mathcal{F}$ if and only if $a = 0$ (in which case the conditional expectation is the zero function). Indeed, the conditional expectation has to be a constant function which, by the requirement of strong μ -measurability, can only be the zero function. For $a \neq 0$ the defining condition (2.24) for conditional expectations is clearly violated.

2.6.c Conditional limit theorems

We proceed with “conditional” versions of the fundamental limit theorems of integration theory. Their proofs follow the same logical pattern as for their unconditional counterparts: we start with monotone convergence and arrive, via Fatou’s lemma, at dominated convergence.

Theorem 2.6.25 (Conditional monotone convergence). *Let $(f_n)_{n \geq 1}$ be a non-decreasing sequence of non-negative functions in $L^0(S)$ and let $f \in L^0(S)$ be σ -integrable over a sub- σ -algebra $\mathcal{F}$ on which μ is σ -finite. If $f_n \uparrow f$ almost everywhere, then*

$$\mathbb{E}(f_n|\mathcal{F}) \uparrow \mathbb{E}(f|\mathcal{F}),$$

the existence of the conditional expectations being part of the assertion.

Proof. The σ -integrability of f over $\mathcal{F}$ implies the same property for each f_n , and hence the existence of the conditional expectations is guaranteed by Theorem 2.6.20.

By monotonicity and the positivity of the conditional expectation operator it follows that we have the almost everywhere inequalities

$$0 \leq \mathbb{E}(f_n|\mathcal{F}) \leq \mathbb{E}(f_{n+1}|\mathcal{F}) \leq \mathbb{E}(f|\mathcal{F}).$$

Hence $(\mathbb{E}(f_n|\mathcal{F}))_{n \geq 1}$ is an almost everywhere bounded non-decreasing sequence, so it has an almost everywhere pointwise limit $g \in L^0(S, \mathcal{F})$ which satisfies $0 \leq g \leq \mathbb{E}(f|\mathcal{F})$. It remains to be shown that $g = \mathbb{E}(f|\mathcal{F})$.

For all $F \in \mathcal{F}_f$ we have $F \in \mathcal{F}_{f_n}$ and

$$\int_F g \, d\mu = \lim_{n \rightarrow \infty} \int_F \mathbb{E}(f_n|\mathcal{F}) \, d\mu = \lim_{n \rightarrow \infty} \int_F f_n \, d\mu = \int_F f \, d\mu,$$

where the first and last steps were based on the usual monotone convergence theorem. This gives the desired result. $\square$

The σ -integrability assumption on f cannot be omitted:

Example 2.6.26. Let $S = (0, 1)$ with the Borel σ -algebra and Lebesgue measure, let $\mathcal{F} = \{\emptyset, (0, 1)\}$, and consider the functions $f(t) = 1/t$ and $f_n(t) = \mathbf{1}_{(1/n, 1)} f$. Each f_n has a conditional expectation with respect to $\mathcal{F}$, but f does not.

Theorem 2.6.27 (Conditional Fatou lemma). *Let μ be σ -finite on the sub- σ -algebra $\mathcal{F}$, let $(f_n)_{n \geq 1}$ be a sequence of non-negative functions in $L^0(S)$, let $f \in L^0(S)$, and assume that all these functions are σ -integrable over $\mathcal{F}$. Suppose furthermore that $f = \liminf_{n \rightarrow \infty} f_n$ almost everywhere. Then, almost everywhere,*

$$\mathbb{E}(f|\mathcal{F}) \leq \liminf_{n \rightarrow \infty} \mathbb{E}(f_n|\mathcal{F}).$$

Proof. Putting $\phi_n := \inf_{m \geq n} f_m$ we have $0 \leq \phi_n \uparrow f = \sup_{n \geq 1} \inf_{m \geq n} f_m$. By the conditional monotone convergence theorem,

$$\begin{aligned} \mathbb{E}(f|\mathcal{F}) &= \mathbb{E}\left(\lim_{n \rightarrow \infty} \phi_n \middle| \mathcal{F}\right) \\ &= \lim_{n \rightarrow \infty} \mathbb{E}(\phi_n|\mathcal{F}) \leq \lim_{n \rightarrow \infty} \inf_{m \geq n} \mathbb{E}(f_m|\mathcal{F}) = \liminf_{n \rightarrow \infty} \mathbb{E}(f_n|\mathcal{F}). \end{aligned}$$

The estimate $\leq$ above was based on the fact that $\phi_n \leq f_m$ for all $m \geq n$ and hence $\mathbb{E}(\phi_n|\mathcal{F}) \leq \mathbb{E}(f_m|\mathcal{F})$ for all $m \geq n$. $\square$

Theorem 2.6.28 (Conditional dominated convergence). *Let μ be σ -finite on the sub- σ -algebra $\mathcal{F}$. Let $(f_n)_{n \geq 1}$ be a sequence in $L^0(S; X)$ and suppose that $\lim_{n \rightarrow \infty} f_n = f$ and $\|f_n\| \leq g$ almost everywhere, where $g \in L^0(S)$ is σ -integrable over $\mathcal{F}$. Then, almost everywhere*

$$\lim_{n \rightarrow \infty} \mathbb{E}(\|f_n - f\||\mathcal{F}) = 0,$$

and

$$\lim_{n \rightarrow \infty} \mathbb{E}(f_n|\mathcal{F}) = \mathbb{E}(f|\mathcal{F}),$$

the existence of these conditional expectations being part of the assertion.

Proof. The existence of the conditional expectations of f and f_n follows from Theorem 2.6.20 and the fact that they, too, must be σ -integrable over $\mathcal{F}$.

Let $g_n := \|f_n - f\|$. Then almost everywhere we have $g_n \rightarrow 0$ and $0 \leq g_n \leq 2g$. Let $\psi_n = \sup_{m \geq n} g_m$. Then $0 \leq \psi_n \leq 2g$ and $\psi_n \downarrow 0$ almost everywhere. Hence $\mathbb{E}(2g - \psi_n|\mathcal{F}) \uparrow \mathbb{E}(2g|\mathcal{F})$ almost everywhere by the conditional monotone convergence theorem and therefore $\mathbb{E}(\psi_n|\mathcal{F}) \downarrow 0$ almost everywhere as $n \rightarrow \infty$. In view of $0 \leq g_n \leq \psi_n$, this proves that $\mathbb{E}(g_n|\mathcal{F}) \downarrow 0$ almost everywhere. Now the final assertion can be derived if one uses Lemma 2.6.19. $\square$

2.6.d Inequalities and identities

The following key inequality is the conditional analogue of Proposition 1.2.11 and extends Lemma 2.6.19.

Proposition 2.6.29 (Conditional Jensen's inequality). *Let μ be σ -finite on the sub- σ -algebra $\mathcal{F}$. Let $\phi : X \rightarrow \mathbb{R}$ be convex and lower semi continuous, let $f \in L^0(S; X)$, and suppose that f and $\phi \circ f$ are σ -integrable over $\mathcal{F}$. Then, almost everywhere,*

$$\phi \circ \mathbb{E}(f|\mathcal{F}) \leq \mathbb{E}(\phi \circ f|\mathcal{F}).$$

Proof. Since f takes its values in a separable subspace almost everywhere, we may assume that X is separable. Then Lemma 1.2.10 provides a sequence of affine functions $\phi_i(x) = \Re \langle x, y_i^* \rangle + a_i$ on X such that $\phi(x) = \sup_i \phi_i(x)$. Thus, almost everywhere,

$$\phi \circ \mathbb{E}(f|\mathcal{F}) = \sup_i \phi_i \circ \mathbb{E}(f|\mathcal{F}) = \sup_i \mathbb{E}(\phi_i \circ f|\mathcal{F}) \leq \mathbb{E}(\phi \circ f|\mathcal{F}).$$

□

Corollary 2.6.30 (L^p -contractivity). *Let μ be σ -finite on the sub- σ -algebra $\mathcal{F}$. Let $p \in [1, \infty)$ and suppose that $f \in L^p(S; X)$. Then the conditional expectations below exist, and satisfy the almost everywhere inequality*

$$\|\mathbb{E}(f|\mathcal{F})\|^p \leq \mathbb{E}(\|f\|^p|\mathcal{F}). \quad (2.29)$$

