Helsinki University of Technology Institute of Mathematics Research Reports
Teknillisen korkeakoulun matematiikan laitoksen tutkimusraporttisarja
Espoo 2001 C16

R-BOUNDEDNESS AND MULTIPLIER THEOREMS

Tuomas Hytönen

TEKNILLINEN KORKEAKOULU
TEKNISKA HÖGSKOLAN
HELSINKI UNIVERSITY OF TECHNOLOGY
TECHNISCHE UNIVERSITÄT HELSINKI
UNIVERSITE DE TECHNOLOGIE D'HELSINKI

Helsinki University of Technology Institute of Mathematics Research Reports
Teknillisen korkeakoulun matematiikan laitoksen tutkimusraporttisarja
Espoo 2001 C16

R-BOUNDEDNESS AND MULTIPLIER THEOREMS

Tuomas Hytönen

Helsinki University of Technology
Department of Engineering Physics and Mathematics
Institute of Mathematics

Tuomas Hytönen: *R-boundedness and Multiplier Theorems*; Helsinki University of Technology Institute of Mathematics Research Reports C16 (2001).

Abstract: *This treatise concentrates on the theory of Fourier multipliers and the machinery required by the modern methods used in this field. We present the classical starting points of the theory, the abstract generalizations and the R-boundedness techniques used in this connection. We also treat the theory of UMD-spaces, which provide a general setting for strong vector-valued multiplier results. This machinery is applied to prove some recent theorems concerning operator-valued Fourier multipliers.*

AMS subject classifications: 42B15 (Primary), 42A61, 46B09, 46B15, 46E40, 60G42 (Secondary).

Keywords: Fourier-multiplier, R-boundedness, UMD-space, Schauder decomposition, martingale, Hilbert transform, Riesz projection, Rademacher functions.

Tuomas.Hytonen@hut.fi

ISBN 951-22-5477-8
ISSN 0784-6460
Espoo, 2001

Helsinki University of Technology
Department of Engineering Physics and Mathematics
Institute of Mathematics
P.O. Box 1100, 02015 HUT, Finland
email:math@hut.fi http://www.math.hut.fi/

Preface

This report constitutes my Master's thesis written at Helsinki University of Technology. The goal of the work was to give a reasonably selfcontained and readable treatise of some of the concepts, tools and results making up the currently vital field of R-boundedness and multiplier theory in UMD-spaces.

I wish to express my sincere thanks to my instructor and supervisor, professor Stig-Olof Londen, for this most fascinating topic, for all the good advice and, above all, for his continuing interest in my work.

I was first guided to the world of R-boundedness by the brilliant lectures of professor Jan Prüss, one of the active developers of the field. Thanks to the foreign contacts of my instructor, I also had the opportunity to benefit from the expertise of professor Ben de Pagter, whose lecture notes were my main source regarding the UMD-theory, and who also kindly answered a number of enquiries via e-mail, and professor Philippe Clément, who recently gave me some fresh ideas going beyond the scope of the present work.

Professor Juha Kinnunen discussed some points of this work with me, and following his advice, I also exchanged a couple of ideas with professor Eero Saksman.

Except at some (very few) places, I make no claim as to novelty. Although I have collected some notes and comments at the end of each chapter, I have not tried to document the origin (or the author) of each particular result.

I had the chance to do this work at the Institute of Mathematics of Helsinki University of Technology. All the people were very kind, and I was even paid for having a great time.

Otaniemi, 23.4.2001

Tuomas Hytönen

Contents

Chapter 0

Introduction

This work concentrates on the two concepts appearing in its title, R-boundedness and multiplier theorems, of which the former one is the tool, the latter the application.

One may pronounce the "R" as either randomized, Riesz or Rademacher; we prefer the first one, since it is probably the most informative. It should guide one's thoughts towards the heart of the matter, the idea of attacking problems of analysis by stochastic means.

It was pointed out by Stein [23], in his introductory words preceding the treatment of classical multiplier results and the Littlewood–Paley theory, that "often the most fruitful way of characterizing various analytic situations – – is in terms of appropriate quadratic expressions". Numerous successful applications show that this is certainly true in the scalar-valued situation, and the same techniques often extend to Hilbert spaces provided one replaces the square of an element by the inner product of a vector by itself. Obviously, however, none of this has any meaning in a general Banach space.

On the other hand, it is known that various quadratic sums can equivalently be treated by means of linear expansions in terms of random weights called Rademacher functions. This is sometimes convenient in the scalar-valued situation, but more importantly, it allows one to generalize the powerful classical machinery far beyond spaces to which the original notion applies.

On this basis, the modern notion of randomized boundedness of operator families appears as a promising tool. This is definitely an understatement, for the recent developments have provided significant generalizations of the classical multiplier theorems, and far reaching applications have emerged also in the important study of maximal L^p-regularity of evolution equations, treated recently e.g. by Arendt and Bu [1], Hieber and Prüss [8] and Weis [25, 26].

Many of the most powerful modern theorems are valid in so-called UMD-spaces, i.e., Banach spaces of unconditional martingale differences. Martingales constitute a class of stochastic processes, so once again, randomization enters the scene. The probabilistic characterization of UMD turns out to be equivalent to the L^p-boundedness of the Hilbert transform, a transformation which is, in a sense, a very representative example of a multiplier operator: The boundedness of this transformation then allows us make the same conclusion concerning many others, the essential assumption directly involving the notion of R-boundedness.

Inspired by these modern developments, it seems appropriate to rephrase the celebration of quadratic expressions as follows: A powerful means of attacking many analytic problems, governed by ultimate determinism, is the act of randomization.

The field of the Fourier multiplier theory is rather wide, and even the classical results would provide material for a treatment of significantly greater length than the present one. There are many books touching these matters, e.g. those of Duoandikoetxea [7], Stein [23], and Stein and Weis [24]. We have only treated the classical theory to show the natural emergence of the notion of multipliers; nevertheless, many of the classical theorems are obtained as special cases of the modern vector-valued results which we concentrate on. Emphasis is also given to the techniques and ideas behind the modern developments, in particular the UMD-theory, which is not so recent as such, but on which most of the modern results are built.

The results appearing here are already known, but the collective presentation is new. A recent work that partly parallels ours is the thesis of Witvliet [28] (a part of which has appeared in the article by Clément et al. [3]); however, [28] concentrates more on the applications to a variety of multiplier theorems (in which direction the treatment goes well beyond ours), but the results from the UMD-theory are only cited there.

The presentation can roughly be divided into three parts: the introductory ideas and results, the UMD-theory, and the modern theorems.

The first four chapters present the starting points of the theory, and we obtain the first multiplier theorems. In Chapter 1, we show that translation invariant transformations $T \in \mathcal{B}(L^p(\mathbb{T}^d); L^q(\mathbb{T}^d))$, where $\mathbb{T}$ is the unit-circle, act on trigonometric polynomials as componentwise multipliers, formally

$$T \sum_{\kappa \in \mathbb{Z}^d} a_\kappa e^{\mathbf{i}2\pi\kappa\cdot(\cdot)} = \sum_{\kappa \in \mathbb{Z}^d} \lambda_\kappa a_\kappa e^{\mathbf{i}2\pi\kappa\cdot(\cdot)},$$

and a similar result is also true for operators acting on $L^p(\mathbb{R}^d)$. The multiplier problem then emerges in an attempt to find a converse of the previous statement, i.e.: Which sequences $\lambda \in \mathbb{C}^{\mathbb{Z}^d}$ give rise to bounded operators $T = T_\lambda$ on $L^p(\mathbb{T}^d)$? For L^1 and L^2, we obtain a definite answer.

The notion of multipliers is generalized to an abstract setting in Chapter 2, after first introducing auxiliary machinery related to decomposing a Banach space X into a countable direct sum of closed subspaces X_k such that each $x \in X$ has a unique representation $x = \sum_{k=1}^\infty x_k$, with $x_k \in X_k$. Of particular importance are unconditional Schauder decompositions, for which each of the series $\sum_{k=1}^\infty x_k = x \in X$, $x_k \in X_k$, is convergent regardless of the permutation of the positive integers in the summation. For such decompositions we show that the multiplier transformation

$$T_\lambda \sum_{k=1}^\infty x_k := \sum_{k=1}^\infty \lambda_k x_k$$

is bounded if and only if $\lambda \in \ell^\infty$.

Having found a connection between multipliers and unconditionality, we need to be able to recognize unconditional decompositions. In Chapter 3, convenient characterizations are obtained by examining randomized norms

$$\left| \sum_{k=1}^n \varepsilon_k x_k \right|_{L^p(\Omega;X)},$$

where ε_k are independent, identically distributed, symmetric, $\{-1,1\}$-valued random variables on the probability space Ω, so called Rademacher functions. A useful tool in this connection is the important Khintchine–Kahane inequality, which essentially establishes the equivalence of all L^p norms in the linear span of these functions.

In Chapter 4, we use the randomized norms to introduce the notion of R-boundedness of a family $\mathcal{T} \subset \mathcal{B}(X;Y)$, defined by the requirement

$$\left| \sum_{k=1}^n \varepsilon_k T_k x_k \right|_{L^p(\Omega;Y)} \leq C \left| \sum_{k=1}^n \varepsilon_k x_k \right|_{L^p(\Omega;X)},$$

for all $T_k \in \mathcal{T}$, $x_k \in X$. We give a survey of the basic properties of R-bounds, and the notion of multiplier transformations is generalized to allow for linear bounded operators in place of the scalar multipliers λ_k. The beautiful interplay of R-boundedness with the stochastic characterization of unconditional decompositions leads to interesting theorems for the generalized multiplier transformations with a rather small effort.

The second part of the work is devoted to the theory of UMD-spaces, which provide a convenient setting for vector-valued Fourier multiplier theorems. In Chapter 5, we study the basic theory of vector-valued random variables, which then allows us to introduce the stochastic processes called martingales and establish their basic properties.

The UMD-spaces are finally defined in Chapter 6, by requiring the boundedness of the so called martingale transforms on $L^p(\Omega; X)$. Somewhat lengthy computations show that the definition is independent of the exponent p, and we also obtain a number of other related equivalent conditions, which are useful in different occasions. The equivalence of the various conditions is the content of Burkholder's theorem, which we prove.

In Chapter 7, we prove a number of results providing us with a large variety of UMD-spaces. In particular, every Hilbert space and every reflexive L^p-space is found to be UMD. Furthermore, we show that every UMD-space is reflexive, which implies, by a result concerning the Radon–Nikodým property cited in the Appendix, that the dual of $L^p(\Gamma; X)$ is $L^{\overline{p}}(\Gamma; X^*)$ whenever X is UMD, a very useful fact in view of some duality arguments in later chapters.

One of the main reasons for the usefulness of the UMD-spaces in the present context is the L^p-boundedness of the Hilbert transform on UMD-valued functions. This theorem is established in Chapter 8. The proof is, despite some technical difficulties, perhaps the most beautiful in this presentation: A totally deterministic result, namely, the boundedness of the Hilbert transform, is shown by considering a process of random walk, on which we apply the martingale inequalities. The random walk starts from the origin and eventually, almost surely, crosses the unit-circle at a random point, and when doing so, it gives us an indication of the largeness of a function and its Hilbert transform in this randomly chosen point. Averaging over all possible paths finally leads to the desired result. (It is also true that the boundedness of the Hilbert transform implies the UMD-condition, but due to limited space, we are forced to omit the proof of this converse result.)

Finally, the third and the last part of the work investigates the modern results concerning Fourier multiplier theorems in a vector-valued setting. A close variant of the Hilbert transform is the Riesz projection, whose multiplier is $\mathbf{1}_{[0,\infty)}$, and this is also readily generalized to the d-dimensional setting in the obvious way. An inclusion–exclusion argument shows that an arbitrary box can be expressed as a signed sum of translates of the positive cone $[0,\infty)^d$, and this idea, combined with the boundedness of the Riesz projection and the basic properties of R-bounds, leads to the R-boundedness of all operators whose multipliers are boxes $[\alpha; \beta)$. This immediately yields a variety of simple multiplier results, and improved theorems follow, once we establish the unconditionality of the Schauder decomposition formed of the so called dyadic blocks of the harmonic components of a function; for $f \in L^p(\mathbb{R}; X)$,

$$S_n f := \mathcal{F}^*(\mathbf{1}_{\cup \pm [2^n, 2^{n+1})} \widehat{f}).$$

The unconditionality of such decompositions in the scalar-valued situation (where this is conventionally stated by means of a square function estimate) is at the heart of the classical multiplier theorems of Marcinkiewicz and Mikhlin; the randomization techniques allow one to extend this result to all UMD-spaces, where it also yields strong multiplier theorems.

In Chapter 10, we have a look at the most recent theorems involving operator-valued Fourier multiplier transformations. The R-boundedness techniques are combined with the UMD-theory in a very fruitful manner. We find that the R-boundedness of the sets

$$\{G(t)\}_{t \neq 0} \qquad \text{and} \qquad \{tG'(t)\}_{t \neq 0}$$

is sufficient for the operator-valued function G to give rise to a bounded multiplier transformation, and the R-boundedness of the first of the above mentioned sets is also necessary.

Since our treatment is almost entirely vector-valued, we have included in the Appendix a rather detailed account of various results from vector-valued analysis that we exploit. We investigate integration of vector-valued functions, the Lebesgue–Bôchner spaces $L^p(\Omega; X)$ and their duals, and we present some fundamental results dealing with vector-valued extensions of operators. Furthermore, we give a survey of some differentiability properties of vector-valued functions and a short introduction to Fourier analysis and the theory of distributions in the vector-valued setting.

Due to limited space, it has been impossible to include all the relevant material in this work. Some additional results related to the topic of each chapter are cited in the Notes and comments -sections. Bibliographical sources are indicated at the same place, and we have also mentioned a few points in the text, which are possibly new as far as we know. A lack of a reference in no place indicates any claim of novelty.

Chapter 1

Emergence of Multipliers

1.1 Introduction

In the classical context, the notion of multipliers emerges in Fourier analysis: It turns out that certain important bounded linear transformations of L^p to L^q, $1 \leq p, q < \infty$, namely those that commute with translations $\tau_h f(x) := f(x-h)$, have a multiplier structure when viewed in the Fourier domain. The set of all such transformations will be denoted by $(L^p(U), L^q(U))$, where U will be either $\mathbb{R}^d$ or the d-torus $\mathbb{T}^d$ to be defined below. Note in particular that $(L^p(U), L^q(U)) \subset \mathcal{B}(L^p(U); L^q(U))$.

The topic of this chapter is to explore how the multipliers enter the scene in a most natural way via these translation invariant transformations. This should provide one answer to the question "Why multipliers?", and motivate the study of their abstract versions in the following chapters.

Analogous results exist concerning both Fourier series and the Fourier transform in the periodic and non-periodic cases, respectively. We begin with the periodic case, which requires less preliminaries.