In particular, we have $\mathbb{E}(f|\mathcal{F}) \in L^p(S; X)$ and

$$\|\mathbb{E}(f|\mathcal{F})\|_p \leq \|f\|_p.$$

Proof. Both f and $\|f\|^p$ are σ -integrable over $\mathcal{F}$ (see Proposition 2.6.3). Therefore, the conditional expectations exist. Now (2.29) follows from Jensen's inequality applied to $\phi(x) = \|x\|^p$, and the norm inequality follows by integrating (2.29) over S and applying Lemma 2.6.22 on the right. □

We continue with some useful identities for conditional expectations. In Propositions 2.6.31 and 2.6.32 we let X_1, X_2, Y be Banach spaces and

$$\beta : X_1 \times X_2 \rightarrow Y$$

be a bounded bilinear mapping. The main examples we have in mind are

$$\begin{aligned} X_1 = \mathbb{K}, \quad X_2 = X, \quad Y = X, \quad \beta(c, x) &= cx && \text{(scalar multiplication);} \\ X_1 = X, \quad X_2 = X^*, \quad Y = \mathbb{K}, \quad \beta(x, x^*) &= \langle x, x^* \rangle && \text{(duality).} \end{aligned}$$

Proposition 2.6.31 (Taking out $\mathcal{F}$ -measurable terms). *Let μ be σ -finite on the sub- σ -algebra $\mathcal{F}$. Suppose that $g \in L^0(S; \mathcal{F}; X_1)$ and that $f \in L^0(S; X_2)$ is σ -integrable over $\mathcal{F}$. Then $\beta(g, f) \in L^0(S; Y)$ is σ -integrable over $\mathcal{F}$, and almost everywhere we have*

$$\mathbb{E}(\beta(g, f)|\mathcal{F}) = \beta(g, \mathbb{E}(f|\mathcal{F})).$$

Proof. Choosing an exhausting sequence of sets $F_i \in \mathcal{F}_f$, which exist by σ -integrability, the function $\beta(g, f)$ is integrable over any of the sets $F_i \cap \{\|g\| \leq j\} \in \mathcal{F}$, which form a countable cover of S . Thus $\beta(g, f)$ is σ -integrable over $\mathcal{F}$, which implies the existence of the conditional expectation by Theorem 2.6.20.

To prove the identity, let us first consider the case that $g = \sum_{i=1}^n x_i \mathbf{1}_{F_i}$ is simple. Then, for $F \in \mathcal{F}_f \subseteq \mathcal{F}_{\mathbb{E}(f|\mathcal{F})}$ we simply compute

$$\begin{aligned} \int_F \beta(g, \mathbb{E}(f|\mathcal{F})) \, d\mu &= \sum_{i=1}^n \beta\left(x_i, \int_{F \cap F_i} \mathbb{E}(f|\mathcal{F}) \, d\mu\right) \\ &= \sum_{i=1}^n \beta\left(x_i, \int_{F \cap F_i} f \, d\mu\right) = \int_F \beta(g, f) \, d\mu, \end{aligned}$$

which shows by definition that $\beta(g, \mathbb{E}(f|\mathcal{F}))$ is a conditional expectation of $\beta(f, g)$ as $\mathcal{F}_f$ is an exhausting ideal for $\mathcal{F}$.

In the general case, we pick a sequence of simple functions

$$g_n \in L^0(S; \mathcal{F}; X_1)$$

such that $g_n \rightarrow g$ pointwise and $\|g_n\| \leq \|g\|$. Then $\beta(g_n, h) \rightarrow \beta(g, h)$ pointwise for $h \in \{f, \mathbb{E}(f|\mathcal{F})\}$, and

$$\|\beta(g_n, f)\| \leq \|\beta\| \cdot \|g_n\| \cdot \|f\| \leq \|\beta\| \cdot \|g\| \cdot \|f\|,$$

which is a σ -integrable over $\mathcal{F}$, as shown by the sets $F_i \cap \{\|g\| \leq j\}$ from the beginning of the proof. Thus the conditional dominated convergence theorem proves that

$$\begin{aligned} \mathbb{E}(\beta(g, f)|\mathcal{F}) &= \lim_{n \rightarrow \infty} \mathbb{E}(\beta(g_n, f)|\mathcal{F}) \\ &= \lim_{n \rightarrow \infty} \beta(g_n, \mathbb{E}(f|\mathcal{F})) = \beta(g, \mathbb{E}(f|\mathcal{F})). \end{aligned}$$

□

Another version of this result, where the function g is allowed to be operator-valued, will be proved in the next chapter (Lemma 3.5.2).

As an application we prove that the conditional expectation has a useful self-adjointness property.

Proposition 2.6.32 (Self-adjointness property). *Suppose that μ is σ -finite on the sub- σ -algebra $\mathcal{F}$. Let $f_1 \in L^0(S; X_1)$ and $f_2 \in L^0(S; X_2)$ be σ -integrable over $\mathcal{F}$. If both $\beta(f_1, \mathbb{E}(f_2|\mathcal{F}))$ and $\beta(\mathbb{E}(f_1|\mathcal{F}), f_2)$ are integrable on some $F \in \mathcal{F}$, then so is $\beta(\mathbb{E}(f_1|\mathcal{F}), \mathbb{E}(f_2|\mathcal{F}))$, and we have*

$$\int_F \beta(f_1, \mathbb{E}(f_2|\mathcal{F})) \, d\mu = \int_F \beta(\mathbb{E}(f_1|\mathcal{F}), f_2) \, d\mu = \int_F \beta(\mathbb{E}(f_1|\mathcal{F}), \mathbb{E}(f_2|\mathcal{F})) \, d\mu.$$

Proof. Denote $g_i := \mathbb{E}(f_i|\mathcal{F})$ for $i = 1, 2$. By Proposition 2.6.31, $\beta(f_1, g_2)$ has a conditional expectation with respect to $\mathcal{F}$ given by

$$\mathbb{E}(\beta(f_1, g_2)|\mathcal{F}) = \beta(g_1, g_2).$$

The left side, and thus the right, is integrable over $F \in \mathcal{F}_{\beta(f_1, g_2)} \cap \mathcal{F}_{\beta(g_1, f_2)}$, and we have

$$\int_F \beta(f_1, g_2) d\mu = \int_F \mathbb{E}(\beta(f_1, g_2)|\mathcal{F}) d\mu = \int_F \beta(g_1, g_2) d\mu.$$

The other identity follows by exchanging the roles of f_1 and f_2 , observing that the right side is symmetric with respect to both f_1 and f_2 . $\square$

Proposition 2.6.33 (Tower property). *Let μ be σ -finite on $\mathcal{G}$, where $\mathcal{G} \subseteq \mathcal{F}$ are sub- σ -algebras. If $h \in L^0(S; X)$ is σ -integrable over $\mathcal{G}$, then the conditional expectations below exist and satisfy*

$$\mathbb{E}(\mathbb{E}(h|\mathcal{F})|\mathcal{G}) = \mathbb{E}(h|\mathcal{G}).$$

Remark 2.6.34. If we drop the condition of σ -finiteness on $\mathcal{G}$ it may happen that exactly one of $\mathbb{E}(h|\mathcal{F})$ and $\mathbb{E}(h|\mathcal{G})$ exists. Let $(S, \mathcal{A}, \mu) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), dx)$, $\mathcal{G} = \{\emptyset, \mathbb{R}\}$, $\mathcal{F} = \sigma([0, \infty))$ and $h(t) = \text{sgn}(t)/(1+t^2)$, we have $\mathbb{E}(h|\mathcal{G}) = 0$ but $\mathbb{E}(h|\mathcal{F})$ does not exist. For the same $\mathcal{G}$, $\mathcal{F} = \sigma(\{[k, k+1), k \in \mathbb{Z}\})$ and $h = \mathbf{1}$, we have $\mathbb{E}(h|\mathcal{F}) = h$, whereas $\mathbb{E}(h|\mathcal{G})$ does not exist. Note that in the second example the measure is σ -finite on $\mathcal{F}$.