1.2 Fourier series and multipliers

We first fix some notation. Denote by $\mathbb{T}^d := \{(e^{\mathrm{i}2\pi x_1}, \ldots, e^{\mathrm{i}2\pi x_d}) : x \in \mathbb{R}^d\} \subset \mathbb{C}^d$ the **d-torus**, which will be identified with the quotient space $\mathbb{R}^d/\mathbb{Z}^d$ in the natural way. The $\mathbb{Z}^d$-periodic functions on $\mathbb{R}^d$ (i.e., those satisfying $f(x+\kappa) = f(x)$ for each $\kappa \in \mathbb{Z}^d$) to be considered will similarly be identified with functions on $\mathbb{T}^d$. A function is said to be continuous on $\mathbb{T}^d$ if its periodic extension is continuous on $\mathbb{R}^d$.

For purposes of integration we require the concept of a **fundamental domain**. This is defined to be any bounded measurable set $D \subset \mathbb{R}^d$ with the property that, for each $\widetilde{x} \in \mathbb{R}^d/\mathbb{Z}^d$, there is exactly one $x \in \widetilde{x} \cap D$. We will then define integration on $\mathbb{T}^d$, for each periodic measurable f and each finite (on $[0,1)^d$, say) periodic Borel measure μ , by the formula

$$\int_{\mathbb{T}^d} f d\mu := \int_D f d\mu. \tag{1.1}$$

By a periodic measure we of course mean μ defined, for $E = \cup_{\kappa \in \mathbb{Z}^d} E \cap ([0,1)^d + \kappa) =: \cup_{\kappa \in \mathbb{Z}^d} E_\kappa$, by $\mu(E) = \sum_{\kappa \in \mathbb{Z}^d} \mu(E_\kappa - \kappa)$. Observe here that $E_\kappa - \kappa \subset [0,1)^d$. Appropriate restriction must be made on the measurable sets with respect to the periodic measure so as to avoid convergence problems of the complex valued series. If the $E_\kappa - \kappa$ are disjoint, then the corresponding series of absolute values is dominated by $|\mu|([0,1)^d)$, the full variation of μ on $[0,1)^d$. This is the case when E is a fundamental domain as defined above. Furthermore, we have the following lemma, without which the definition of integration above would not make very much sense.

Lemma 1.1. *The integral in (1.1) is independent of the fundamental domain D and thus well-defined.*

Proof. It suffices to consider positive μ, since in every case $\mu \ll |\mu|$, and the integration is defined by $\int f d\mu = \int f \frac{d\mu}{d|\mu|} d|\mu|$. Also, it is only necessary to verify the claim for f being the periodic extension of $\mathbf{1}_E$, where $E \subset [0,1)^d$ is measurable, since the rest follows by standard approximation arguments. (Observe that the set $[0,1)^d$ here is an example of a fundamental domain.) The explicit expression of f will now be $\mathbf{1}_{\cup_{\kappa\in\mathbb{Z}^d}(E+\kappa)}$. We will further decompose D into $\cup_{\ell\in\mathbb{Z}^d} D_\ell$, where $D_\ell := D \cap \left([0,1)^d + \ell\right)$. Then we have

$$\begin{aligned}\int_D f d\mu &= \int_{\mathbb{R}^d} \mathbf{1}_{\cup_{\kappa\in\mathbb{Z}^d}(E+\kappa)} \mathbf{1}_D d\mu = \mu(\{\bigcup_{\kappa\in\mathbb{Z}^d}(E+\kappa)\} \cap \{\bigcup_{\ell\in\mathbb{Z}^d} D_\ell\}) = \mu(\bigcup_{\kappa\in\mathbb{Z}^d}(E+\kappa)\cap D_\kappa) \\ &= \sum_{\kappa\in\mathbb{Z}^d} \mu((E+\kappa)\cap D_\kappa) = \sum_{\kappa\in\mathbb{Z}^d} \mu(E\cap(D_\kappa-\kappa)) = \mu(\bigcup_{\kappa\in\mathbb{Z}^d} E\cap(D_\kappa-\kappa)) = \mu(E\cap\bigcup_{\kappa\in\mathbb{Z}^d}(D_\kappa-\kappa)),\end{aligned}$$

where the periodicity of μ was used in the third to last step. From the fact that D is a fundamental domain and the definition of D_κ it follows that $\cup_{\kappa\in\mathbb{Z}^d}(D_\kappa - \kappa) = [0,1)^d \supset E$. Thus the result of the previous computation is just $\mu(E)$ independently of D. □

Now, for a finite Borel measure μ on $\mathbb{T}^d$, the κth Fourier coefficient, $\kappa \in \mathbb{Z}^d$, of μ is defined by

$$a_\kappa := \int_{\mathbb{T}^d} e^{-\mathbf{i}2\pi\kappa\cdot x} d\mu(x), \tag{1.2}$$

and whenever a_κ are as above, we write

$$\mu \sim \sum_{\kappa\in\mathbb{Z}^d} a_\kappa e^{\mathbf{i}2\pi\kappa\cdot x}. \tag{1.3}$$

Since on $\mathbb{T}^d$ the function spaces L^p can be identified with subspaces of finite Borel measures (by $d\mu := f dm$), the definition of Fourier coefficients immediately extends to such functions. For the moment, the series representation (1.3) may be viewed as a purely formal expression, although, as is well known, important convergence results exist.

Now we state the first classical theorem, which shows the emergence of multipliers.

Theorem 1.2. *If $T \in (L^p(\mathbb{T}^d), L^q(\mathbb{T}^d))$, $1 \le p, q < \infty$, then there exists a unique $\lambda \in \ell^\infty(\mathbb{Z}^d)$ such that for $L^p \ni f(x) \sim \sum_{\kappa\in\mathbb{Z}^d} a_\kappa e^{\mathbf{i}2\pi\kappa\cdot x}$ we have*

$$Tf(x) \sim \sum_{\kappa\in\mathbb{Z}^d} \lambda_\kappa a_\kappa e^{\mathbf{i}2\pi\kappa\cdot x} \tag{1.4}$$

and $|\lambda|_{\ell^\infty(\mathbb{Z}^d)} \le |T|_{\mathcal{B}(L^p(\mathbb{T}^d);L^q(\mathbb{T}^d))}$.

Proof. Since trigonometric polynomials are dense in $L^p(\mathbb{T}^d)$, it suffices to verify the theorem for them, and due to linearity, a proof for an arbitrary $e^{\mathbf{i}2\pi\kappa\cdot x}$, $\kappa \in \mathbb{Z}^d$, will do. Note that these have unity norm in $L^p(\mathbb{T}^d)$.

Let $\phi_\kappa := Te^{\mathbf{i}2\pi\kappa\cdot(\cdot)}$. Then, using the commutativity of T with translations, we have, for each h, $\phi_\kappa(\cdot - h) = \tau_h\phi_\kappa = \tau_h Te^{\mathbf{i}2\pi\kappa\cdot(\cdot)} = T\tau_h e^{\mathbf{i}2\pi\kappa\cdot(\cdot)} = Te^{\mathbf{i}2\pi\kappa\cdot(\cdot-h)} = e^{-\mathbf{i}2\pi\kappa\cdot h}Te^{\mathbf{i}2\pi\kappa\cdot(\cdot)} = e^{-\mathbf{i}2\pi\kappa\cdot h}\phi_\kappa$. This equality holds in the L^p sense, i.e., for all h, for a.e. x, $\phi_\kappa(x-h) - e^{-\mathbf{i}2\pi\kappa\cdot h}\phi_\kappa(x)$.

Now, we would rather like to first fix an x and then have the previous equality for at least a.e. h. This indeed follows: Integrating first with respect to x and then with respect to h the identity $\left|\phi_\kappa(x-h) = e^{-\mathbf{i}2\pi\kappa\cdot h}\phi_\kappa(x)\right| = 0$, we have

$$\int_{\mathbb{T}^d} dh \int_{\mathbb{T}^d} dx \left|\phi_\kappa(x-h) - e^{-\mathbf{i}2\pi\kappa\cdot h}\phi_\kappa(x)\right| = 0,$$

and Fubini's theorem allows us to change the order of integration. The non-negative integrand yielding a zero integral must then vanish for a.e. x, for a.e. h.

It follows that we can fix an x_0 (in fact, almost any) and obtain for a.e. h, thus for almost every $y := x_0 - h$, the above equality, i.e., $\phi_\kappa(y) = e^{-\mathbf{i}2\pi\kappa\cdot(x_0-y)}\phi_\kappa(x_0) =: e^{\mathbf{i}2\pi\kappa\cdot y}\lambda_\kappa$. This means that $Te^{\mathbf{i}2\pi\kappa\cdot x} = \lambda_\kappa e^{\mathbf{i}2\pi\kappa\cdot x}$, and since $T \in \mathcal{B}(L^p; L^q)$, we have $\left|Te^{\mathbf{i}2\pi\kappa\cdot x}\right|_{L^q} = |\lambda_\kappa| \left|e^{\mathbf{i}2\pi\kappa\cdot x}\right|_{L^q} \leq |T|_{\mathcal{B}(L^p;L^q)} \left|e^{\mathbf{i}2\pi\kappa\cdot x}\right|_{L^p}$. It is then clear, recalling that $\left|e^{\mathbf{i}2\pi\kappa\cdot x}\right|_{L^q} = \left|e^{\mathbf{i}2\pi\kappa\cdot x}\right|_{L^p} = 1$, that $\lambda := \{\lambda_\kappa\}_{\kappa\in\mathbb{Z}^d} \in \ell^\infty(\mathbb{Z}^d)$ and $|\lambda|_{\ell^\infty(\mathbb{Z}^d)} \leq |T|_{\mathcal{B}(L^p,L^q)}$. The uniqueness of λ is evident from the construction. □

The reason for calling the T in Theorem 1.2 a (Fourier) multiplier operator should be evident from (1.4). Note, however, that the result of this theorem is only one-sided. This can be improved in the important special cases of L^1 and L^2, which we will now look at. A particularly simple characterization of bounded multiplier operators exists in L^2, where the "if" of Theorem 1.2 can be replaced by "if and only if".

Corollary 1.3. *$T \in (L^2(\mathbb{T}^d), L^2(\mathbb{T}^d))$ if and only if T satisfies (1.4) with $\lambda \in \ell^\infty(\mathbb{Z}^d)$. Furthermore, $|T|_{\mathcal{B}(L^2(\mathbb{T}^d))} = |\lambda|_{\ell^\infty(\mathbb{Z}^d)}$.*

Proof. In $L^2(\mathbb{T}^d)$, a Fourier series $\sum_{\kappa\in\mathbb{Z}^d} a_\kappa e^{\mathbf{i}2\pi\kappa\cdot x}$ converges (in the L^2 norm) if and only if $\sum_{\kappa\in\mathbb{Z}^d} |a_\kappa|^2$ converges (thus the enumeration of $\mathbb{Z}^d$ here is irrelevant, a matter to be discussed in more detail in Section 2.2). Furthermore, if $f \sim \sum_{\kappa\in\mathbb{Z}^d} a_\kappa e^{\mathbf{i}2\pi\kappa\cdot x}$, the Fourier series converges to f, and $|f|_{L^2(\mathbb{T}^d)} = \sum_{\kappa\in\mathbb{Z}^d} |a_\kappa|^2$.

Now, if $\lambda \in \ell^\infty(\mathbb{Z}^d)$, then $\sum_{\kappa\in\mathbb{Z}^d} |\lambda_\kappa a_\kappa|^2 \leq |\lambda|^2_{\ell^\infty(\mathbb{Z}^d)} \sum_{\kappa\in\mathbb{Z}^d} |a_\kappa|^2$, and thus the multiplier operator T defined by (1.4) is a linear operator satisfying $|Tf|_{L^2} \leq |\lambda|_{\ell^\infty} |f|_{L^2}$. From the convergence of the Fourier series and the fact that each multiplier λ_κ certainly commutes with translations it follows that the multiplier operator $T \in (L^2, L^2)$. The rest of the assertion is now a consequence of Theorem 1.2. □

A complete characterization, if not as simple as in the L^2 case, can also be obtained for $(L^1(\mathbb{T}^d), L^1(\mathbb{T}^d))$. The characterization will involve the convolution of a function with a measure; we briefly review the relevant theory, but refer to the appropriate literature for a thorough treatment.

The convolution of finite Borel measures μ, ν on $\mathbb{T}^d$ is defined as the unique Borel measure $\mu * \nu$ on $\mathbb{T}^d$ such that

$$\int_{\mathbb{T}^d} g d(\mu * \nu) = \int_{\mathbb{T}^d}\int_{\mathbb{T}^d} g(x+y) d\mu(x) d\nu(y) \tag{1.5}$$

for every continuous g. It is obvious from this that $\nu * \mu = \mu * \nu$, and it also follows that $|\mu * \nu|(\mathbb{T}^d) \leq |\mu|(\mathbb{T}^d)\,|\nu|(\mathbb{T}^d)$. When one of the measures, say μ, is absolutely continuous with respect to the Lebesgue measure m, and thus has a Radon–Nikodým derivative f, then also $\mu * \nu \ll m$ and

$$\frac{d(\mu * \nu)}{dm} = \int_{\mathbb{T}^d} f(x-y) d\nu(y).$$

If both μ and ν are absolutely continuous, then $\frac{d(\mu*\nu)}{dm} = \frac{d\mu}{dm} * \frac{d\nu}{dm}$, where $*$ in this last expression denotes the ordinary convolution of functions. Thus the two definitions agree, and we can also regard the convolutions of functions with functions and functions with measures as special cases of the general convolution 1.5, identifying f with fdm and $\nu \ll m$ with $\frac{d\nu}{dm}$. In particular, we may define $f * \nu := \nu * f := \frac{d(\mu*\nu)}{dm}$, with $\mu := fdm$. A more explicit expression for this is given above.

Of our interest is the immediate consequence of the definition (1.5) concerning the Fourier coefficients of a convolution:

$$\mu * \nu \sim \sum_{\kappa\in\mathbb{Z}^d} a_\kappa b_\kappa e^{\mathbf{i}2\pi\kappa\cdot x} \qquad \text{for} \quad \mu \sim \sum_{\kappa\in\mathbb{Z}^d} a_\kappa e^{\mathbf{i}2\pi\kappa\cdot x}, \quad \nu \sim \sum_{\kappa\in\mathbb{Z}^d} b_\kappa e^{\mathbf{i}2\pi\kappa\cdot x}. \tag{1.6}$$

By the considerations above, this is equally valid for the convolutions of functions with functions and functions with measures.

Now we proceed towards the multiplier theorem in $L^1(\mathbb{T}^d)$. The proof will exploit some properties of the (modified) Poisson kernel, which are given in the following lemma.