Proof. The existence of $g := \mathbb{E}(h|\mathcal{G})$ and $f := \mathbb{E}(h|\mathcal{F})$ follows by Theorem 2.6.20 since h is σ -integrable over $\mathcal{G}$, and hence over $\mathcal{F}$, and μ is σ -finite on $\mathcal{G}$, and hence on $\mathcal{F}$. Moreover,

$$\mathcal{G}_f = \mathcal{F}_f \cap \mathcal{G} \supseteq \mathcal{F}_h \cap \mathcal{G} = \mathcal{G}_h$$

so that f is also σ -integrable over $\mathcal{G}$ and hence has a conditional expectation.

For $G \in \mathcal{G}_h \subseteq \mathcal{F}_h$, we have

$$\int_G g d\mu = \int_G \mathbb{E}(h|\mathcal{G}) d\mu = \int_G h d\mu = \int_G \mathbb{E}(h|\mathcal{F}) d\mu = \int_G f d\mu,$$

and thus g is the conditional expectation of f with respect to $\mathcal{G}$, as we wanted to show. $\square$

Independence and conditional expectations

The next two propositions are concerned with independence. Readers not familiar with this notion from probability theory may safely skip them.

Proposition 2.6.35 (Independence I). *Let $(S, \mathcal{A}, \mu)$ be a probability space. If $f \in L^1(S; X)$ is independent of the sub- σ -algebra $\mathcal{F}$, then its conditional expectation with respect to $\mathcal{F}$ is given by the constant function $\mathbb{E}(f|\mathcal{F}) = \mathbb{E}f$.*

Proof. The result is a consequence of the following identities, valid for all $F \in \mathcal{F}$:

$$\int_F f \, d\mu = \mathbb{E}(\mathbf{1}_F f) = \mathbb{E}(\mathbf{1}_F) \mathbb{E}f = \int_F \mathbb{E}f \, d\mu.$$

□

A more refined version of this proposition reads as follows.

Proposition 2.6.36 (Independence II). *Let $(S, \mathcal{A}, \mu)$ be a probability space and let $\mathcal{F}$ and $\mathcal{G}$ be sub- σ -algebras. Let $f \in L^0(S; X)$ be σ -integrable over $\mathcal{F}$. If $\mathcal{G}$ and $\sigma(f, \mathcal{F})$ are independent, then f has a conditional expectation with respect to $\sigma(\mathcal{G}, \mathcal{F})$ and*

$$\mathbb{E}(f | \sigma(\mathcal{G}, \mathcal{F})) = \mathbb{E}(f | \mathcal{F}).$$

Proof. For all $F \in \mathcal{F}_f$ and $G \in \mathcal{G}$ we have $F \in \mathcal{F}_{\mathbb{E}(f|\mathcal{F})}$ and

$$\begin{aligned} \int_{F \cap G} \mathbb{E}(f | \mathcal{F}) \, d\mu &= \int_S \mathbf{1}_G \cdot (\mathbf{1}_F \mathbb{E}(f | \mathcal{F})) \, d\mu \\ &\stackrel{(i)}{=} \mu(G) \int_F \mathbb{E}(f | \mathcal{F}) \, d\mu = \mu(G) \int_F f \, d\mu \\ &\stackrel{(ii)}{=} \int_S \mathbf{1}_G \cdot (\mathbf{1}_F f) \, d\mu = \int_{F \cap G} f \, d\mu. \end{aligned}$$

In (i) we used the independence of $\mathcal{G}$ and $\mathcal{F}$, in (ii) the independence of $\mathcal{G}$ and $\mathbf{1}_F f$. Let

$$\mathcal{C} = \{H \in \sigma(\mathcal{F}, \mathcal{G}) : H \subseteq F \cap G \text{ for some } F \in \mathcal{F}_f \text{ and } G \in \mathcal{G}\}.$$

Since S can be covered by countably many sets in $\mathcal{F}_f$, S can be covered by countably many sets in $\mathcal{C}$. It follows that $\mathcal{C}$ is an exhausting ideal for $\mathcal{H} := \sigma(\mathcal{F}, \mathcal{G})$.

By Proposition 2.6.9 It remains to be shown that for all $C \in \mathcal{C}$ we have

$$\int_C \mathbb{E}(f | \mathcal{F}) \, d\mu = \int_C f \, d\mu.$$

Fix a set $C \in \mathcal{C}$. By assumption we have $C \subseteq F_0 \cap G_0$ for some $F_0 \in \mathcal{F}_f$ and $G_0 \in \mathcal{G}$. Fix this set G_0 and consider the restricted σ -algebra $\mathcal{H}|_{G_0}$ on G_0 . The collection of all $F \cap G$ with $F \in \mathcal{F}_f$ and $G \subseteq G_0$ is closed under taking finite intersections and generates $\mathcal{H}|_{G_0}$. Hence by Dynkin's Lemma A.1.3 we conclude that

$$\int_H \mathbb{E}(f | \mathcal{F}) \, d\mu = \int_H f \, d\mu \quad \forall H \in \mathcal{H}|_{G_0}.$$

Since $C \in \mathcal{H}|_{G_0}$, this concludes the proof. □

As an illustration we include the following lemma which will be used in Theorems 3.3.10 and 4.4.1.

Lemma 2.6.37. *Let $(S, \mathcal{A}, \mu)$ be a probability space. Let $\xi_1, \dots, \xi_n$ be independent and identically distributed random variables in $L^0(S; X)$. Let*

$$\bar{\xi}_n := \frac{1}{k} \sum_{k=1}^n \xi_k.$$

Let $\mathcal{G} \subseteq \mathcal{A}$ be a σ -algebra which is independent of $\sigma(\xi_1, \bar{\xi}_n)$. Then

$$\mathbb{E}(\xi_j | \sigma(\bar{\xi}_n, \mathcal{G})) = \mathbb{E}(\xi_j | \sigma(\bar{\xi}_n)) = \bar{\xi}_n$$

for all $j = 1, \dots, n$.

Proof. The first identity follows from Proposition 2.6.36. Turning to the second identity, we claim that for any Borel set $B \subseteq X$ the following identity holds

$$\int_{\{\bar{\xi}_n \in B\}} \xi_j \, d\mathbb{P} = \int_{\{\bar{\xi}_n \in B\}} \xi_k \, d\mathbb{P}, \quad 1 \leq j, k \leq n. \quad (2.30)$$

Taking the average on the right-hand side of (2.30) with respect to k , keeping j fixed, the identity $\mathbb{E}(\xi_j | \sigma(\bar{\xi}_n)) = \bar{\xi}_n$ follows.

To prove (2.30) it suffices to take $j = 1$. For $k = 1, \dots, n$ let $h_k : X^n \rightarrow X$ be given by

$$h_k(x_1, \dots, x_n) := \mathbf{1}_{g(x_1, \dots, x_n) \in B} \otimes x_k \quad \text{with} \quad g(x_1, \dots, x_n) := \frac{1}{n} \sum_{j=1}^n x_j.$$

Then the following symmetry property holds: $h_j \circ \phi_{j,k} = h_k$, where $\phi_{j,k} : X^n \rightarrow X^n$ flips the j th and k th variable.

Let $\nu = \mathbb{P} \circ (\xi_1, \dots, \xi_n)^{-1}$ be the image measure on X^n . By the properties of the ξ_k 's we have $\nu \circ \phi_{j,k}^{-1} = \nu$. Therefore, by substitution (see Proposition 1.2.6), we find

$$\begin{aligned} \int_{\{\bar{\xi}_n \in B\}} \xi_1 \, d\mathbb{P} &= \int_{X^n} h_1 \, d\nu = \int_{X^n} h_1 \, d\nu \circ \phi_{1,k}^{-1} \\ &= \int_{X^n} h_1 \circ \phi_{1,k} \, d\nu = \int_{X^n} h_k \, d\nu = \int_{\{\bar{\xi}_n \in B\}} \xi_k \, d\mathbb{P}. \end{aligned}$$

□

Uniform integrability of conditional expectations

Definition 2.6.38 (Uniform integrability). *Let $1 \leq p < \infty$. A subset $T \subseteq L^p(S; X)$ is said to be uniformly p -integrable (or just uniformly integrable when $p = 1$) if*

(i) *For all $\varepsilon > 0$, there exists an $r > 0$ such that*

$$\sup_{f \in T} \int_S \mathbf{1}_{\{\|f\| > r\}} \|f\|^p \, d\mu \leq \varepsilon.$$

(ii) For every $\varepsilon > 0$, there is a set $B \in \mathcal{A}$ of finite measure such that

$$\sup_{f \in T} \int_{\mathbb{C}B} \|f\|^p d\mu \leq \varepsilon.$$

The reader may check that these conditions are satisfied if and only if the family $\{\|f\|^p : f \in T\}$ is uniformly integrable in the sense of Definition A.3.1.

Proposition 2.6.39. *Let $(\mathcal{F}_i)_{i \in I}$ be a family of sub- σ -algebras of $\mathcal{A}$ such that μ is σ -finite on their intersection $\mathcal{F} := \bigcap_{i \in I} \mathcal{F}_i$, and let $1 \leq p < \infty$. Then for all $f \in L^p(S; X)$ the family $\{\mathbb{E}(f|\mathcal{F}_i) : i \in I\}$ is uniformly p -integrable.*

It will follow from the proof that the sets B in condition (ii) may be chosen from $\mathcal{F}$.

Remark 2.6.40. The assumption of σ -finiteness of $\mathcal{F}$ cannot be weakened to σ -finiteness on each of the $\mathcal{F}_i$. Indeed, let $(S, \mathcal{A}, \mu) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), dx)$. Let

$$\mathcal{D}_n := \{2^{-n}([0, 1) + k) : k \in \mathbb{Z}\}, \quad n \in \mathbb{Z},$$

be the standard dyadic cubes of side-length 2^{-n} , and let $\mathcal{F}_n := \sigma(\mathcal{D}_n)$ the σ -algebra generated by them. Then $(\mathcal{F}_n)_{n \in \mathbb{Z}}$ is σ -finite filtration. For $f := \mathbf{1}_{[0, 1)}$ and $n \geq 0$, we have $f_{-n} := \mathbb{E}(f|\mathcal{F}_{-n}) = 2^{-n} \mathbf{1}_{[0, 2^n)}$. Given a set $B \in \mathcal{B}(\mathbb{R})$ of finite measure, most of the mass of f_{-n} eventually gets concentrated outside the set B as $n \rightarrow \infty$, and therefore condition (ii) cannot be satisfied. Notice that $\mathcal{F} = \{\emptyset, (-\infty, 0), [0, \infty), \mathbb{R}\}$ is purely infinite in this case.

Proof of Proposition 2.6.39. First we observe that $\|\mathbb{E}(f|\mathcal{F}_i)\|^p \leq \mathbb{E}(\|f\|^p|\mathcal{F}_i)$ by the conditional Jensen inequality. Thus by Proposition A.3.4 it suffices to consider the case $X = \mathbb{R}$ and $p = 1$, and we need to show the uniform integrability of $\{\mathbb{E}(|f|\mathcal{F}_i) : i \in I\}$ for every $f \in L^1(S)$.

Fix $\varepsilon > 0$ arbitrary and choose $\delta > 0$ such that $\mu(A) \leq \delta$ implies $\int_A |f| d\mu \leq \varepsilon$. Fix $i \in I$ and write $g_i := \mathbb{E}(|f|\mathcal{F}_i)$ for brevity. Set $r := \|f\|_1/\delta$ and note that $\mu(|g_i| > r) \leq \|g_i\|_1/r = \delta$. Noting the $\mathcal{F}_i$ -measurability of $\{|g_i| > r\}$, we conclude that

$$\int_S \mathbf{1}_{\{|g_i| > r\}} |g_i| d\mu = \int_S \mathbf{1}_{\{|g_i| > r\}} |f| d\mu \leq \varepsilon.$$

Choose an increasing sequence of sets $B_k \in \mathcal{F}$ of finite measure that exhaust S ; they exist by assumption. Then $\int_{\mathbb{C}B_k} |f| d\mu \rightarrow 0$ by dominated convergence. Given $\varepsilon > 0$, choose $B := B_k$ so large that the mentioned integral is at most ε . Then, for any $i \in I$, we have, using the conditional Jensen inequality and the defining property of the conditional expectation on $\mathbb{C}B \in \mathcal{F} \subseteq \mathcal{F}_i$,

$$\int_{\mathbb{C}B} |\mathbb{E}(f|\mathcal{F}_i)| d\mu \leq \int_{\mathbb{C}B} \mathbb{E}(|f|\mathcal{F}_i) d\mu = \int_{\mathbb{C}B} |f| d\mu \leq \varepsilon.$$

□

2.7 Notes

References on the general L^p -extension problem include [Bukhvalov \[1975\]](#), [Hernandez \[1984\]](#), [Herz \[1971\]](#), [Pisier \[1986a, 2010\]](#), [Viot \[1981\]](#).

Section 2.1

The following characterisation of “ L^p -extension spaces” is due to [Kwapień \[1972b\]](#) and establishes a converse to Proposition 2.1.2.

Theorem 2.7.1 (Kwapień). *Let $1 \leq p \leq \infty$ be fixed. For a Banach space X the following assertions are equivalent:*

- (1) *for any bounded operator $T \in \mathcal{L}(L^p(0, 1))$, $T \otimes I_X$ extends to a bounded operator on $L^p(0, 1; X)$;*
- (2) *there exists a measure space $(S, \mathcal{A}, \mu)$ such that X is isomorphic to a closed subspace of a quotient of a space $L^p(S)$.*

A good starting point for the structure theory of L^p -spaces is the Handbook article by [Alspach and Odell \[2001\]](#); classical papers on the subject include [Rosenthal \[1970, 1973\]](#).

Theorem 2.1.3 admits an extension to regular operators, and in this formulation a converse holds. Recall that a bounded operator $T : L^{p_1}(S_1) \rightarrow L^{p_2}(S_2)$ is called *regular* if there exists a positive operator $S : L^{p_1}(S_1) \rightarrow L^{p_2}(S_2)$ such that

$$|Tx| \leq S|x| \quad \forall x \in L^{p_1}(S_1). \quad (2.31)$$

For such an operator we define

$$\|T\|_{\text{reg}} := \inf \{ \|S\| : S \in \mathcal{L}(L^{p_1}(S_1), L^{p_2}(S_2)), S \geq 0, \text{ and (2.31) holds} \}.$$

Theorem 2.7.2 (Bukhvalov). *Let $1 \leq p_1, p_2 < \infty$. For a bounded operator T from $L^{p_1}(S_1)$ to $L^{p_2}(S_2)$ the following assertions are equivalent:*

- (1) *$T \otimes I_X$ extends to a bounded operator from $L^{p_1}(S_1; X)$ to $L^{p_2}(S_2; X)$ for every Banach space X ;*
- (2) *$T \otimes I_{\ell^1}$ extends to a bounded operator from $L^{p_1}(S_1; \ell^1)$ to $L^{p_2}(S_2; \ell^1)$;*
- (3) *T is regular.*

In this situation, for every Banach space X we have

$$\|T \otimes I_X\| \leq \|T\|_{\text{reg}} \leq \|T \otimes I_{\ell^1}\|.$$

An analogue for adjoint operators $T : L^\infty(S_1) \rightarrow L^p(S_2)$ in the style of Theorem 2.1.7 can be formulated as well.

The proof of Theorem 2.7.2 uses the fact, to be proved shortly, that T is regular if and only if there is a constant C such that for all $x_1, \dots, x_N$ in $L^{p_1}(S_1)$,

$$\left\| \sum_{n=1}^N |Tx_n| \right\|_{L^{p_2}(S_2)} \leq C \left\| \sum_{n=1}^N |x_n| \right\|_{L^{p_1}(S_1)}. \quad (2.32)$$

Moreover, $\|T\|_{\text{reg}}$ equals the least admissible constant C .

Proof. (3) $\Rightarrow$ (1): This follows by repeating the proof of Theorem 2.1.3 line by line; this argument produces the estimate $\|T \otimes I_X\| \leq \|T\|_{\text{reg}}$.