Lemma 1.4. *The modified Poisson kernel*

$$P_r(x) := \sum_{\kappa\in\mathbb{Z}^d} r^{|\kappa|_1} e^{\mathbf{i}2\pi\kappa\cdot x}, \qquad 0<r<1,$$

satisfies $P_r \geq 0$ *and* $|P_r|_{L^1(\mathbb{T}^d)} = 1$.

The actual Poisson kernel has the Euclidean norm of κ in the exponent instead of the 1-norm. The relevant properties in view of Theorem 1.5 below are identical, but they are easier to establish for the modified kernel.

Proof. Note first that

$$P_r(x) = \sum_{k_1,\ldots,k_d\in\mathbb{Z}} r^{\sum_{j=1}^d |k_j|} e^{\mathbf{i}2\pi\sum_{j=1}^d k_j x_j} = \prod_{j=1}^d \sum_{k_j\in\mathbb{Z}} r^{|k_j|} e^{\mathbf{i}2\pi k_j x_j},$$

and that each of the d factors in this last product is absolutely convergent geometric series for $r \in (0,1)$. This guarantees the convergence of the series in the definition of P_r whichever enumeration of $\mathbb{Z}$ we take, and it also justifies the change of the order of summation and integration in the following computation:

$$\int_{\mathbb{T}^d} \sum_{\kappa\in\mathbb{Z}^d} r^{|\kappa|_1} e^{\mathbf{i}2\pi\kappa\cdot x} dx = \sum_{\kappa\in\mathbb{Z}^d} \int_{[0,1]^d} r^{|\kappa|_1} e^{\mathbf{i}2\pi\kappa\cdot x} dx = \sum_{\kappa\in\mathbb{Z}^d} \delta_{k,0} = 1.$$

The second assertion immediately follows from this, once we prove that $P_r \geq 0$, since then the last computation gives the L^1 norm of P_r.

For the positiveness of P_r, we note that each of the d factors in the product expression for P_r can be summed in a closed form (dropping for the moment the subscript j for convenience)

$$1 + \sum_{k=1}^\infty r^k (e^{\mathbf{i}2\pi kx} + e^{-\mathbf{i}2\pi kx}) = 1 + 2\Re\sum_{k=1}^\infty r^k e^{\mathbf{i}2\pi kx} = \Re\frac{1+re^{\mathbf{i}2\pi x}}{1-re^{\mathbf{i}2\pi x}} = \frac{1-r^2}{1-2r\cos(2\pi x)+r^2}.$$

Clearly both the numerator and the denominator are positive for $r \in (0,1)$, and the assertion is hence proved. ☐

Now comes the L^1 theorem.

Theorem 1.5. *An operator* T *is in* $(L^1(\mathbb{T}^d), L^1(\mathbb{T}^d))$ *if and only* T *satisfies (1.4) with* $\lambda = \{\lambda_\kappa\}_{\kappa\in\mathbb{Z}^d}$ *consisting of the Fourier coefficients of a finite Borel measure* μ *on* $\mathbb{T}^d$*; i.e.,* λ_κ *is given by (1.2) with* λ_κ *in place of* a_κ*. Furthermore,* $Tf = \mu * f$ *and* $|\mu|(\mathbb{T}^d) = |T|_{\mathcal{B}(L^1(\mathbb{T}^d))}$.

Proof. Assume $T \in \mathcal{B}(L^1(\mathbb{T}^d))$ and that T commutes with translations. Since $|P_r|_{L^1} = 1$, we have $|TP_r|_{L^1} \leq |T|_{\mathcal{B}(L^1)}$, so $\{TP_r\}_{0<r<1}$ is bounded in L^1. We then consider $L^1(\mathbb{T}^d)$ embedded in the set of finite Borel measures on $\mathbb{T}^d$, with the obvious identification $d\nu := f dm$ for $f \in L^1$. Since the space of finite Borel measures on $\mathbb{T}^d$ is the dual of $C(\mathbb{T}^d)$, we conclude from the Banach–Alaoglu theorem [19] that the closure of the bounded set $\{TP_r\}_{0<r<1}$ is weak*-compact. Hence each sequence of this set, in particular each sequence of the form $\{TP_{r_j}\}_{j=1}^\infty$, $r_j \to 1$ has a subsequence, still denoted by $\{TP_{r_j}\}_{j=1}^\infty$, which is weakly* convergent to a finite Borel measure μ. This means that

$$\int_{\mathbb{T}^d} TP_{r_j} g dx \to \int_{\mathbb{T}^d} g d\mu \qquad g \in C(\mathbb{T}^d).$$

Then for any $\epsilon > 0$ there is a $g \in C(\mathbb{T}^d)$, $|g|_{L^\infty} \leq 1$, such that $|\mu|(T) \leq \int g d\mu + \epsilon = \lim \int TP_{r_j} g dx \leq \lim |TP_{r_j}|_{L^1} \leq |T|_{\mathcal{B}(L^1)}$.

It seems reasonable but requires verification to show that the μ here is the one we want. To this end, we substitute $g(x) := e^{-\mathbf{i}2\pi\kappa\cdot x}$ into the limit statement above; then the left-hand side becomes, by definition, the κth Fourier coefficient of TP_{r_j}, and this is $\lambda_\kappa r_j^{|\kappa|_1}$ by Theorem 1.2 and the series representation of the Poisson kernel, from which the Fourier coefficient are obvious. As $r_j \to 1$, this clearly converges to λ_κ for each $\kappa \in \mathbb{Z}^d$. On the other hand, $\int_{\mathbb{T}^d} e^{-\mathbf{i}2\pi\kappa\cdot x} d\mu(x)$ is, again by definition, the κth Fourier coefficient of μ. Thus $\mu \sim \sum_{\kappa\in\mathbb{Z}^d} \lambda_\kappa e^{\mathbf{i}2\pi\kappa\cdot x}$, where the λ_κ are the multipliers related to T by Theorem 1.2.

On the other hand, if μ is a finite Borel measure with Fourier coefficients $\{\lambda_\kappa\}_{\kappa\in\mathbb{Z}^d}$, then clearly $f \mapsto \mu * f$ is linear and $|\mu * f|_{L^1} \le |\mu|(\mathbb{T}^d)\,|f|_{L^1}$. Furthermore,

$$(\mu * \tau_h f)(x) = \int_{\mathbb{T}^d} (\tau_h f)(x-y)d\mu(y) = \int_{\mathbb{T}^d} f((x-h)-y)d\mu(y) = (\mu * f)(x-h) = \tau_h(\mu * f)(x),$$

whence $\mu * \cdot \in (L^1(\mathbb{T}^d), L^1(\mathbb{T}^d))$. By the property (1.6), it is clear that the multipliers related to $\mu * \cdot$ are the Fourier coefficients λ. Since the multipliers are unique, we combine the two parts of the proof to deduce that $T \in (L^1(\mathbb{T}^d), L^1(\mathbb{T}^d))$ if and only if $T = \mu * \cdot$ for a finite Borel measure μ, and the multipliers of T are the Fourier coefficients of μ. The equality $|T|_{\mathcal{B}(L^1(\mathbb{T}^d))} = |\mu|(\mathbb{T}^d)$ is evident. □

We then leave the Fourier series and proceed to investigate Fourier transforms. These require some preliminaries, which we give in the following sections.

1.3 Some L^p results

We need a couple of L^p concepts to conveniently handle the multipliers in the non-periodic case. The first one of these is the notion of L^p derivative. In a more general context, a perhaps more elegant definition would identify $f \in L^p$ with a tempered distribution, which always has a derivative in the distribution sense, so that no new type of derivative would be needed. We nevertheless stick here to the somewhat old-fashioned notion, because it fits so naturally to the analysis of the linear operators bounded from L^p to L^q and commuting with the translations τ_h; indeed, both of these concepts appear explicitly in the definition of the L^p derivative below, and it will be clear how to exploit these properties in the sequel.

Definition 1.6. *If, for $f \in L^p(\mathbb{R}^d)$ and $h = h_j e_j$, the difference quotients $-\frac{1}{h_j}(\tau_h f - f)$ converge in the L^p norm to a function g, then g is called the L^p* ***(partial) derivative*** *of f with respect to x_j. The L^p derivatives of any order are defined iteratively in the usual way. We denote by $D^\alpha_{L^p} f$ the L^p derivative of f of order $\alpha \in \mathbb{N}^d$, whenever it exists.*

The L^1 derivatives are particularly important because of the property

$$\widehat{D^\alpha_{L^1} f}(x) = (\mathbf{i}2\pi)^{|\alpha|_1} x^\alpha \widehat{f}(x), \tag{1.7}$$

for $f \in L^1$. For $\psi \in \mathcal{S}$, the L^p derivative coincides with the usual one: $D^\alpha_{L^p}\psi = D^\alpha\psi$.

The following lemma gives a criterion showing that sufficient differentiability in the L^p sense implies continuity. The related norm estimate will also be exploited in connection with the multiplier theorems.

Lemma 1.7. *If $f \in L^p(\mathbb{R}^d)$ has all L^p derivatives, $1 \le p \le \infty$, of order $|\alpha|_1 \le n+1$, then there is a continuous g such that $f = g$ a.e. and*

$$|g(0)| \le C(n,p) \sum_{|\alpha|_1 \le n+1} |D^\alpha_{L^p} f|_{L^p}. \tag{1.8}$$

The assertion does not really depend on the particular choice of the point 0, but we will exploit the result in the form stated.

Proof. We first establish the assertion for $p = 1$ and then use an appropriate extrapolation for the general case.

We first observe that $(1+|x|^2)^{\frac{1}{2}(n+1)} \leq (1+\sum_{k=1}^{d}|x_\kappa|)^{n+1} \leq C(n)\sum_{|\alpha|_1\leq n+1}|x^\alpha|$, and thus

$$(1+|x|^2)^{\frac{1}{2}(n+1)}\left|\widehat{f}(x)\right| \leq C\sum_{|\alpha|_1\leq n+1}\left|x^\alpha \widehat{f}(x)\right| = C\sum_{|\alpha|_1\leq n+1}\left|\widehat{D^\alpha_{L^1}f}(x)\right| \leq C\sum_{|\alpha|_1\leq n+1}\left|D^\alpha_{L^1}f\right|_{L^1}.$$

(The property $\left|\widehat{h}\right|_{L^\infty} \leq |h|_{L^1}$, an immediate consequence of the definition of Fourier transform, was used in the last step.) Note that we use C to denote constants depending only on d (and later p), not necessarily the same at every occurrence.

Now the assumption on the L^1 derivatives of f says that the quantity in the right-hand side of the previous inequality is a finite number. Since $(1+|x|^2)^{-\frac{1}{2}(n+1)} \in L^1(\mathbb{R}^d)$, we then see that $\widehat{f} \in L^1$. When this is the case, we know that the inverse Fourier transform of $\widehat{f}$ defines a uniformly continuous function g, which is equal to f in the L^1 sense, i.e., almost everywhere. Furthermore, $|g(0)| \leq |f|_{L^\infty} \leq \left|\widehat{f}\right|_{L^1} \leq C\sum_{|\alpha|_1\leq n+1}\left|D^\alpha_{L^1}f\right|_{L^1}$. Thus the proof of the case $p = 1$ is complete. (Here we actually had a uniform bound for all $|g(x)|$, $x \in \mathbb{R}$.)

For general $p > 1$, multiply an $f \in L^p$ by a $\psi \in \mathcal{D}$. Then $f\psi \in L^p$ is compactly supported and thus the L^1 norms of $f\psi$ as well as of $\frac{1}{h_j}(\tau_h f\psi - f\psi)$ are bounded by the corresponding L^p norms (times a constant depending on the size of the support of ψ). Thus the convergence of the difference quotients in L^p implies convergence in L^1, and it follows that $f\psi$ has derivatives in L^1; iterative application of this observation yields all the same derivatives as for f in L^p. (Also observe that the L^1 and L^p derivatives, when both exist, must coincide, since convergence in either norm implies pointwise convergence a.e. of a subsequence.) Hence $f\psi$ satisfies the assumptions of the lemma for the case $p = 1$ already proved, and we conclude that there is a continuous g_ψ such that $f\psi = g_\psi$, and g_ψ satisfies (1.8) with $f\psi$ in place of f and $p = 1$. To obtain the desired bound we then compute

$$\begin{aligned}\left|D^\alpha_{L^1}(f\psi)\right|_{L^1} = \left|\sum_{\beta\leq\alpha}\binom{\alpha}{\beta}D^{\alpha-\beta}\psi D^\beta_{L^p}f\right|_{L^1} &\leq \int_{\operatorname{supp}\psi}\sum_{\beta\leq\alpha}\binom{\alpha}{\beta}\left|D^{\alpha-\beta}\psi(x)\right|\left|D^\beta_{L^p}f(x)\right|dx \\ &\leq C\sup_{|\gamma|_1\leq n+1}|\psi|_{\gamma,0}\sum_{\beta\leq n+1}\int_{\operatorname{supp}\psi}\left|D^\beta_{L^p}f\right|dx \leq C(\psi)\sum_{\beta\leq n+1}\left|D^\beta_{L^p}f\right|_{L^p}\end{aligned}$$

If we now choose (and fix for the moment) the ψ in such a way that $\psi|_{B(0,r)} = 1$ and $\psi|_{B(0,R)^c} = 0$ for $0 < r < R < \infty$, we find that $f\psi = f$ in $B(0,r)$ and in particular $g_\psi = f$ a.e. in $B(0,r)$. Furthermore, we have a definite bound of the desired form for $|g(0)|$ given by the previous computation. Observe that while this bound depends on ψ, we can fix a definite ψ as above; this ψ will do for any $f \in L^p$, so we have the desired estimate.

Once this is done, we can then take a suitable $\psi \in \mathcal{D}$ with r, R arbitrarily large to show that f is equal to a continuous function g a.e. in every ball. We can deduce the continuity a.e. in every ball centered at the origin, thus continuity a.e. in $\mathbb{R}$, but the constant in the desired inequality may now blow up. This does not matter, since we already deduced that inequality with a definite constant.

The proof is complete. □

Another important concept in L^p theory and elsewhere is the idea of "approximating the identity". This often occurs in applications in the form of convolving a function with something smooth so as to produce an approximation of the function for which certain, otherwise formal, manipulations like differentiation can be performed. The following lemma does not directly involve any concepts of smoothness, but it nevertheless gives a useful approximation criterion to be applied later.

Lemma 1.8. *If $\phi \in L^1(\mathbb{R}^d)$ and for each $\epsilon > 0$ we define $\phi^\epsilon := \epsilon^{-n}\phi(\epsilon^{-1}x)$, then for every $f \in L^p$, $1 \le p < \infty$ or $f \in C_0 \subset L^\infty$ the following convergence holds:*

$$|f * \phi^\epsilon - af|_{L^p} \to 0 \text{ as } \epsilon \to 0, \tag{1.9}$$

where $a := \int_{\mathbb{R}^d} \phi dx$.