(1) $\Rightarrow$ (2): This implication is trivial.

(2) $\Rightarrow$ (3): Let $x_1, \dots, x_N \in L^{p_1}(S_1)$ be given. Then $x := (x_n)_{n=1}^N$ defines an element of $L^{p_1}(S; \ell_1)$ and $Tx := (Tx_n)_{n=1}^N$ satisfies $Tx = (T \otimes I_{\ell^1})x$ and

$$\begin{aligned} \left\| \sum_{n=1}^N |Tx_n| \right\|_{L^{p_2}(S_2)} &= \|(T \otimes I_{\ell^1})x\|_{L^{p_2}(S_2; \ell^1)} \\ &\leq \|T \otimes I_{\ell^1}\| \|x\|_{L^{p_1}(S_1; \ell^1)} = \|T \otimes I_{\ell^1}\| \left\| \sum_{n=1}^N |x_n| \right\|_{L^{p_1}(S_1)}. \end{aligned}$$

This shows that (2.32) holds, with constant $\|T \otimes I_{\ell^1}\|$. Therefore T is regular and $\|T\|_{\text{reg}} \leq \|T \otimes I_{\ell^1}\|$. $\square$

Let us now prove the characterisation of regular operators alluded to in (2.32). Recall that an operator T from $L^{p_1}(S_1)$ to $L^{p_2}(S_2)$ is called *regular* if there exists a positive operator $V : L^{p_1}(S_1) \rightarrow L^{p_2}(S_2)$ such that $|Tx| \leq S|V|$ for all $x \in L^{p_1}(S_1)$. For such an operator we define $\|T\|_{\text{reg}}$ as the infimum of the norms $\|V\|$, where V ranges over all positive operators from $L^{p_1}(S_1)$ to $L^{p_2}(S_2)$ satisfying $|Tx| \leq V|x|$ for all $x \in L^{p_1}(S_1)$.

Proposition 2.7.3. *Let $1 \leq p_1, p_2 \leq \infty$ and let $T : L^{p_1}(S_1) \rightarrow L^{p_2}(S_2)$ be a bounded linear operator. The following assertions are equivalent:*

- (1) T is regular;
- (2) there exists a constant C such that for all $x_1, \dots, x_N$ in $L^{p_1}(S_1)$,

$$\left\| \sum_{n=1}^N |Tx_n| \right\|_{L^{p_2}(S_2)} \leq C \left\| \sum_{n=1}^N |x_n| \right\|_{L^{p_1}(S_1)};$$

- (3) T is the difference of two positive operators.

Moreover, $\|T\|_{\text{reg}}$ is the infimum of admissible constants C in (2).

Sketch. The implications (3) $\Rightarrow$ (1) $\Rightarrow$ (2) are easy. The latter also shows that (2) holds with constant $C \leq \|T\|_{\text{reg}}$; the converse inequality follows from (2) by taking $N = 1$.

The proof of (2) $\Rightarrow$ (3) is taken from [Bukhvalov \[1975, Lemma 1.7\]](#) and uses the fact that in the spaces $L^p(S)$ every norm bounded set A that is directed upward (in the sense that for any two $y, y' \in A$ there exists $z \in A$

such that $z \geq y$ and $z \geq y'$) has a supremum. This property, in turn, implies that every non-empty set that is bounded above has a supremum (see, e.g., Meyer-Nieberg [1991, Section 2.4]).

Let us put $X = L^{p_1}(S_1)$ and $Y = L^{p_2}(S_1)$ for brevity. For $0 \leq x \in X$ consider the set

$$A_x = \left\{ \sum_{n=1}^N |Tx_n| : x = \sum_{n=1}^N x_n, x_n \geq 0 \right\}.$$

Using the Riesz decomposition property of Banach lattices (see Aliprantis and Burkinshaw [2006, Theorem 1.15]) this set is easily seen to be directed upward and, by our assumption, it is norm bounded. It follows that $\sup A_x$ exists. If $|z| \leq x$, then

$$|Tz| \leq |Tz^+| + |Tz^-| + T(x - |z|) \leq \sup A_x.$$

Thus we may define a mapping $T^+ : X^+ \rightarrow Y^+$ by $T^+x := \sup\{Tz : 0 \leq z \leq x\}$. Standard arguments from the theory of Banach lattices (see Aliprantis and Burkinshaw [2006, Chapter 1]) show that T^+ extends to a bounded positive linear operator from X to Y by setting, for any $x \in X$, $T^+x := T^+x^+ - T^+x^-$. Similarly, $T^-x := \sup\{-Tz : 0 \leq z \leq x\}$ defines a bounded positive operator from X to Y , and we have $T = T^+ - T^-$. $\square$

It is known that every bounded operator $T : L^1 \rightarrow L^p$ is regular (this follows, e.g., from Meyer-Nieberg [1991, Proposition 3.3.14]). Proposition 2.1.1 then appears as a special case of Theorem 2.7.2.

The idea to use Lemma 2.1.6 to prove Theorem 2.1.7 was shown to us by N. Lindemulder. Further results along this line may be found in Lindemulder [2016].

Our proof of Theorem 2.1.9 is adapted from García-Cuerva and Rubio de Francia [1985]. In its present form it is due to Marcinkiewicz and Zygmund [1939]; the case $p_1 = p_2$ (without sharp constant) was shown earlier by Paley [1932]. Some further developments are presented in Garling [2007], which also contains a wealth of material relating to other topics treated in this chapter.

In Example 2.1.15 we have dealt with the Fourier transform on the real line. A similar example can be given for the Fourier transform on the circle. In this formulation, the first to notice the non-extendability of the Hausdorff–Young inequality to certain Banach spaces X was Bochner [1933a]. His example is interesting enough to be mentioned:

Example 2.7.4. The Fourier transform on the circle $\mathbb{T} = [0, 1)$ is defined by $\mathcal{F}f(k) := \widehat{f}(k)$, $k \in \mathbb{Z}$, with

$$\mathcal{F}f(k) := \int_0^1 f(t) e^{-2\pi i k t} dt.$$

It is elementary to see that $\mathcal{F}$ an isometry from $L^2(\mathbb{T})$ onto $\ell^2(\mathbb{Z})$. We wish to check that $\mathcal{F} \otimes I_{L^1(\mathbb{T})}$ does not extend to a bounded operator from $L^2(\mathbb{T}; L^1(\mathbb{T}))$ to $\ell^2(\mathbb{Z}; L^1(\mathbb{T}))$.

Suppose the contrary. Fix an arbitrary $f \in L^1(\mathbb{T})$ and define $g_f \in L^2(\mathbb{T}; L^1(\mathbb{T}))$ by $(g_f(t))(s) := f(t + s \bmod 1)$. By approximation one sees that its image under $\mathcal{F} \otimes I_{L^1(\mathbb{T})}$ must be the sequence $(\widehat{g_f}(k))_{k \in \mathbb{Z}}$, where $\widehat{g_f}(k) \in L^1(\mathbb{T})$ is the function $s \mapsto e^{2\pi i k s} \widehat{f}(k)$, so that

$$\|(\mathcal{F} \otimes I_{L^1(\mathbb{T})})g_f\|_{\ell^2(\mathbb{Z}; L^1(\mathbb{T}))}^2 = \sum_{k \in \mathbb{Z}} \|e^{2\pi i k(\cdot)} \widehat{f}(k)\|_{L^1(\mathbb{T})}^2 = \sum_{k \in \mathbb{Z}} |\widehat{f}(k)|^2.$$

By assumed boundedness of $\mathcal{F} \otimes I_{L^1(\mathbb{T})}$, it follows that

$$\|f\|_2^2 = \sum_{k \in \mathbb{Z}} |\widehat{f}(k)|^2 \leq \|\mathcal{F} \otimes I_{L^1(\mathbb{T})}\|^2 \|g_f\|_{L^2(\mathbb{T}; L^1(\mathbb{T}))}^2 = \|\mathcal{F} \otimes I_{L^1(\mathbb{T})}\|^2 \|f\|_1^2.$$

an obvious contradiction.