The cases of particular interest are $a = 0$ and $a = 1$; these are really all there is in the lemma, since the rest is only a matter of normalization.

Proof. A simple change of variable shows that also $\int_{\mathbb{R}^d} \phi^\epsilon dx = a$ for each $\epsilon > 0$. Thus we can compute

$$\left(\int_{\mathbb{R}^d} |f * \phi^\epsilon(x) - af(x)|^p \, dx\right)^{\frac{1}{p}} = \left(\int_{\mathbb{R}^d} \left|\int_{\mathbb{R}^d} (f(x-y) - f(x))\phi^\epsilon(y) dy\right|^p dx\right)^{\frac{1}{p}}$$
$$\le \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |f(x-y) - f(x)|^p \, dx\right)^{\frac{1}{p}} |\phi^\epsilon(y)| \, dy = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |f(x-\epsilon t) - f(x)|^p \, dx\right)^{\frac{1}{p}} |\phi(t)| \, dt,$$

where Minkowski's integral inequality was applied. Inside the integral we now have the difference norm $|f(\cdot - \epsilon t) - f|_{L^p}$, which is certainly dominated by $2\,|f|_{L^p}$ for each ϵ, and thus the whole of the integrand is dominated by the integrable function $2\,|f|_{L^p}\,|\phi(\cdot)|$. Furthermore, as $\epsilon \to 0$, $|f(\cdot - \epsilon t) - f|_{L^p} \to 0$, certainly if f is continuous with compact support. For general f, we use the density of C_c in each of L^p, $1 \le p < \infty$ and in C_0 to write f as a sum $g + h$, where $g \in C_c$ and the L^p norm of h can be chosen as small as one likes. The assertion of the lemma now follows from the Lebesgue theorem of dominated convergence. □

Another useful approximation criterion is the following.

Lemma 1.9. *For $f \in L^p$, $p \in [1, \infty]$, the norm can be obtained by*

$$|f|_{L^p(\mathbb{R}^d)} = \sup_{\substack{\psi \in \mathcal{D} \\ |\psi|_{L^{\overline{p}}} \le 1}} \int_{\mathbb{R}^d} f\psi dx.$$

Proof. For $p \in (1, \infty)$, this follows from the duality of L^p and $L^{\overline{p}}$ together with the density of $\mathcal{D}$ in $L^{\overline{p}}$.

A function $f \in L^1$ can be approximated by an appropriate simple function $\sum_{j=1}^k z_j \mathbf{1}_{E_j}$ in the L^1 norm as closely as one likes. Then $g := \sum_{j=1}^k \overline{z_j}/\,|z_j|\,\mathbf{1}_{E_j}$ is in L^∞ with norm 1, and the integral of g times the given approximation of f gives the L^1 norm of this approximation exactly. The finite number of indicators $\mathbf{1}_{E_j}$ in the representation of g can then be approximated by functions in $\mathcal{D}$ (Lemma A.19).

For $f \in L^\infty$, for each $\epsilon > 0$ we find a set E of positive measure such that $|f| > |f|_{L^\infty} - \epsilon$ on E. Then $g := \frac{\overline{f}}{|f| m(E)} \mathbf{1}_E \in L^1$ with unity norm, $\int_{\mathbb{R}^d} fg dx > |f|_{L^\infty} - \epsilon$, and $g \in L^1$ can be approximated by functions of $\mathcal{D}$ (Lemma A.19). □

The following simple result is occasionally useful when dealing with translations.

Lemma 1.10. *If $f \in L^p(\mathbb{R}^d)$, $p \in [1, \infty)$, then $|f + \tau_h f|_{L^p(\mathbb{R}^d)} \to 2^{\frac{1}{p}}\,|f|_{L^p(\mathbb{R}^d)}$ as $|h| \to \infty$.*

Note that this is not true in L^∞; a simple counterexample is given by the constant function 1, for which $1 + \tau_h 1 = 2$ for all h.

Proof. Since, for $f \in L^p$, $\int_{B(0;R)^c} |f|^p \, dx \to 0$ as $R \to \infty$, we can, for a given $\epsilon > 0$, choose R so that $\int_{B(0;R)^c} |f|^p \, dx < \epsilon$. Consider $h \in B(0; 2R)^c$. Then

$$|f + \tau_h f|^p_{L^p(\mathbb{R}^d)} \ge \left|(f + \tau_h f)\mathbf{1}_{B(0;R)}\right|^p_{L^p(\mathbb{R}^d)} + \left|(f + \tau_h f)\mathbf{1}_{B(h;R)}\right|^p_{L^p(\mathbb{R}^d)}$$
$$\ge \left(\left|f\mathbf{1}_{B(0;R)}\right|_{L^p(\mathbb{R}^d)} - \left|(\tau_h f)\mathbf{1}_{B(0;R)}\right|_{L^p(\mathbb{R}^d)}\right)^p + \left(\left|(\tau_h f)\mathbf{1}_{B(h;R)}\right|_{L^p(\mathbb{R}^d)} - \left|f\mathbf{1}_{B(h;R)}\right|_{L^p(\mathbb{R}^d)}\right)^p$$
$$> \left(\left|f\mathbf{1}_{B(0;R)}\right|_{L^p(\mathbb{R}^d)} - \epsilon\right)^p + \left(\left|\tau_h f\mathbf{1}_{B(h;R)}\right|_{L^p(\mathbb{R}^d)} - \epsilon\right).$$

In the first inequality, we simply reduce the domain of integration; the second follows from the triangle inequality for the L^p norm, i.e., Minkowski's inequality. The third one exploits the choice of R and h, for we have $B(h;R) \subset B(0;R)^c$ and similarly $B(0;R) \subset B(h;R)^c$; clearly $\int_{A+h} |\tau_h f|^p \, dx = \int_A |f|^p \, dx$ for any measurable A. (Recall $\tau_h f(x) := f(x-h)$.) For $|h|$ sufficiently large we further have

$$\left|f\mathbf{1}_{B(0;R)}\right|_{L^p(\mathbb{R}^d)} = \left|\tau_h f\mathbf{1}_{B(h;R)}\right|_{L^p(\mathbb{R}^d)} > \left|f\mathbf{1}_{B(0;R)}\right|_{L^p(\mathbb{R}^d)} - \epsilon$$

for any given $\epsilon > 0$. Combining this estimate with the previous one and employing the arbitrariness of ϵ to pass to the limit $\epsilon \to 0$, we deduce that

$$\liminf_{|h|\to\infty} |f + \tau_h f|^p_{L^p(\mathbb{R}^d)} \geq 2\,|f|^p_{L^p(\mathbb{R}^d)}\,.$$

The reverse inequality with lim sup in place of lim inf is obvious from the triangle inequality. □

These results at our disposal, we are better prepared to look at the continuous analogues of the multiplier theorems introduced in Section 1.2.

1.4 Fourier transforms and multipliers

The multiplier theorems related to Fourier transforms are conveniently dealt with by using the theory of (tempered) distributions. That is, we will first allow for generalized multipliers, in a sense to be made precise below, and then show that these can be identified with proper functions under certain circumstances. The part of the theory of distributions required here is presented in the Appendix, Section A.6. The setting there is vector valued for later purposes; however, this does not essentially complicate the matters. Also, the theory needed in the present context is quite standard and found in many textbooks.

We begin with the continuous analogue of Theorem 1.2, where the multipliers are merely distributions, as noted above, but improved forms for special cases will be encountered immediately afterwards.

Theorem 1.11. *If $B \in (L^p(\mathbb{R}^d), L^q(\mathbb{R}^d))$, $1 \leq p, q \leq \infty$, then there exists a unique $u \in \mathcal{S}^*$ such that $\widehat{B\phi} = \widehat{u}\widehat{\phi}$ for each $\phi \in \mathcal{S}$.*

Note that we have written the theorem in the given form in order to emphasize the multiplier nature of the result. By Lemma A.43, a statement equivalent to the last equation is given by $B\phi = u * \phi$, and this is what we are going to prove.

Proof. Take a $\phi \in \mathcal{S}$. As a first step, we are going to show that $B\phi$ has L^q derivatives of arbitrary order and that, in fact, $D^\alpha_{L^q} B\phi = B(D^\alpha \phi)$. To see this, take $h = h_j e_j$, e_j the jth vector in the standard basis, and observe that $-\frac{1}{h_j}(\tau_h B\phi - B\phi) = B(-\frac{1}{h_j}(\tau_h \phi - \phi)) \to B(\frac{\partial \phi}{\partial x_j})$ in L^q as $h \to 0$. The linearity and commutativity with translations of B were used in the first step and continuity from L^p to L^q in the second; recall that convergence of the difference quotient to the derivative in $\mathcal{S}$ implies the corresponding convergence in L^p by Lemma A.28. The result for general α follows by induction.

Now we are in a position to apply Lemma 1.7: $B\phi$ has L^q derivatives of arbitrary order, in particular, those of order at most $d+1$, and thus $B\phi$ equals a.e. a continuous functions with which it is henceforth identified (in L^q, we merely pick another member of the equivalence class). Now we employ first the inequality provided by Lemma 1.7, then the commutativity of B and D^α established above, and finally Lemma A.28 to estimate.

$$\begin{aligned} |B\phi(0)| &\leq C \sum_{|\alpha|_1 \leq n+1} |D^\alpha_{L^q}(B\phi)|_{L^q} \leq |B|_{\mathcal{B}(L^p;L^q)}\, C \sum_{|\alpha|_1 \leq n+1} |D^\alpha \phi|_{L^p} \\ &\leq A\,|B|_{\mathcal{B}(L^p;L^q)}\, C \sum_{|\alpha|_1 \leq n+1} \sum_{k=0}^{d} |\phi|_{0,\beta_\kappa}\,. \end{aligned}$$

Here, A, N, C and β are independent of ϕ, so this computation shows that $\phi \mapsto B\phi(0)$ is a continuous (obviously linear) functional on $\mathcal{S}$, thus a member $\widetilde{u}$ of $\mathcal{S}^*$. (The reflection is just for convenience later on.)

Finally, we will show that the $u \in \mathcal{S}^*$ we introduced here is just the u in the assertion that we want. Thus take a $\phi \in \mathcal{S}$ and apply Lemma A.44 to compute $u * \phi(x) = \left\langle u, \tau_x \widetilde{\phi} \right\rangle = \left\langle u, \widetilde{\tau_{-x}\phi} \right\rangle = \langle \widetilde{u}, \tau_{-x}\phi \rangle = B\tau_{-x}\phi(0) = \tau_{-x}B\phi(0) = B\phi(x)$. The assumption that B commutes with translations was used in the second to last step. Note that the point evaluation here certainly makes sense, since $u * \phi$ is continuous, in fact C^∞, by Lemma A.44.

The theorem is now proved. $\square$

As with the multipliers for Fourier series (and nearly always), the L^2 case presents the most beautiful symmetry.

Theorem 1.12. *$B \in (L^2(\mathbb{R}^d), L^2(\mathbb{R}^d))$ if and only if $\widehat{u} \in L^\infty$, where u is as in Theorem 1.11. In particular, $\widehat{u}$ is a proper function. Furthermore, $|\widehat{u}|_{L^\infty(\mathbb{R}^d)} = |B|_{\mathcal{B}(L^2(\mathbb{R}^d))}$.*

Proof. The "if" part follows immediately, since a bounded function g can always be used to define an operator $B \in (L^2(\mathbb{R}^d), L^2(\mathbb{R}^d))$ by $\widehat{B\phi} := g\widehat{\phi}$; in L^2 the definition can be stated in terms of Fourier transforms as here due to Plancherel's theorem. The boundedness of B is obvious by the same theorem, while the commutativity with translations is perhaps most easily seen investigating them, too, in the Fourier domain and recalling that $\widehat{\tau_h f}(x) = e^{-\mathbf{i}2\pi h\cdot x}\widehat{f}(x)$. It is obvious that the multipliers $e^{\mathbf{i}2\pi h\cdot x}$ commute with the multiplier $g(x)$.

Let us then assume that $B \in (L^2(\mathbb{R}^d), L^2(\mathbb{R}^d))$ and $u \in \mathcal{S}^*$ is the related tempered distribution provided by Theorem 1.11. We define $\varphi_0(x) := e^{-\pi|x|^2} \in \mathcal{S}$; this function is useful because it is a fixed point of the Fourier transform operator: $\widehat{\varphi_0} = \varphi_0$, as an easy standard computation shows. (Note that we define the Fourier transform by $\widehat{\psi}(x) = \int_{\mathbb{R}^d} \psi(y)e^{-\mathbf{i}2\pi x\cdot y}dy$; for other close variants of the definition found in the literature, there exists a suitable φ_0 with the same property, as well.)

We further let $f := \widehat{u}\widehat{\varphi_0} = \widehat{u * \varphi_0}$. This is in L^2 by the Plancherel theorem, since (by Theorem 1.11) $u * \varphi_0 = B\varphi_0 \in L^2$, as $\varphi_0 \in L^2$ and $B \in \mathcal{B}(L^2)$. Finally, let $g := \frac{f}{\varphi_0}$. Now certainly g is a proper function (as the quotient of two functions, with the denominator strictly positive). Our intention is to show that $g = \widehat{u}$. As always with distributions, this is done by investigating with a test function; due to the density of $\mathcal{D}$ in $\mathcal{S}$, it in fact suffices to verify the claim for an arbitrary $\psi \in \mathcal{D}$.

Observe that also $\frac{\psi}{\varphi_0} \in \mathcal{D}$, this is a property to be used in a minute. Now $\widehat{u}(\psi) = \widehat{u}(\widehat{\varphi_0}\frac{\psi}{\varphi_0}) = \widehat{u}\widehat{\varphi_0}(\frac{\psi}{\varphi_0}) = \int_{\mathbb{R}^d} f\frac{\psi}{\varphi_0}dx = \int_{\mathbb{R}^d} g\psi dx$, but this is just what we went for.

It is now shown that $\widehat{u} = g$ is a proper function; all that remains is its essential boundedness. To this end, take a $\phi \in \mathcal{S}$ and compute $\left|g\widehat{\phi}\right|_{L^2} = \left|\widehat{u * \phi}\right|_{L^2} = |u * \phi|_{L^2} = |B\phi|_{L^2} \leq |B|_{\mathcal{B}(L^2)} |\phi|_{L^2}$. Since $\mathcal{S}$ is dense in L^2, the inequality holds for all $\phi \in L^2$, and this can clearly only be the case if $\widehat{u} = g \in L^\infty$ with $|\widehat{u}|_{L^\infty} \leq |B|_{\mathcal{B}(L^2)}$. It is easy to see that, in fact, the equality holds. $\square$

We then come to the L^1 case.

Theorem 1.13. *An operator B is in $(L^1(\mathbb{R}^d), L^1(\mathbb{R}^d))$ if and only if the related distribution u given by Theorem 1.11 is a finite Borel measure μ. In this case, $|\mu|(\mathbb{R}^d) = |B|_{\mathcal{B}(L^1(\mathbb{R}^d))}$.*

Proof. If $u = \mu$ is a finite Borel measure, then the convolution with an $f \in L^1$ is given by

$$\mu * f(x) := f * \mu(x) := \int_{\mathbb{R}^d} f(x-y)d\mu(y), \tag{1.10}$$

and this satisfies $|\mu * \cdot|_{\mathcal{B}(L^1(\mathbb{R}^d))} \leq |\mu|(\mathbb{R}^d)$ as is readily verified integrating the previous expression with respect to x over $\mathbb{R}^d$ and employing Fubini's theorem to change the order of integration. From (1.10) it is also clear that $\mu * \cdot$ commutes with translations.

To show the converse, we take the $u \in \mathcal{S}^*$ related to B as in Theorem 1.11. We then define $u_\epsilon := u * W(\cdot, \epsilon^2)$, where

$$W(x, \alpha) := \frac{1}{(4\pi\alpha)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4\alpha}}, \quad \alpha > 0, \tag{1.11}$$

is the Gauss–Weierstrass kernel. Two properties of this kernel interest us: $W(\cdot, \alpha) \in \mathcal{S}$ for each $\alpha > 0$, and $\int_{\mathbb{R}^d} W(x, \alpha)dx = 1$ as is readily verified by standard computations. We also note that $W(\cdot, \epsilon^2) = W^\epsilon(\cdot, 1)$, where the superscript ϵ has the same meaning as in Lemma 1.8; in fact, we are going to use that lemma pretty soon.

We first observe that $|u_\epsilon|_{L^1} = \left|BW(\cdot, \epsilon^2)\right|_{L^1} \leq |B|_{\mathcal{B}(L^1)} \left|W(\cdot, \epsilon^2)\right|_{L^1} = |B|_{\mathcal{B}(L^1)}$. Thus u_ϵ is a bounded sequence in L^1. We then consider L^1 embedded in the space of finite Borel measures on $\mathbb{R}^d$ (by the natural identification $d\nu := fdm$ for $f \in L^1$). We know that this space of finite Borel measures is the dual of $C_0(\mathbb{R}^d)$. As in the proof of Theorem 1.5 for the periodic case, we invoke the Banach–Alaoglu theorem to deduce the existence of a sequence $\{\epsilon_\kappa\}_{k=1}^\infty$ with $\epsilon_\kappa \to 0$, for which u_{ϵ_κ} converges weakly* to a finite Borel measure μ, i.e.,

$$\int_{\mathbb{R}^d} g u_{\epsilon_\kappa} dm \to \int_{\mathbb{R}^d} g d\mu$$

for each $g \in C_0(\mathbb{R}^d)$.

We claim that this μ is the one we are looking for, and for this we must show that $u(\psi) = \int_{\mathbb{R}^d} \psi d\mu$ for $\psi \in \mathcal{S}$. To this end, we investigate $\psi_\epsilon := \psi * W(\cdot, \epsilon^2) = \psi * W^\epsilon(\cdot, 1)$. We wish to show that $\psi_\epsilon \to \psi$ in $\mathcal{S}$ as $\epsilon \to 0$. If we can do this, then the rest of the argument follows easily: We have $u(\psi_\epsilon) = u(\psi * W(\cdot, \epsilon^2)) = u(\widetilde{W}(\cdot, \epsilon^2) * \psi) = u * W(\cdot, \epsilon^2)(\psi)$ and for $\epsilon = \epsilon_\kappa$, the right-hand side tends to $\int_{\mathbb{R}^d} \psi d\mu$, whereas the left-hand side tends to $u(\psi)$; thus these quantities are equal by the uniqueness of the limit. (We used above the symmetry $W = \widetilde{W}$ to add the reflection without altering the value.)

We hence turn to investigate the desired convergence of ψ_ϵ. By Lemma 1.8, for $D^\alpha \psi \in C_0$, we know that $D^\alpha \psi_\epsilon(x) = \int_{\mathbb{R}^d} D^\alpha \psi(x-y) W^\epsilon(y,1)dy \to D^\alpha \psi(x)$ uniformly. Furthermore, since $x^\beta = (x - y + y)^\beta = \sum_{\gamma \leq \beta} \binom{\beta}{\gamma} (x-y)^{\beta-\gamma} y^\gamma$, we can compute

$$x^\beta \psi_\epsilon(x) = \int_{\mathbb{R}^d} (x-y)^\beta \psi(x-y) W^\epsilon(y,1)dy + \sum_{0 \neq \gamma \leq \beta} \binom{\beta}{\gamma} \int_{\mathbb{R}^d} (x-y)^{\beta-\gamma} \psi(x-y) y^\gamma W^\epsilon(y,1)dy.$$

The first term here converges to $x^\alpha \psi(x)$ uniformly in x by the same argument as before. For the other terms we observe that $y^\gamma W^\epsilon(y,1) = \epsilon^{|\gamma|_1} (\epsilon^{-1} y)^\gamma W^\epsilon(y,1) = \epsilon^{|\gamma|_1} (y^\gamma W(y,1))^\epsilon$. Here, $y^\gamma W(y,1)$ is an integrable function on $\mathbb{R}^d$. Denoting its integral over $\mathbb{R}^d$ by a_γ, Lemma 1.8 shows that $\int_{\mathbb{R}^d} (x-y)^{\beta-\gamma} \psi(x-y)(y^\gamma W(y,1))^\epsilon dy \to a_\gamma x^{\beta-\gamma} \psi(x)$ as $\epsilon \to 0$ uniformly in x. Thus $\epsilon^{|\gamma|_1}$ times this quantity tends to zero for $\gamma \neq 0$. Hence all that remains in the limit is the first term, which has the desired convergence. The convergence of $x^\beta D^\alpha \psi_\epsilon(x)$ follows by the same argument by writing $D^\alpha \psi$ in place of ψ. Thus the convergence $\psi_\epsilon \to \psi$ in $\mathcal{S}$ is established, and the proof is complete. □

For certain values of p and q, we have a characterization of $(L^p(\mathbb{R}^d), L^q(\mathbb{R}^d))$ even simpler than in the cases of L^1 and L^2 examined above. It turns out that all bounded linear operators from L^p to L^q commuting with translations must vanish if $q < p < \infty$. Thus the only interesting cases with finite exponents can occur for $p \leq q$. (Note that we have already found a large number of non-trivial operators in (L^1, L^1) as well as in (L^2, L^2), so that one need not be afraid of the whole work being a characterization of the zero operator.)

Proposition 1.14. *For $1 \leq q < p < \infty$, we have $(L^p(\mathbb{R}^d), L^q(\mathbb{R}^d)) = \{0\}$.*

Proof. For an arbitrary but fixed $f \in L^p$, we examine the function $f + \tau_h f$. Given $B \in (L^p, L^q)$, we have $B(f + \tau_h f) = Bf + \tau_h Bf$. Then $|Bf + \tau_h Bf|_{L^q} \leq |B|_{\mathcal{B}(L^p;L^q)} |f + \tau_h f|_{L^p}$. Taking the limit $|h| \to \infty$ in the previous inequality and employing Lemma 1.10, we deduce $2^{\frac{1}{q}} |Bf|_{L^q} \leq$

$2^{\frac{1}{p}} |B|_{\mathcal{B}(L^p;L^q)} |f|_{L^p}$. Assuming $B \neq 0$ and taking the supremum over all $f \in L^p$ with $|f|_{L^p} \leq 1$ we obtain $2^{\frac{1}{q}} \leq 2^{\frac{1}{p}}$, which is clearly false for $q < p$; thus the assumption $B \neq 0$ resulted in a contradiction and must be false. $\square$

To conclude this section, we present a duality property of the spaces $(L^p(\mathbb{R}^d), L^q(\mathbb{R}^d))$.

Proposition 1.15. *For $p, q \in [1, \infty]$, we have*

$$(L^p(\mathbb{R}^d), L^q(\mathbb{R}^d)) = (L^{\overline{q}}(\mathbb{R}^d), L^{\overline{p}}(\mathbb{R}^d)),$$

and the identity mapping is an isometric isomorphism between the two spaces.

Proof. For $B \in (L^p, L^q)$ $(p, q \in [1, \infty])$, Theorem 1.11 provides us with a unique $u \in \mathcal{S}^*$ such that $B\phi = u * \phi$ for all $\phi \in \mathcal{S}$. Then, for $\phi, \psi \in \mathcal{S}$,

$$\int_{\mathbb{R}^d} (u * \phi)\psi dx = \langle u * \phi, \psi \rangle = \left\langle u, \widetilde{\phi} * \psi \right\rangle = \left\langle u, \widetilde{\phi * \widetilde{\psi}} \right\rangle = \left\langle \widetilde{u}, \widetilde{\psi} * \phi \right\rangle = \langle \widetilde{u} * \psi, \phi \rangle = \int_{\mathbb{R}^d} (\widetilde{u} * \psi)\phi dx \quad (1.12)$$

The first and last steps make sense, since by Lemma A.44, $u * \phi$ (similarly $\widetilde{u} * \psi$) can be identified with a slowly increasing function; the product of such a function with $\psi \in \mathcal{S}$ (or ϕ) is again in $\mathcal{S}$, in particular in L^1.

If $u * \cdot \in (L^p, L^q)$, it follows by Hölder's inequality that

$$\int_{\mathbb{R}^d} (\widetilde{u} * \psi)\phi dx \leq |u * \phi|_{L^q} |\psi|_{L^{\overline{q}}} \leq |u * \cdot|_{\mathcal{B}(L^p;L^q)} |\phi|_{L^p} |\psi|_{L^{\overline{q}}}.$$

Taking now the supremum over all $\phi \in \mathcal{S}$ with $|\phi|_{L^p} \leq 1$ gives, by Lemma 1.9, $|\widetilde{u} * \psi|_{L^{\overline{p}}} \leq |u * \cdot|_{\mathcal{B}(L^p;L^q)} |\psi|_{L^{\overline{q}}}$ for $\psi \in \mathcal{S}$. But this shows that $\widetilde{u} * \cdot \in (L^{\overline{q}}, L^{\overline{p}})$ whenever $u * \cdot \in (L^p, L^q)$, and $|\widetilde{u} * \cdot|_{\mathcal{B}(L^{\overline{q}};L^{\overline{p}})} \leq |u * \cdot|_{\mathcal{B}(L^p;L^q)}$. The reverse inclusion and inequality follow by changing the roles of p and q with their conjugate exponents, and the proof is complete after observing that $u * \cdot$ is in one of the spaces (L^p, L^q) if and only if $\widetilde{u} * \cdot$ is, and the norms agree. $\square$

Corollary 1.16. *$B = u * \cdot \in (L^\infty(\mathbb{R}^d), L^\infty(\mathbb{R}^d))$ if and only if u is a finite Borel measure μ. When this is the case, $|\mu|(\mathbb{R}^d) = |B|_{\mathcal{B}(L^\infty(\mathbb{R}^d))}$.*

Proof. This is immediate from Theorem 1.13 and Proposition 1.15. $\square$

1.5 Notes and comments

This chapter is mostly based on Stein and Weiss [24]. Proofs of the relevant results related to the convolutions of measures are indicated in Rudin [20].

The introduction of the modified Poisson kernel for the proof of Theorem 1.5 simplifies the background work. The properties of this kernel which are relevant in this context are the same as those of the actual Poisson kernel, but they are established more readily for the modified version; the easy computations in Lemma 1.4 follow [20].

The chapter is definitely not exhaustive in the treatment of classical theorems on Fourier multipliers. More results, for instance concerning interrelations between $(L^p(\mathbb{T}^d), L^q(\mathbb{T}^d))$ and $(L^p(\mathbb{R}^d), L^q(\mathbb{R}^d))$ are covered in Stein and Weiss [24].

Chapter 2

Decompositions of Banach Spaces

2.1 Introduction

In this chapter, we present an abstract framework which will guide us in generalizing the classical multiplier theorems. We are motivated by the form of the multipliers acting on functions on the torus: The function is first decomposed into the harmonic components, and the action of the multiplier operator is defined separately for each of these components. In order to generalize this idea, we now investigate ways of decomposing vectors of a Banach space X (a complete linear normed space) into sums $x = \sum_k x_k$, where $x_k \in X_k$ and X_k are distinct subspaces of X. An important special case of such a decomposition is the representation of a vector in terms of a linear basis.

A well-known fundamental result in the theory of linear spaces says that every non-trivial vector space has a Hamel basis, a linearly independent collection of vectors, such that every vector of the space can be expressed as a finite linear combination of the basis vectors. This is a very general result, a consequence of the very axioms of a vector space (and the axiom of choice), and no reference to any convergence or topology is involved. However, since convergence plays a prominent role in all analysis and since Banach spaces are not only linear normed spaces (thus metric and hence topological spaces with a rich structure) but also complete, it should not be surprising that the Hamel bases are not the most useful tool in this connection. In Hilbert spaces, an even more restricted class of linear spaces, an orthonormal basis is of course desirable, but it is clear that this concept is not as such applicable in a general Banach space.

The concept which proves to be useful is that of a Schauder basis, and its generalization, the Schauder decomposition. A Schauder basis is a countable sequence $\{e_n\}_{n=1}^{\infty} \subset X$, X Banach, such that each $x \in X$ has a unique representation $x = \sum_{k=1}^{\infty} \xi_k e_k$ with ξ_k scalars. As a preparation for definitions of various types of convergence of the infinite sum above and for the formulation of various properties of these bases, we next examine some properties of series.

2.2 Convergence of series

The following definition presents the various kinds of convergence of interest.