Incidentally, this example also shows that the partial Fourier sums $S_N g$ of a function in $g \in L^2(\mathbb{T}; X)$ need not converge to g in the norm of $L^2(\mathbb{T}; X)$, an observation from [Bochner and Taylor \[1938\]](#). Indeed, with the above notation we have $S_N g_f(k) = \sum_{|n| \leq N} e^{2\pi i k n t} e^{i k(\cdot)} \widehat{f}(k)$. Now convergence $S_N g_f \rightarrow g_f$ in $L^2(\mathbb{T}; L^1(\mathbb{T}))$ would lead to the contradiction

$$\begin{aligned} \|f\|_{L^2(\mathbb{T})}^2 &= \lim_{N \rightarrow \infty} \int_0^1 \left| \sum_{|k| \leq N} e^{2\pi i k n t} \widehat{f}(k) \right|^2 dt \\ &\leq \lim_{N \rightarrow \infty} \int_0^1 \|S_N g_f(t)\|_{L^1(\mathbb{T})}^2 dt = \|g_f\|_{L^2(\mathbb{T}; L^1(\mathbb{T}))}^2 \leq \|f\|_{L^1(\mathbb{T})}^2. \end{aligned}$$

Example [2.1.15](#) is from [Peetre \[1969\]](#). Example [2.1.16](#) is a variation of an example in [Defant \[1989\]](#). Example [2.1.17](#) is due to [Rosiński and Suchanecki \[1980\]](#). Previous examples of bounded functions which fail to admit a stochastic integral were given in [Yor \[1974\]](#). Theorem [2.1.20](#) is due to [Hoffmann-Jørgensen and Pisier \[1976\]](#).

Weighted norm inequalities

In harmonic analysis, weighted norm inequalities in the scalar case often imply the boundedness of vector-valued extensions. This principle was first observed by [Rubio de Francia \[1984\]](#). This rich and deep subject is covered in the monographs [García-Cuerva and Rubio de Francia \[1985\]](#) and [Cruz-Uribe, Martell, and Pérez \[2011\]](#). A typical example is as follows. Let $1 < p_0 < \infty$ and suppose T is a bounded operator on $L^{p_0}(\mathbb{R}^d, w)$ for all Muckenhoupt A_{p_0} -weights w , with a bound on the norms which depends on the A_{p_0} -constant of w in a suitable ‘monotone’ way. Then T extends to a bounded operator on $L^p(\mathbb{R}^d, w; \ell^q)$

for all $1 < p, q < \infty$ and all A_p -weights w . This result remains true for sub-linear operators, and in this formulation it covers the Hardy–Littlewood maximal operator. Instead of ℓ^q it is possible to use more general classes of Banach function spaces; see [Rubio de Francia \[1986\]](#). Recall that a *Banach function space* over a measure space $(S, \mathcal{A}, \mu)$ is a subspace $X \subseteq L^0(S)$ equipped with a complete norm that respects the pointwise ordering, in the sense that if $|f| \leq |g|$ for some $f \in L^0(S)$ and $g \in X$, then $f \in X$ with $\|f\|_X \leq \|g\|_X$.

Section 2.2

The chronologically first result on the interpolation of operators in the modern sense is the Riesz–Thorin Theorem [2.2.1](#). It is originally due to [Riesz \[1927\]](#), whose proof is quite different from the present one; it is completely “elementary” (although perhaps not easy), as it has nothing to do with complex variables but relies at the bottom on a careful analysis of the conditions for equality in Hölder’s inequality. In the introduction to his article, [Riesz \[1927\]](#) indicates the work of [Jensen \[1906\]](#) on convex function inequalities as his source of inspiration. It was [Thorin \[1938\]](#) who introduced the now-standard complex-variable framework for this result, foreshadowing the abstract notion of complex interpolation. Our proof of Theorem [2.2.1](#) is standard and may be found in most textbooks on interpolation theory. We mention [Bergh and Löfström \[1976\]](#), [Kreĭn, Petunĭn, and Semĕnov \[1982\]](#), [Lunardi \[2009\]](#), [Triebel \[1978\]](#).

The Marcinkiewicz Interpolation Theorem [2.2.3](#) was announced without proof in a two-page note by [Marcinkiewicz \[1939a\]](#). The full theorem, already formulated in [Marcinkiewicz \[1939a\]](#), also deals with operators T mapping L^p into another L^q . such a mapping is called *quasi-linear* if, for all $f \in L^p$ and $c \in \mathbb{K}$,

$$\|T(cf)\| = |c|\|T(f)\| \quad \text{almost everywhere,}$$

and there is a constant $C \geq 0$ such that for all $f, g \in L^p$,

$$\|T(f + g)\| \leq C(\|T(f)\| + \|T(g)\|) \quad \text{almost everywhere.}$$

Theorem 2.7.5 ([Marcinkiewicz \[1939a\]](#)). *Let $T : L^{p_0} + L^{p_1} \rightarrow L^0$ be a quasi-linear operator that maps $T : L^{p_i} \rightarrow L^{q_i, \infty}$ for $i = 0, 1$, where $p_i \leq q_i$ and $p_0 < p_1$. Then*

$$T : L^p \rightarrow L^q \quad \text{for} \quad \frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad \frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q}, \quad \theta \in (0, 1).$$

According to [Zygmund \[1956, p. 228\]](#), the proof of this theorem was indicated in Marcinkiewicz’s letter to Zygmund in June 1939, and finally published by [Zygmund \[1956\]](#). The argument found there is already the one presented in most modern textbooks; we have followed the same lines, except for some fine-tuning of the constants which may be new. Our proof also gives Theorem

2.7.5, but a variation of Lemma 2.2.5 is needed which involves a Minkowski type argument.

The theorem of Marcinkiewicz [1939a] apparently attracted little attention until it was revisited by Zygmund [1956]. It seems that Zygmund himself had forgotten about the result in the first place, as he later adds that his fundamental work on singular integrals, Calderón and Zygmund [1952], “implicitly uses the same ideas”, whereas “an explicit application of [Marcinkiewicz’s interpolation theorem] would have shortened the proof” (Zygmund [1956, p. 235]). The key novelty of Marcinkiewicz theorem, allowing applications beyond the scope of the Riesz–Thorin theorem, was the use of weak-type norms. In the preface to Marcinkiewicz [1964], Zygmund writes that this was “suggested in all probability by an old result of Kolmogorov [1925] about properties of conjugate functions.” An extension of Marcinkiewicz’s result to Lorentz spaces is due to Hunt [1964].

Comprehensive treatments of vector-valued function spaces and their interpolation theory are contained in Amann [1995], König [1986], Kreĭn, Petunĭn, and Semĕnov [1982], Schmeisser [1987]. Lemma 2.2.7 is due to Calderón [1964] with a different proof. Our presentations of Theorems 2.2.6 and 2.2.10 follow Triebel [1978, Section 1.18.4] and Kreĭn, Petunĭn, and Semĕnov [1982, Theorem IV.2.11], respectively. The Clarkson inequalities (Corollary 2.2.9) are from Clarkson [1936]. The proof presented here was kindly shown to us by J. Voigt. The idea of applying complex interpolation to the mapping $(f, g) \mapsto (f+g, f-g)$ appears in Kato and Takahashi [1997], where the validity of the Clarkson inequalities in $L^p(S; X)$ is related to the (Rademacher) type and cotype constants of X . Further results along this line are given in Milman [1984], where Clarkson inequalities in $L^p(S; X)$ are related to the Fourier type of X with respect to the group $\{0, 1\}$ with addition modulo 2.

Section 2.3

Hardy and Littlewood introduced the maximal function that now bears their name in their famous 1930 paper Hardy and Littlewood [1930]. They proved its L^p -boundedness as a consequence of an elaborate argument using decreasing rearrangements, which were introduced for that purpose in the same paper. The latter’s usefulness is explained as follows:

“The problem is most easily grasped when stated in the language of cricket, or any other game in which a player compiles a series of scores of which an average is recorded.”

They then consider the batsman’s satisfaction as an increasing function of his average over the number of inning played to date. This leads them to the notion of decreasing rearrangements, for it is easy to see that

“(The batsman’s) total satisfaction for the season (...) is a maximum, for a given stock of innings, when the innings are played in decreasing order.”