Definition 2.1. *Let $\{x_n\}_{n=1}^{\infty} \subset X$ be a sequence in the linear normed space X. We call the series $\sum_{n=1}^{\infty} x_n$*

1. ***convergent to** $x \in X$, if $\lim_{N\to\infty} \left| \sum_{n=1}^{N} x_n - x \right|_X = 0$, denoted by $\sum_{n=1}^{\infty} x_n = x$;*

2. ***unconditionally convergent**, if $\sum_{n=1}^{\infty} x_{\sigma(n)}$ converges for every $\sigma \in S_{\mathbb{Z}_+}$, i.e., every permutation of the natural numbers;*

3. ***summable to*** *$x \in X$, if for all $\epsilon > 0$ there is an $F_0 \subset \mathbb{Z}_+$, $\#F_0 < \infty$, such that for all $F_0 \subset F \subset \mathbb{Z}_+$, $\#F < \infty$, we have*
$$\left|\sum_{k\in F} x_k - x\right|_X < \epsilon;$$

4. ***shrinking,*** *if for all $\epsilon > 0$ there is an $F_0 \subset \mathbb{Z}_+$, $\#F_0 < \infty$, such that for all $F \subset \mathbb{Z}_+ \setminus F_0$, $\#F < \infty$, we have*
$$\left|\sum_{k\in F} x_k\right|_X < \epsilon;$$

5. ***absolutely convergent,*** *if the series $\sum_{n=1}^\infty |x_n|_X$ of positive real numbers is convergent.*

In a Banach space X, several but not all of these convergence types are equivalent. Recall that the property that absolute convergence implies convergence characterizes completeness of normed linear spaces. Thus in a Banach space it is clear that an absolutely convergent series is also unconditionally convergent, since the absolutely convergent series can be permuted according to elementary analysis. While the converse is true in the finite-dimensional setting, this is not the case in general. For instance, let $\mathcal{H}$ be an infinite-dimensional Hilbert space with an orthonormal system $\{e_n\}_{n=1}^\infty \subset \mathcal{H}$. Then the series $\sum_{n=1}^\infty \alpha_n e_n$ converges if and only if $\sum_{n=1}^\infty |\alpha_n|^2$ converges, and as this last expression is a series of positive reals, it converges if and only if each of its permutations converges. Thus the same is true for $\sum_{n=1}^\infty \alpha_n e_n$. However, the series $\sum_{n=1}^\infty \alpha_n e_n$ converges absolutely if and only if $\sum_{n=1}^\infty |\alpha_n|$ converges. Since sequences $\{\alpha_n\}_{n=0}^\infty \in \ell^2 - \ell^1$ are abundant, a standard example being $\{\frac{1}{n}\}_{n=1}^\infty$, it is clear in this case that absolute convergence is a strictly stronger property (and typically too strong to be of interest) than unconditional convergence. In fact, a result by Dvoretzky and Rogers asserts that the equivalence of absolute and unconditional convergence is characteristic of finite-dimensional Banach spaces [8]. (The discussion above shows this in the case of a Hilbert space.)

Lemma 2.2. *Let $\{x_k\}_{k=1}^\infty \subset X$, where X is a Banach space. The following statements are equivalent:*

1. *$\sum_{k=1}^\infty x_k$ is unconditionally convergent.*
2. *$\sum_{k=1}^\infty x_k$ is summable.*
3. *$\sum_{k=1}^\infty x_k$ is shrinking.*
4. *$\sum_{k=1}^\infty \lambda_k x_k$ is convergent for each bounded sequence $\{\lambda_k\}_{k=1}^\infty \in \ell^\infty(\mathbb{Z}_+)$.*
5. *$\sum_{k=1}^\infty \epsilon_k x_k$ is convergent for each sequence $\{\epsilon_k\}_{k=1}^\infty \in \{-1,1\}^{\mathbb{Z}_+}$.*
6. *$\sum_{k=1}^\infty \delta_k x_k$ is convergent for each sequence $\{\delta_k\}_{k=1}^\infty \in \{0,1\}^{\mathbb{Z}_+}$.*

Furthermore, the sum of an unconditionally convergent series is independent of the permutation, even if X is merely a linear normed space (not necessarily complete).

Proof. We first derive the uniqueness of the sum of an unconditionally convergent series from the corresponding result for real or complex numbers. If $\sum_{k=1}^\infty x_{\sigma(k)}$ converges for an arbitrary permutation σ, then so does $\sum_{k=1}^\infty \langle x^*, x_{\sigma(k)}\rangle$ for each $x^* \in X^*$. Now, in the scalar case, the sum must be independent of the permutation, thus $\langle x^*, \sum_{k=1}^\infty x_{\sigma(k)}\rangle$ is the same for any two $\sigma \in S_{\mathbb{Z}_+}$ for *all* $x^* \in X^*$. Since X^* separates the points of X, the sum must be equal for any two permutations.

$1 \Rightarrow 2$. Let $x := \sum_{k=1}^\infty x_k$, where the series converges unconditionally. We claim that $\sum_{k=1}^\infty x_k$ is summable to x, and that for a given ϵ we can take $F_0 := \{1,\dots,N\}$ for some N. If this is not the case, then for arbitrarily large N we find a finite $F \subset \mathbb{Z}_+ + N$ so that

$\left|\sum_{k\in F\cup F_0} x_k - x\right|_X > \epsilon$. On the other hand, since $\sum_{k=1}^{\infty} x_k$ is convergent to x, we have, for large enough n, that $\left|\sum_{k\in F_0} x_k - x\right|_X < \frac{1}{2}\epsilon$. For such n and the corresponding F_0 and F as above, we have $\left|\sum_{k\in F} x_k\right|_X \geq \left|\sum_{k\in F_0\cup F} x_k - x\right|_X - \left|\sum_{k\in F_0} x_k - x\right|_X > \epsilon - \frac{1}{2}\epsilon = \frac{1}{2}\epsilon$. Since these considerations were valid for all large enough n, we pick some n_1 and the corresponding finite F_1 and inductively choose $n_k > \max F_{k-1}$ and a suitable finite F_k. Let $N_k := \{n_k, \ldots, n_{k+1} - 1\}$ and let $\sigma \in S_{\mathbb{Z}_+}$ permute each of the N_k in such a way that the appropriate number of lowest values $n_k, \ldots, n_k'$ of N_k are mapped onto $F_k \subset N_k$. Then the series $\sum_{k=1}^{\infty} x_{\sigma(k)}$ cannot converge, since it fails the Cauchy criterion as the sequence of differences of its partial sums $\sum_{j=n_k}^{n_k'} x_{\sigma(k)} = \sum_{k\in F_k} x_k$ does not tend to 0.

$2 \Rightarrow 3$. This follows immediately by taking, for a given $\epsilon > 0$, the set F_0 as in the definition of summability. For a finite $F \subset \mathbb{Z}_+ \setminus F_0$ we then have $\left|\sum_{k\in F} x_k\right|_X \leq \left|\sum_{k\in F\cup F_0} x_k - x\right|_X + \left|\sum_{k\in F_0} x_k - x\right|_X < 2\epsilon$.

$3 \Rightarrow 4$. Investigate a fixed but arbitrary sequence $\lambda \in \ell^\infty$. By separately considering each of the four sequences in the decomposition $\lambda_k = |\lambda|_{\ell^\infty}(\lambda_k^+ - \lambda_k^- + \mathbf{i}\mu_k^+ - \mathbf{i}\mu_k^-)$ if necessary, we may assume, without loss of generality, that $\lambda_k \in [0,1]$. If the sequence in 4 is not convergent and thus not Cauchy, we can find an $\epsilon > 0$ and arbitrarily large n, m such that $\left|\sum_{k=n}^{m} \lambda_k x_k\right|_X > \epsilon$. Here $(\lambda_n, \ldots, \lambda_m) \in [0,1]^N = \operatorname{conv}\{0,1\}^N$, where $N = m - n + 1$. Denoting by $\{\alpha^j\}_{j=1}^{2^N}$ an enumeration of the vertices of the hypercube $[0,1]^N$, it is then possible to express any point of the cube (the convex hull of the vertices) as a convex combination of the vertices, in particular, $(\lambda_n, \ldots, \lambda_m) = \sum_{j=1}^{2^N} \nu^j \alpha^j$, i.e., $\lambda_k = \sum_{j=1}^{2^N} \nu^j \alpha_k^j$ for $k = n, \ldots, m$, where $\alpha_k^j \in \{0,1\}$ and $\nu^j \geq 0$ with $\sum_{j=1}^{2^N} \nu^j = 1$. Then

$$\begin{aligned}\epsilon < \left|\sum_{k=n}^{m} \lambda_k x_k\right|_X &= \left|\sum_{j=1}^{2^N} \nu^j \sum_{k=n}^{m} \alpha_k^j x_k\right|_X \leq \sum_{j=1}^{2^N} \nu^j \left|\sum_{k=n}^{m} \alpha_k^j x_k\right|_X \\ &\leq \max_{\alpha\in\{0,1\}^N} \left|\sum_{k=n}^{m} \alpha_k x_k\right|_X = \left|\sum_{k\in F} x_k\right|_X,\end{aligned}$$

for a certain finite set $F \subset \{n, \ldots, m\}$. Since this happens infinitely often, $\sum_{k=1}^{\infty} x_k$ cannot be shrinking.

$4 \Rightarrow 5$. This is trivial.

$5 \Rightarrow 6$. This follows directly, since any δ sequence is obtained from a certain ϵ sequence by $\delta_k = \frac{1}{2}(\epsilon_k + 1)$, and since the convergence of the series $\sum_{k=1}^{\infty} x_k$ is included in the assumption 5 (taking $\epsilon_k := 1$ for all k). If 5 holds, then $\sum_{k=1}^{\infty} \delta_k x_k = \frac{1}{2}\sum_{k=1}^{\infty} \epsilon_k x_k + \frac{1}{2}\sum_{k=1}^{\infty} x_k$, and the left-hand side converges, since the right-hand side does.

$6 \Rightarrow 1$. If $\sum_{k=1}^{\infty} x_{\sigma(k)}$ is not convergent, then the sequence of partial sums is not Cauchy, and it is not true that $\left|\sum_{k=n}^{m} x_{\sigma(k)}\right|_X \to 0$ as $n, m \to \infty$. Thus for some $\epsilon > 0$ we find arbitrarily large n, m such that $\left|\sum_{k=n}^{m} x_{\sigma(k)}\right|_X \geq \epsilon$. Pick $N_1 := \{n_1, \ldots, m_1\}$ with such n_1, m_1. Inductively choose $N_k := \{n_k, \ldots, m_k\}$ such that $p_k := \min \sigma(N_k) > \max \sigma(N_{k-1}) =: q_{k-1}$. (This is possible, since $\sigma(n) \to \infty$ as $n \to \infty$ for any $\sigma \in S_{\mathbb{Z}_+}$.) Define a δ sequence by $\delta_j := 1$ if $j \in \cup_{k=1}^{\infty} \sigma(A_k)$ and 0 otherwise. Then $\sum_{k=1}^{\infty} \delta_k x_k$ cannot converge, since the differences of partial sums $\sum_{j=p_k}^{q_k} \delta_j x_j = \sum_{j\in A_k} x_{\sigma(j)}$ do not tend to 0.

□

In the proof of the implication $3 \Rightarrow 4$ we obtained a byproduct, which we state as a separate result for later reference. It is obvious that the same argument applies with $[0,1]$ replaced by any interval $[a,b]$; furthermore, completeness was not relevant in this step.

Lemma 2.3. *For $\lambda_1, \ldots, \lambda_N \in [a,b]$, $x_1, \ldots, x_N \in X$, where X is a linear normed space, the following inequality holds:*

$$\left|\sum_{k=1}^{N} \lambda_k x_k\right|_X \leq \max_{\alpha\in\{a,b\}^N} \left|\sum_{k=1}^{N} \alpha_k x_k\right|.$$

Furthermore, the proof of this fact essentially relied on the obvious geometric fact $[0,1]^n = \operatorname{conv}\{0,1\}^n$. This can also be viewed as a consequence of a more general result concerning convex hulls of Cartesian products. (Note that $[0,1]^n = [0,1] \times \cdots \times [0,1]$.) We also find this a natural place to state this result, which has some use later on:

Lemma 2.4. *Let W_i, $i = 1, \ldots, n$, be subsets of (possibly different) vector spaces. Then*

$$\operatorname{conv}(W_1 \times \cdots \times W_n) = \operatorname{conv}(W_1) \times \cdots \times \operatorname{conv}(W_n).$$

Proof. If $(x_i)_{i=1}^n \in \prod_{i=1}^n \operatorname{conv}(W_i)$, then the x_i are convex combinations $x_i = \sum_{j_i=1}^N \lambda_i^{j_i} w_i^{j_i}$ of $w_i^{j_i} \in W_i$. (We can choose the same N for each i, possibly after augmenting some terms with $\lambda_i^{j_i} = 0$.) We can always write the previous expression in the form $x_i = \sum_{j_1,\ldots,j_n=1}^N \prod_{k=1}^n \lambda_k^{j_k} w_i^{j_i}$, since $\sum_{j_k=1}^N \lambda_k^{j_k} = 1$, and thus

$$(x_i)_{i=1}^n = \sum_{j_1,\ldots,j_n=1}^N \prod_{k=1}^n \lambda_k^{j_k} (w_i^{j_i})_{i=1}^n \in \operatorname{conv}(\prod_{i=1}^n W_i),$$

since obviously (any enumeration of) $\{\prod_{k=1}^n \lambda_k^{j_k}\}_{j_1,\ldots,j_n=1}^N$ is a proper set of coefficients for a convex combination. Thus $\prod_{i=1}^n \operatorname{conv}(W_i) \subset \operatorname{conv}(\prod_{i=1}^n W_i)$.

For the converse, it is easy to see that the Cartesian product of convex sets is again convex. Also, it is clear that $\prod_{i=1}^n \operatorname{conv}(W_i) \supset \prod_{i=1}^n W_i$, and it follows that $\prod_{i=1}^n \operatorname{conv}(W_i) \supset \operatorname{conv}(\prod_{i=1}^n W_i)$, since this last expression is the smallest convex set containing the Cartesian product of the W_i, $i = 1, \ldots, n$. □

2.3 Schauder decompositions

We begin with the definitions of Schauder decompositions and bases, making use of the concepts of convergence introduced above.

Definition 2.5. *A **Schauder decomposition** of a Banach space X is a sequence $D = \{D_k\}_{k=1}^\infty \subset \mathcal{B}(X)$ of projections such that*

1. *$D_k D_\ell = 0$ for $k \neq \ell$,*
2. *$x = \sum_{k=1}^\infty D_k x$ for all $x \in X$.*

The operators $P_n = \sum_{k=1}^n D_k$ are called the partial sum projections corresponding to D.

*A **Schauder basis** is a sequence $\{e_k\}_{k=1}^\infty$ such that each $x \in X$ has a unique representation $x = \sum_{k=1}^\infty \xi_k e_k$, $\xi_k \in K$.*

*A Schauder decomposition or basis is called **unconditional**, if the series in the corresponding definition converges unconditionally for every $x \in X$. The **range** of a Schauder decomposition is the set $\operatorname{ran}(D) := \cup_{n=1}^\infty \operatorname{ran} P_n$, i.e., all the vectors whose series representation in the decomposition has only finitely many non-zero terms.*

The following result gives a simple but useful characterization of Schauder decompositions.