Instead of the batsman's average for the complete season to date, they proceed to consider his maximum average for any consecutive series of innings to date, for which bounds are established that ultimately lead to the L^p -boundedness of the maximal function.

The limitation of this argument is that it works in dimension $d = 1$ only. The extension to higher dimensions through the use of a covering lemma is due to Wiener [1939]. The version with weights presented in Theorem 2.3.2, which works for functions with values in an arbitrary Banach space X , should be carefully distinguished from the Hardy–Littlewood theorem for A_p -weights hinted at in the Notes to Section 2.1. The latter states that if $1 < p < \infty$ and w is a Muckenhoupt A_p -weight on $\mathbb{R}^d$, then the Hardy–Littlewood function M is bounded on $L^p(\mathbb{R}^d; w)$ and, more generally, on $L^p(\mathbb{R}^d, w; \ell^q)$ for all $1 < q < \infty$.

The literature contains many versions of Theorem 2.3.8, making it difficult to trace its exact origins. The reader is referred to Stein and Weiss [1971] for a fuller discussion. It was shown in Burkholder [1964] that, in many situations, maximal inequalities are not only sufficient, but also necessary to obtain the almost everywhere convergence of a given sequence of functions.

Section 2.4

With the exception of the material on Fourier type, the results of this section are standard. Vector-valued spaces of test functions and the corresponding classes of (tempered) distributions are discussed in Amann [1995].

The notion of Fourier type goes back to Peetre [1969], where most of the results presented here can be found. Peetre was concerned with improvement of the classical inclusion relations

$$(X_0, X_1)_{\theta, 1} \hookrightarrow [X_0, X_1]_{\theta} \hookrightarrow (X_0, X_1)_{\theta, \infty}$$

between the real and complex interpolation spaces in the situation where X_0 and X_1 have Fourier type p_0 and p_1 , respectively. He was able to prove that under this assumption one has the embeddings

$$(X_0, X_1)_{\theta, p_{\theta}} \hookrightarrow [X_0, X_1]_{\theta} \hookrightarrow (X_0, X_1)_{\theta, p'_{\theta}},$$

where $\frac{1}{p_{\theta}} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$. Peetre's proof is reproduced in Appendix C.

The equivalence of Fourier type for the line and the circle of Proposition 2.4.20 is due to Kwapien [1972a] whose approach we follow. A detailed proof that the function considered in (2.17) takes its minimum in the points $\frac{1}{2} + \mathbb{Z}$ is written out in Gijswijt and Van Neerven [2016]. The proof is long and tedious, and ultimately relies on a variant of the classical Hardy–Littlewood–Pólya inequality.

The notion of Fourier type is connected with the probabilistic notions of type and cotype (see Definition 4.3.12; a detailed study will follow in Volume II). It is straightforward to see that every Banach space with Fourier type p has type p and cotype q . It is a deep theorem of Bourgain [1988] that the following converse holds:

Theorem 2.7.6 (Bourgain). *If X has non-trivial type, then X has non-trivial Fourier type.*

Here, ‘non-trivial’ means ‘belonging to $(1, 2]$ ’. It is not true, however, that non-trivial type p implies non-trivial Fourier type p ; Bourgain’s proof gives a non-linear dependence of the Fourier type on the type and its constant. Another proof of Bourgain’s theorem can be found in [Hinrichs \[1996\]](#).

In [König \[1991\]](#), various summability properties of the Fourier coefficients of X -valued functions were established under Fourier type assumptions on X . For instance, it was shown that X has non-trivial Fourier type if and only if the Fourier coefficients of every $f \in C^1(\mathbb{T}; X)$ are absolutely summable.

Expository references for this material are [García-Cuerva, Kazaryan, Kolyada, and Torrea \[1998\]](#) and [Pietsch and Wenzel \[1998\]](#). Bourgain’s result has various application to vector-valued harmonic analysis; in the presence of non-trivial Fourier type it is often possible to improve certain parameters, much in the way as we saw above in connection with interpolation.

Section 2.5

A classical reference on Sobolev spaces is the book by Adams, which appeared in its second edition as [Adams and Fournier \[2003\]](#). Gentle introductions may be found in [Brezis \[2011\]](#) and [Evans \[2010\]](#). A thorough treatment of fractional Sobolev spaces which includes various other related classes of function spaces is presented in [Runst and Sickel \[1996\]](#) and [Triebel \[1978\]](#). A self-contained introduction is presented in [Di Nezza, Palatucci, and Valdinoci \[2012\]](#). Sobolev spaces of vector-valued functions are treated mostly in connection with the theory of evolution equations; see for instance [Amann \[1995\]](#), [Prüss and Simonett \[2016\]](#). A detailed discussion of the theory of vector-valued Sobolev spaces can be found in [Schmeisser \[1987\]](#), [Triebel \[1997\]](#), and references therein.

The results connecting weak derivatives and almost everywhere derivatives are mostly classical, although it is hard to find a systematic reference in the literature. For more on the connection with the Radon–Nikodým property we refer to [Diestel and Uhl \[1977\]](#). The proofs of [Lemma 2.5.10](#) and [2.5.11](#) are taken from [Stein and Shakarchi \[2005\]](#). We learnt the trick in the proof of $(3) \Rightarrow (2)$ of [Theorem 2.5.12](#) from [Arendt, Batty, Hieber, and Neubrander \[2011\]](#) whose presentation we follow. A direct real variable proof of [Lemma 2.5.13](#) that avoids the use of the Carathéodory theorem and the Radon–Nikodým theorem can be found in many textbooks. A detailed presentation is given in [Stein and Shakarchi \[2005\]](#).

The proof of [Theorem 2.5.17](#) is taken from [Lunardi \[2009\]](#). The result can be generalised to open domains D in $\mathbb{R}^d$ that admit suitable extension operators, such as domains with a uniformly C^1 boundary. For such domains one then obtains

$$(L^p(D; X), W^{1,p}(D; X))_{\theta,p} = W^{\theta,p}(D; X)$$

with equivalent norms.

Section 2.6

Conditional expectations in a σ -finite setting have been considered by many authors. Some aspects of our approach seem to be new, however, even in the scalar-valued situation. Our approach is related to that of [Stromberg \[1994\]](#) and [Tanaka and Terasawa \[2013\]](#), the principal difference being that we take the collection $\mathcal{F}_f$ of sets over which f is integrable as the point of departure rather than the collection of sets of finite μ -measure, our philosophy being that what really matters is the integrability properties of f on a given set, and not so much the size of the set.

[Neveu \[1975\]](#) takes a different approach which we will outline and compare with ours. Let first $(S, \mathcal{A}, \mathbb{P})$ be a probability space and let $\mathcal{F} \subseteq \mathcal{A}$ be a sub- σ -algebra. For arbitrary $\mathcal{A}$ -measurable $f : S \rightarrow [0, \infty]$, Neveu checks that $\mathbb{E}(f|\mathcal{F})$ may be defined as the almost everywhere limit of $\mathbb{E}(f \wedge n|\mathcal{F})$ as $n \rightarrow \infty$, in the sense that this limit is the unique (up to a $\mathbb{P}$ -null set) $\mathcal{F}$ -measurable function with values in $[0, \infty]$ such that

$$\int_F \mathbb{E}(f|\mathcal{F}) d\mathbb{P} = \int_F f d\mathbb{P} \quad \forall F \in \mathcal{F}. \quad (2.33)$$

As a second step, let $(S, \mathcal{A}, \mu)$ be a σ -finite measure space and let $\mathcal{F} \subseteq \mathcal{A}$ be a sub- σ -algebra. We do not assume μ to be σ -finite on $\mathcal{F}$. Fix a strictly positive integrable function $w : S \rightarrow (0, \infty)$. Then the measure $\mathbb{P} := w\mu$ is a probability measure and we have $\mu = w^{-1}\mathbb{P}$ (where of course $w^{-1} = 1/w$). On $\mathcal{F}$, by (2.33) we have $\mu = \mathbb{E}(w^{-1}|\mathcal{F})\mathbb{P}$, the conditional expectation being taken with respect to $\mathbb{P}$. The set

$$F_\mu := \{\mathbb{E}(w^{-1}|\mathcal{F}) < \infty\}$$

belongs to $\mathcal{F}$ and has the following properties:

- (i) $\mu|_{\mathcal{F}}$ is σ -finite on F_μ ;
- (ii) $\mu|_{\mathcal{F}}$ is purely infinite on $\mathbb{C}F_\mu$.