Lemma 2.6. *A collection $D = \{D_k\}_{k=1}^\infty \subset \mathcal{B}(X)$ satisfying $D_k D_\ell = \delta_{k\ell} D_\ell$ is a Schauder decomposition of X if and only if $\overline{\operatorname{ran}}(D) = X$ and the partial sum projectors P_n are uniformly bounded. Then the projectors D_k are also uniformly bounded.*

Proof. If D is a Schauder decomposition, then $\operatorname{ran}(D) \ni P_n x \to x$ as $n \to \infty$ for all $x \in X$, so obviously $\operatorname{ran}(D)$ is dense. From $P_n x \to x$ it also follows that $|P_n x|_X$ is bounded in n for each fixed x, and the boundedness in n of $|P_n|_{\mathcal{B}(X)}$ follows from the principle of uniform boundedness. That D_n, too, are uniformly bounded follows from the identity $D_n = P_n - P_{n-1}$.

Conversely, suppose $\operatorname{ran}(D)$ is dense and $|P_n|_{\mathcal{B}(X)} \leq K$ for all $n \in \mathbb{Z}_+$. To show that D is a Schauder decomposition, we must prove that $x = \sum_{k=1}^{\infty} D_k x$, i.e., that $P_n x \to x$ as $n \to \infty$ for each $x \in X$. As usual, let $\epsilon > 0$ be given.

By the density of $\operatorname{ran}(D)$, we can find $x_k \in \operatorname{ran}(D_k)$, $k = 1, \ldots, N$, such that the estimate $\left|\sum_{k=1}^{N} x_k - x\right|_X < \epsilon_1 := \frac{\epsilon}{K+1}$ is valid. For $n \geq N$, $\sum_{k=1}^{N} x_k \in \operatorname{ran}(P_n)$; thus $P_n \sum_{k=1}^{N} x_k = \sum_{k=1}^{N} x_k$. We then compute

$$|P_n x - x|_X = \left| P_n \left(x - \sum_{k=1}^{N} x_k \right) + \left(\sum_{k=1}^{n} x_k - x \right) \right|_X < (K+1)\epsilon_1 = \epsilon.$$

Thus for given $\epsilon > 0$ there is an N such that for all $n \geq N$ we have $|P_n x - x|_X < \epsilon$, but this is just what we went for. □

From the similarity of the definitions of a Schauder decomposition and basis, one might suspect that they have something to do with each other. The following lemma verifies these suspicions by showing that Schauder bases are essentially a special case of Schauder decompositions with $\dim \operatorname{ran}(D_k) = 1$.

Lemma 2.7. *Let $\{X_k\}_{k=1}^{\infty}$ be a collection of closed linear subspaces of a Banach space X. If each $x \in X$ has a unique representation $x = \sum_{k=1}^{\infty} x_k$ with $x_k \in X_k$ (i.e., $X = \bigoplus_{k=1}^{\infty} X_k$ is a* ***direct product*** *of the subspaces X_k), then $\{D_k\}_{k=1}^{\infty}$ defined by $D_k x := x_k$ is a Schauder decomposition of X.*

Conversely, if $\{D_k\}_{k=1}^{\infty}$ is a Schauder decomposition, then $X_k := \operatorname{ran}(D_k)$ defines a sequence of subspaces as in the first assertion of the lemma. The correspondence is one-to-one.

Proof. Once the existence parts are shown, the uniqueness is obvious from the statement. The first assertion is the interesting part of the lemma, since the converse statement is almost trivial: For projections D_k, $\operatorname{ran}(D_k)$ is closed, and the series expression for x is of course given by the Schauder decomposition $x = \sum_{k=1}^{\infty} D_k x$. If there is another expression $\sum_{k=1}^{\infty} x_k = x$, then $D_\ell x = D_\ell \sum_{k=1}^{\infty} D_k x_k = \sum_{k=1}^{\infty} D_\ell D_k x_k = D_\ell x_\ell = x_\ell$, i.e., x_ℓ is uniquely determined by x. Above we used the boundedness of D_ℓ to bring the operator inside the sum, together with the projection property $D_k x_k = x_k$ for $x_k \in \operatorname{ran}(D_k)$, and finally the property 1 of Schauder decompositions.

Clearly the D_k defined as above are linear and satisfy the properties 1 and 2 of a Schauder decomposition, so it suffices to show that the D_k are bounded. As in the proof of Lemma 2.6, we do this by first investigating the partial sums. Let us thus define a new norm on X by $\|x\|_X := \sup_{n \in \mathbb{Z}_+} |\sum_{k=1}^{n} x_k|_X$. Since $\lim_{n\to\infty} |\sum_{k=1}^{n} x_k|_X = |x|_X$, we see that for each $x \in X$, $|x|_X \leq \|x\|_X < \infty$. Now the identity mapping $I : (X, \|\cdot\|_X) \to (X, |\cdot|_X)$ is a bounded linear bijection. If we can show that X with the new norm is a Banach space, then it follows from the open mapping theorem that the inverse I^{-1} is also bounded, and thus the two norms are equivalent. Once this is established, we know that $|\sum_{k=1}^{n} D_k x|_X \leq \|x\|_X \leq C\,|x|_X$ and the boundedness of D_n follows by considering the difference of two partial sums as in the proof of Lemma 2.6.

For the completeness, let $\|x^m - x^n\|_X \to 0$ as $n, m \to \infty$. Then

$$\left| \sum_{k=1}^{N} (x_k^m - x_k^n) \right|_X \to 0 \text{ as } n, m \to \infty \tag{2.1}$$

uniformly in N. By considering the difference of (2.1) as such and with $N-1$ in place of N, we deduce that $|x_N^n - x_N^m|_X \to 0$, where $\{x_N^n\}_{n=1}^{\infty} \subset X_N$. As a closed subspace of the Banach space $(X, |\cdot|_X)$, X_N is itself a Banach space, and so the Cauchy sequence converges: $x_N^n \to x_N$, and this holds for each $N \in \mathbb{Z}_+$. Letting n tend to ∞ in (2.1), it follows that $\left|\sum_{k=1}^{N} (x_k^m - x_k)\right|_X \to 0$ as $m \to \infty$ uniformly in N. Since $\sum_{k=1}^{\infty} x_k^m$ converges for each m, it follows that $\sum_{k=1}^{\infty} x_k$ converges to the limit of the Cauchy sequence investigated. □

The result concerning the new norm $\|\cdot\|_X$, which was the essential ingredient of the proof of Lemma 2.7 appears interesting enough to be stated on its own. After doing this we proceed to some corollaries.

Lemma 2.8. *Let $\{D_k\}_{k=1}^\infty$ be a Schauder decomposition of the Banach space X. Then the new norm*

$$\|x\|_X := \sup_{n\in\mathbb{Z}_+} \left|\sum_{k=1}^n D_k x\right|_X$$

is equivalent to $|\cdot|_X$.

Corollary 2.9. *For a Schauder basis $\{e_k\}_{k=1}^\infty$, there is a Schauder decomposition $\{D_k\}_{k=1}^\infty$ with $\operatorname{ran}(D_k) = \operatorname{span}(e_k)$ and $D_k x = \xi_k e_k$ for $x = \sum_{k=1}^\infty \xi_k e_k$. For a Schauder decomposition $\{D_k\}_{k=1}^\infty$ with $\dim \operatorname{ran}(D_k) = 1$ for each k, there exists a Schauder basis $\{e_k\}_{k=1}^\infty$ such that $0 \neq e_k \in \operatorname{ran}(D_k)$. The correspondence is one-to-one, up to scaling of the basis vectors e_k. If the basis is unconditional, so is the decomposition, and vice versa.*

Proof. Since $\operatorname{span}(e_k)$ is a closed subspace, the unconditionality part is the only thing not dealt with in Lemma 2.7. However, this is also immediate, since the definition of unconditionality simply requires the unconditional convergence of the same series in both cases. □

Knowing the relation between Schauder bases and decompositions, we can state an analogue of Lemma 2.6 to give a useful criterion for a sequence to be a basis.

Corollary 2.10. *A sequence $\{e_k\}_{k=1}^\infty \subset X$ of non-zero vectors is a Schauder basis of X if and only if $\operatorname{span}\{e_k\}_{k=1}^\infty$ is dense and there is a constant K such that*

$$\left|\sum_{k=1}^n \xi_k e_k\right|_X \leq K \left|\sum_{k=1}^m \xi_k e_k\right|_X \tag{2.2}$$

for all integers $m \geq n \geq 1$ and all scalars ξ_k, $k = 1, \ldots, m$.

The smallest possible constant K is called the **basis constant** of $\{e_k\}_{k=1}^\infty$. If $K = 1$ (obviously, it cannot be less), then the basis is said to be **monotone**. For instance, each countable orthonormal basis of a Hilbert space is monotone. Another example will be given in Example 3.1.

By definition, $\operatorname{span}\{e_k\}_{k=1}^\infty$ is always dense in its own closure. If $\{e_k\}_{k=1}^\infty$ is a basis of this closure of its linear span, then $\{e_k\}_{k=1}^\infty$ is called a **basic sequence**, independently of whether $\overline{\operatorname{span}}\{e_k\}_{k=1}^\infty = X$ or not. Dropping the density requirement from the assertion of Corollary 2.10, it also gives a characterization of all basic sequences $\{e_k\}_{k=1}^\infty \subset X$.

Proof. If $\{e_k\}_{k=1}^\infty$ is a Schauder basis, let D be the corresponding decomposition. We know from Lemma 2.6 that $\operatorname{span}\{e_k\}_{k=1}^\infty = \operatorname{ran}(D)$ is dense and (2.2) is a direct consequence of the uniform boundedness of the partial sum projectors P_n.

To prove the converse, we first note that it follows from (2.2) that the e_k are linearly independent. Indeed, suppose $F \subset \mathbb{Z}_+$ is finite and $\xi_k \neq 0$ for some $k \in F$. Let $p \in F$ be the smallest number with this property, and set $\xi_k := 0$ for $k \notin F$. Then we have $\left|\sum_{k\in F} \xi_k e_k\right|_X = \left|\sum_{k=1}^{\max F} \xi_k e_k\right|_X \geq K^{-1} \left|\sum_{k=1}^p \xi_k e_k\right|_X = K^{-1} |\xi_k| \, |e_k|_X > 0$.

Thus every $x \in \operatorname{span}\{e_k\}_{k=1}^\infty$ has a unique representation $x = \sum_{k=1}^m \xi_k e_k$ for some $m \in \mathbb{Z}_+$ and scalars ξ_k, $k = 1, \ldots, m$. We then define $D_\ell x := \xi_\ell e_\ell$ if $\ell \leq m$ and 0 otherwise. We want to show that $D := \{D_k\}_{k=1}^\infty$ is a sequence of operators satisfying the conditions of Lemma 2.6. On $\operatorname{span}\{e_k\}_{k=1}^\infty$, $D_k D_\ell = \delta_{k\ell}$ and the sum operators $P_n := \sum_{k=1}^n D_k$ are uniformly bounded by the assumption (2.2). Thus $D_n = P_n - P_{n-1}$ are uniformly bounded as linear operators on $\operatorname{span}\{e_k\}_{k=1}^\infty$ and they can be extended to uniformly bounded linear operators on $\overline{\operatorname{span}}\{e_k\}_{k=1}^\infty = X$. The same conclusion follows for P_n. Lemma 2.6 then shows that D is a Schauder decomposition of X; thus $\{e_k\}_{k=1}^\infty$ is a Schauder basis by Corollary 2.9. □

Example 2.11. *The sequence $\{t^k\}_{k=0}^{\infty}$ is linearly independent and dense on $C[0,1]$ with supremum norm, but not a basis.*

Thus the linear independence of a dense sequence does not guarantee that it is a basis. We saw in the proof of Corollary 2.10 that the condition (2.2) implies linear independence of $\{e_k\}_{k=1}^{\infty}$. By the same Corollary, this example shows that the converse is not true.

Proof. The linear independence of $\{t^k\}_{k=0}^{\infty}$ is obvious, since a polynomial of non-zero coefficients does not vanish identically. The density follows from Weierstrass' approximation theorem. However, only a real analytic (in particular, infinitely differentiable) function f has a convergent power series representation $f(t) = \sum_{k=0}^{\infty} a_k t^k$, whereas continuous non-differentiable functions on $[0,1]$ are abundant. $\square$

Next we describe some alternative characterizations of unconditional Schauder decompositions (thus by Corollary 2.9 also of unconditional bases). Lemma 2.12 is closely related to Lemma 2.2 above. For compactness of notion we introduce, for each $\lambda \in \ell^{\infty}(\mathbb{Z}_+)$, the operators

$$T_{\lambda} \sum_{k=1}^{\infty} D_k x := \sum_{k=1}^{\infty} \lambda_k D_k x. \tag{2.3}$$

If $D = \{D_k\}_{k=1}^{\infty}$ is an unconditional Schauder decomposition, this is well-defined for every $x \in X$ by Lemma 2.2(4), i.e., the series on the right is convergent. For any Schauder decomposition (not necessarily unconditional), the definition is valid for $x \in \operatorname{ran} D$. The operator T_{λ} in (2.3) is, in fact, the first example of abstract multiplier operators, many more of which will be encountered in the sequel.

As a corollary of Lemma 2.12 below, we obtain the first abstract multiplier theorem.

Lemma 2.12. *Let $D = \{D_k\}_{k=1}^{\infty}$ be a Schauder decomposition of the Banach space X. Then the following statements are equivalent:*

1. *D is unconditional.*
2. *The family $\{T_{\delta}\}_{\delta \in \{0,1\}^{\mathbb{Z}_+}}$ of operators on $\operatorname{ran} D$ is uniformly bounded.*
3. *The family $\{T_{\lambda}\}_{\lambda \in \ell^{\infty}(\mathbb{Z}_+)}$ of operators on $\operatorname{ran} D$ is uniformly bounded.*
4. *The family $\{T_{\epsilon}\}_{\epsilon \in \{-1,1\}^{\mathbb{Z}_+}}$ of operators on $\operatorname{ran} D$ is uniformly bounded.*

Proof. $1 \Rightarrow 2$. Now that unconditional convergence is assumed, the operators T_{δ} are well-defined on all of X. The uniform boundedness is established by repeated application of the uniform boundedness principle, the first goal being the boundedness of each T_{δ} itself. We first consider the partial sum operators T_{δ}^n defined as the nth partial sum of (2.3). Then clearly each T_{δ}^n is bounded as a finite sum of bounded operators D_k. Furthermore, $T_{\delta}^n x \to T_{\delta} x$ for each $x \in X$; thus $|T_{\delta}^n x|_X$ is bounded for every $x \in X$, and the principle of uniform boundedness implies the boundedness in n of $|T_{\delta}^n|_{\mathcal{B}(X)}$. Then the strong limit T_{δ} is also bounded.