Recall that the second property means that $\mu|_{\mathcal{F}}$ only takes the values 0 and ∞ on the $\mathcal{F}$ -measurable subsets of $\mathbb{C}F_\mu$. Then Neveu observes that $\mathbb{E}(w^{-1}f|\mathcal{F})$ is strictly positive μ -almost everywhere on F_μ and that, for an $\mathcal{A}$ -measurable function $f : S \rightarrow [0, \infty]$, the prescription

$$\mathbb{E}_\mu(g|\mathcal{F}) := \begin{cases} \frac{\mathbb{E}(w^{-1}f|\mathcal{F})}{\mathbb{E}(w^{-1}|\mathcal{F})} & \text{on } F_\mu \\ 0 & \text{on } \mathbb{C}F_\mu \end{cases} \quad (2.34)$$

defines the conditional expectation of g with respect to $\mathcal{F}$, taken with respect to μ , in the sense that it agrees with the orthogonal projection in $L^2(S, \mathcal{F}, \mu)$

onto $L^2(S, \mu)$ and is the unique (up to a $\mathbb{P}$ -null set) $\mathcal{F}$ -measurable function with values in $[0, \infty]$ such that

$$\int_F \mathbb{E}_\mu(g|\mathcal{F}) \, d\mu = \int_F g \, d\mu \quad \forall F \in \mathcal{F}, \, F \subseteq F_\mu. \quad (2.35)$$

In particular, the definition of $\mathbb{E}_\mu(g|\mathcal{F})$ is independent of the choice of w . This conditional expectation defines a positive linear contraction from $L^p(S, \mathcal{F}, \mu)$ onto $L^p(S, \mu)$ for each $p \in [1, \infty]$.

To compare this approach with ours, we first observe that if $(S, \mathcal{A}, \mu)$ is a σ -finite measure space and $\mathcal{F} \subseteq \mathcal{A}$ is a sub- σ -algebra, then by Proposition A.1.4 there exist disjoint sets $S_0, S_1 \in \mathcal{F}$ such that $S_0 \cup S_1 = S$ such that the restriction of μ to $\mathcal{F}|_{S_0}$ is σ -finite and the restriction of μ to $\mathcal{F}|_{S_1}$ is purely infinite. To see that Neveu's approach coincides with ours, we simply note that $S_0 = F_\mu$ and $S_1 = \mathbb{C}F_\mu$ up to μ -null sets, for there can be only one such partition of S up to μ -null sets. Now it follows from (2.35) and uniqueness that on F_μ the conditional expectation given by (2.34) agrees with ours.

The extension of this approach to the vector-valued case would create some difficulties, since the step which extends the conditional expectation on $L^2(S, \mathbb{P})$ to $L^0(S, \mathbb{P})$ is based on a monotonicity argument. This can easily be circumvented by using the fact (see Proposition 2.6.2) that any function in $L^0(S; X)$ is σ -integrable and then define the conditional expectation along the lines of (2.28), but this would bring us straight to the construction we have taken. Let us also emphasise that our approach develops conditional expectations from first principles.

Yet another approach could be based on extending the regular conditional probability (see Kallenberg [2002, Chapter 6]) and to use it to define the vector-valued conditional expectation. Using this method, “conditional versions” of several classical theorems can be derived immediately.

Subsections 2.6.c and 2.6.d follow a more or less standard pattern. An alternative proof of Proposition 2.6.29 (for functions $f \in L^p(S; X)$) can be found in Haase [2007].

Conditional expectations in $L^p(S; L^q(T; X))$

Let $(S, \mathcal{A}, \mu)$ and $(T, \mathcal{B}, \nu)$ be measure spaces. For a sub- σ -algebra $\mathcal{F}$ of the product σ -algebra $\mathcal{A} \times \mathcal{B}$, we define $L^p_{\mathcal{F}}(S; L^q(T; X))$ to be the closed subspace in $L^p(S; L^q(T; X))$ of all functions having a strongly $\mathcal{F}$ -measurable representative. The following result is proved in L   and Van Neerven [2016].

Theorem 2.7.7. *Let $(A, \mathcal{A}, \mu)$ and $(B, \mathcal{B}, \nu)$ be probability spaces, let X be a Banach space, and let $1 < p, q < \infty$. Then the conditional expectation operator $\mathbb{E}(\cdot|\mathcal{F})$ on $L^1(S \times T)$ restricts to a bounded projection on the space $L^p(S; L^q(T; X))$ with range $L^p_{\mathcal{F}}(S; L^q(T; X))$. The norm of this projection is bounded by a constant independent of X .*

The next example, due to [Qiu \[2012\]](#), shows that the conditional expectation may fail to be contractive. It will be refined in Chapter 4 (see Proposition [4.3.14](#)).

Example 2.7.8. Let $A = B = \{0, 1\}$ with $\mathcal{A} = \mathcal{B} = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}$ and $\mu = \nu$ the measure on $\{0, 1\}$ that gives each point mass $\frac{1}{2}$, and let

$$\mathcal{F} = \{\{(0, 0), (1, 0)\}, \{(0, 1)\}, \{(1, 1)\}\}.$$

Let $f : A \times B \rightarrow \mathbb{R}$ be defined by $f(0, 0) = 0$, $f(1, 0) = 1$, $f(0, 1) = 1$, $f(1, 1) = 0$. Then

$$\mathbb{E}(f|\mathcal{F})(0, 0) = \frac{1}{2}, \quad \mathbb{E}(f|\mathcal{F})(1, 0) = \frac{1}{2}, \quad \mathbb{E}(f|\mathcal{F})(0, 1) = 1, \quad \mathbb{E}(f|\mathcal{F})(1, 1) = 0.$$

Hence in this example we have

$$\begin{aligned} \|f\|_{L^p(S; L^2(T))} &= \left(\left(\frac{1}{2}\right)^{p/2} + \left(\frac{1}{2}\right)^{p/2} \right)^{1/p}, \\ \|\mathbb{E}(f|\mathcal{F})\|_{L^\infty(S; L^2(T))} &= \left(\left(\frac{1}{8}\right)^{p/2} + \left(\frac{5}{8}\right)^{p/2} \right)^{1/p}. \end{aligned}$$

Consequently, for large enough p the conditional expectation fails to be contractive in $L^p(S; L^2(T))$.

As an immediate consequence of Proposition [2.7.7](#) we obtain the following duality result, due to [Lü, Yong, and Zhang \[2012\]](#). Their proof, however, requires an additional assumption on $\mathcal{F}$.

Corollary 2.7.9. *Let $(A, \mathcal{A}, \mu)$ and $(B, \mathcal{B}, \nu)$ be a probability space and a σ -finite measure space respectively, and let X be a Banach space whose dual has the Radon-Nikodým property, and let $1 < p, q < \infty$. Then we have a natural isometric isomorphism of Banach spaces*

$$(L^p_{\mathcal{F}}(S; L^q(T; X)))^* = L^{p'}_{\mathcal{F}}(S; L^{q'}(T; X)).$$

This corollary cannot be extended to $p = 1$, even when $q = 2$. In fact it is shown in [Lü and Van Neerven \[2016\]](#) that if $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space supporting a Brownian motion $(B_t)_{t \in [0, 1]}$, and if $\mathcal{P}$ is the so-called progressive σ -algebra associated with it (see, e.g., [Chung and Williams \[1990\]](#) for its definition), then

$$L^\infty_{\mathcal{P}}(\Omega; L^2(0, 1)) \subsetneq (L^1_{\mathcal{P}}(\Omega; L^2(0, 1)))^*. \quad (2.36)$$

Analysis in Banach Spaces

Volume I: Martingales and Littlewood-Paley Theory

Hytönen, T.; van Neerven, J.; Veraar, M.; Weis, L.

2016, XVII, 614 p. 3 illus., Hardcover

ISBN: 978-3-319-48519-5