Now from the fact that $\sum_{k=1}^{\infty} x_k$ is shrinking (Lemma 2.2(3)) it follows that for sufficiently large N, the finite set F_0 corresponding to a given $\epsilon > 0$ satisfies $F_0 \subset \{1, \ldots, N\}$, and thus for any finite $F \subset \mathbb{Z}_+ + N$ we have $\left|\sum_{k \in F} D_k x\right|_X < \epsilon$, thus $\left|\sum_{k=N+1}^{M} \delta_k D_k x\right|_X < \epsilon$ for any $M > N$, and eventually $\left|\sum_{k=N+1}^{\infty} \delta_k D_k x\right|_X \leq \epsilon$. A combination of assumption 1 and Lemma 2.2(6) yields that this series converges. Now each T_{δ} is a sum of two operators $T_{\delta}^n + S_{\delta}^n$, where T_{δ}^n is the nth partial sum as above and S_{δ}^n is the rest. We just saw that the $S_{\delta}^n x$ are all bounded in δ for a fixed x and a large enough n. Fix such an n and observe that there are only finitely many different operators T_{δ}^n (indeed, 2^n, one for each $\delta \in \{0,1\}^n$). It then follows that $|T_{\delta} x|_X$ is uniformly bounded in δ for a fixed x, and one more application of the uniform boundedness principle shows that the operator norm $|T_{\delta}|_{\mathcal{B}(X)}$ is bounded in δ.

$2 \Rightarrow 3$. By using the decomposition $\lambda_k = |\lambda|_{\ell^\infty} (\lambda_k^+ - \lambda_k^- + \mathbf{i}\mu_k^+ - \mathbf{i}\mu_k^-)$ if necessary, we may assume $\lambda_k \in [0,1]$. For $\operatorname{ran}(D) \ni x = \sum_{k=1}^n D_k x_k$ we apply Lemma 2.3 to deduce $|T_\lambda x|_X = |\sum_{k=1}^n \lambda_k D_k x|_X \le \max_{\delta\in\{0,1\}^n} |\sum_{k=1}^n \delta_k D_k x|_X = |T_\delta x|_X \le C\,|x|_X$, where the last inequality follows, of course, from the assumption 2.

$3 \Rightarrow 4$. This is trivial.

$4 \Rightarrow 1$. Now $|\sum_{k=n}^m \epsilon_k D_k x|_X \le C\,|\sum_{k=n}^m D_k x|_X \to 0$ as $n, m \to \infty$, since $\sum_{k=1}^\infty D_k x$ converges. Thus the sequence of partial sums of $\sum_{k=n}^m \epsilon_k D_k x$ is Cauchy and hence convergent for any $\epsilon \in \{-1,1\}^{\mathbb{Z}_+}$. By Lemma 2.2(5), this is equivalent to the unconditional convergence of $\sum_{k=n}^m \epsilon_k D_k x$. □

Remark 2.13. *1. An equivalent statement to the conditions 2 to 4 of Lemma 2.12 is to say that there is a $C > 0$ such that*

$$\left|\sum_{k=1}^n \epsilon_k x_k\right|_X \le C\left|\sum_{k=1}^n x_k\right|_X \tag{2.4}$$

for $x_k \in \operatorname{ran}(D_k)$, $k = 1,\dots,n$, for all $n \in \mathbb{Z}_+$ and $\epsilon \in \{-1,1\}^{\mathbb{Z}_+}$, and similarly with δ and λ replacing ϵ. The smallest C for which (2.4) is valid is called the ***unconditional constant*** *of D and denoted by C_D.*

2. *Since by Lemma 2.12 the validity of (2.4) implies unconditionality of the Schauder decomposition D, the series involved in (2.4) with $x_k := D_k x$ are convergent for any $x \in X$. It is legitimate to pass to the limit $n \to \infty$ to deduce the same inequality for infinite series.*

3. *By substituting $\epsilon_k x_k$ in place of x_k in (2.4), it follows that the inequality in fact is two-sided:*

$$C^{-1}\left|\sum_{k=1}^n x_k\right|_X \le \left|\sum_{k=1}^n \epsilon_k x_k\right|_X \le C\left|\sum_{k=1}^n x_k\right|_X, \quad \epsilon_k = \pm 1, \quad x_k \in \operatorname{ran}(D_k). \tag{2.5}$$

Now we state the first abstract multiplier theorem. This simple result will be exploited in deriving more powerful theorems in the sequel.

Corollary 2.14. *If $D = \{D_k\}_{k=1}^\infty$ is an unconditional Schauder decomposition of the Banach space X, then the operator T_λ defined by (2.3) is bounded, i.e., $T_\lambda \in \mathcal{B}(X)$, if and only if $\lambda \in \ell^\infty$.*

Proof. The "if" part is already included in Lemma 2.12(3) and Remark 2.13(2). To show the "only if" part, take $\lambda \notin \ell^\infty$ and pick a sequence $\{k_j\}_{j=1}^\infty \subset \mathbb{Z}_+$ such that $|\lambda_{k_j}| \to \infty$. Finally take $x_{k_j} \in \operatorname{ran}(D_{k_j})$ normalized to unity norm and conclude that $|T_\lambda x_{k_j}|_X = |\lambda_{k_j} x_{k_j}|_X = |\lambda_{k_j}| \to \infty$, thus T_λ is not bounded. □

We conclude this section with a simple construction to yield new Schauder decompositions of X from existing ones.

Definition 2.15. *Let $D = \{D_k\}_{k=1}^\infty$ be a Schauder decomposition of the Banach space X, and $\{n_k\}_{k=1}^\infty \subset \mathbb{Z}_+$ be a strictly increasing sequence. Then the sequence $D' := \{D'_k\}_{k=1}^\infty$ of operators defined by $D'_k := \sum_{\ell=n_{k-1}+1}^{n_k} D_\ell$ is called a* ***blocking*** *of D.*

It is easy to verify that a blocking of a Schauder decomposition is again a Schauder decomposition. Furthermore, the blocking is unconditional if the original decomposition is.

2.4 Notes and comments

This chapter mostly follows the presentation in Hieber and Prüss [8] and in Witvliet [28]. A number of results have also been taken from James [11].

Bases of Banach spaces have been studied for quite a long time. James' paper contains more classical theorems concerning this subject than we have cited here. The notion of unconditional convergence, according to Hille and Phillips [9], dates back to the work of Orlicz in 1933.

Chapter 3

Randomized Norms

3.1 Introduction

We now introduce some probabilistic aspects to the analysis of Schauder decompositions. These pave the way for the extremely important concept of R-boundedness that will be used later on. Recall that R stands for "randomized". Furthermore, the randomization concepts help us in describing some relations between a Schauder decomposition $D = \{D_k\}_{k=1}^{\infty}$ and the collection of dual operators $D^* := \{D_k^*\}_{k=1}^{\infty}$. This gives us an abstract framework for powerful duality arguments, which are often fruitful in applications, since many of the most common operators of analysis have either a self-adjoint or a skew-adjoint nature.

Of particular interest will be independent, identically distributed, symmetric $\{-1, 1\}$-valued random variables on a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, i.e., a measure space (any set) Ω with a σ-algebra $\mathfrak{F} \subset 2^{\Omega}$ and a probability measure $\mathbb{P} : \mathfrak{F} \to [0, 1]$. These will be denoted by ε_k, $k \in \mathbb{Z}_+$, and referred to as the **Rademacher functions**. Such random variables are of course abundant, a standard example being $r_k(t) := \operatorname{sgn} \sin(2^k \pi t)$, $k \in \mathbb{Z}_+$, on $\Omega = [0, 1]$ with $\mathbb{P} = m$, the Lebesgue measure and $\mathfrak{F} = \mathfrak{M}[0, 1]$, the corresponding σ-algebra of measurable subsets of $[0, 1]$. (Sometimes, the name of Rademacher is only used for the particular functions r_k. We will use the name for all independent, identically distributed, symmetric $\{-1, 1\}$-valued random variables; usually, we are only concerned about the joint distribution and not the "internal structure" of the functions.)

For a sequence $x_1, \ldots, x_n \in X$, the randomized norm $|\sum_{k=1}^{n} \varepsilon_k x_k|_{L^p(\Omega;X)}$ will be of interest. Here $L^p(\Omega; X)$ is the space of all $\mathfrak{F}$-measurable functions $f : \Omega \to X$ whose pointwise norm is integrable in the pth power. A more detailed account of vector valued integration is given in the Appendix, Section A.2; in the present context, however, this is hardly needed: Observe that the randomized norms $|\sum_{k=1}^{n} \varepsilon_k x_k|_{L^p(\Omega;X)}$ only involve simple functions (and in fact norms of such functions), so only a naïve idea of integration is required; indeed

$$\left|\sum_{k=1}^{n} \varepsilon_k x_k\right|_{L^p(\Omega;X)} := \left(\int_{\Omega} \left|\sum_{k=1}^{n} \varepsilon_k(\omega) x_k\right|_X^p d\mathbb{P}(\omega)\right)^{\frac{1}{p}} = \left(\frac{1}{2^n} \sum_{\eta \in \{-1,1\}^n} \left|\sum_{k=1}^{n} \eta_k x_k\right|_X^p\right)^{\frac{1}{p}}. \tag{3.1}$$

Substituting $\epsilon_k x_k$ in place of x_k in this equation, $\epsilon \in \{-1, 1\}^n$, and observing that the summation on the right is thereafter over exactly the same values (possibly in a different order), we find that the randomized norm remains invariant under such transformations; explicitly

$$\left|\sum_{k=1}^{n} \epsilon_k \varepsilon_k x_k\right|_{L^p(\Omega;X)} = \left|\sum_{k=1}^{n} \varepsilon_k x_k\right|_{L^p(\Omega;X)}.$$

In probabilistic language, this illustrates the fact that $\{\epsilon_k \varepsilon_k\}_{k=1}^{n}$ and $\{\varepsilon_k\}_{k=1}^{n}$ have identical joint distributions. (The argument above applies to any function of the ε_k and not just to the randomized norms considered.)

A simple property of the Rademacher functions, which relates them to the context of Schauder bases, is the following:

Example 3.1. *For $x_k \in X \setminus \{0\}$ and ε_k Rademacher functions on Ω, $\{\varepsilon_k x_k\}_{k=1}^{\infty}$ is a monotone basic sequence in $L^p(\Omega; X)$, $p \in [1, \infty)$.*

Proof. Applying the triangle inequality, we find that

$$\left|\sum_{k=1}^{n} \varepsilon_k x_k\right|_{L^p(\Omega;X)} \leq \frac{1}{2}\left|\sum_{k=1}^{n} \varepsilon_k x_k + x_{n+1}\right|_{L^p(\Omega;X)} + \frac{1}{2}\left|\sum_{k=1}^{n} \varepsilon_k x_k - x_{n+1}\right|_{L^p(\Omega;X)} = \left|\sum_{k=1}^{n+1} \varepsilon_k x_k\right|_{L^p(\Omega;X)},$$

where the last equality follows from the fact that ε_{n+1} attains both of the values 1 and -1 with probability $\frac{1}{2}$, independently of the other ε_k. The assertion, i.e., the inequality (2.2) with $e_k = \varepsilon_k x_k$ and $K = 1$, follows from the previous computation by iteration, and observing that the same computation is valid with x_k replaced by $\xi_k x_k$. ☐

3.2 Randomization and duality of decompositions

We now give the first result, an easy consequence of Lemma 2.12 and Remark 2.13, concerning the relation of randomization and unconditionality.

Lemma 3.2. *For a Schauder decomposition $D = \{D_k\}_{k=1}^{\infty}$ of a Banach space X, the following conditions are equivalent:*

1. *D is unconditional.*
2. *For every $p \in [1, \infty)$ there exists a $C_p > 0$ such that*

$$C_p^{-1}\left|\sum_{k=1}^{n} x_k\right|_X \leq \left|\sum_{k=1}^{n} \varepsilon_k x_k\right|_{L^p(\Omega;X)} \leq C_p\left|\sum_{k=1}^{n} x_k\right|_X \tag{3.2}$$

   for all $x_k \in \operatorname{ran}(D_k)$, $k \in \mathbb{Z}_+$ and all $n \in \mathbb{Z}_+$.

3. *There exists one $p \in [1, \infty)$ and some $C_p > 0$ such that (3.2) holds.*

Proof. 1 $\Rightarrow$ 2 follows by observing that, for a fixed $\omega \in \Omega$, (2.5) is valid with $\epsilon_k := \varepsilon_k(\omega)$ (by Lemma 2.12 and Remark 2.13) and integrating the pth power of this inequality over Ω with respect to $\mathbb{P}$, recalling that $\mathbb{P}(\Omega) = 1$ for a probability measure.

2 $\Rightarrow$ 3 is obvious.

3 $\Rightarrow$ 1. Here we use the fact that, for a sequence $\{\epsilon_k\}_{k=1}^{\infty} \subset \{-1, 1\}^{\mathbb{Z}_+}$, the joint distributions of $\varepsilon_1, \ldots, \varepsilon_n$ and $\epsilon_1\varepsilon_1, \ldots, \epsilon_n\varepsilon_n$ are identical. Thus, writing the inequality (3.2) with $\epsilon_k x_k$ in place of x_k, it follows that

$$\left|\sum_{k=1}^{n} \epsilon_k x_k\right|_X \leq C_p\left|\sum_{k=1}^{n} \epsilon_k\varepsilon_k x_k\right|_{L^p(\Omega;X)} = C_p\left|\sum_{k=1}^{n} \varepsilon_k x_k\right|_{L^p(\Omega;X)} \leq C_p^2\left|\sum_{k=1}^{n} x_k\right|_X.$$

Thus the operators T_ϵ on $\operatorname{ran}(D)$ (defined by (2.3)) are uniformly bounded by $|T_\epsilon|_{\mathcal{B}(X)} \leq C_p^2$ for $\epsilon \in \{-1, 1\}^{\mathbb{Z}_+}$, and the claim follows from Lemma 2.12(4), ☐

In Lemma 3.2 we saw that the value of the exponent p in randomization was quite irrelevant. This is an example of a more general phenomenon related to the randomized norms $|\sum_{k=1}^{n} \varepsilon_k x_k|_{L^p(\Omega;X)}$. In fact, the inequality of Khintchine and Kahane states that

$$\left|\sum_{k=1}^{n} \varepsilon_k x_k\right|_{L^q(\Omega;X)} \leq K_{q,p}\left|\sum_{k=1}^{n} \varepsilon_k x_k\right|_{L^p(\Omega;X)} \tag{3.3}$$

for some constant $K_{q,p}$ depending only on $p, q \in (0,\infty)$, in any linear normed space X. In the present work we are mainly concerned about exponents p in the range $[1,\infty)$. The inequality is trivial for $q \leq p$, with $K_{q,p} = 1$, since the L^p norm on a probability space is an increasing function of p by Jensen's inequality. We will prove the Khintchine–Kahane inequality for $q > p$ in Section 3.3.

Motivated by Lemma 3.2 and the Khintchine–Kahane inequality, we give the following definition akin to the unconditionality of a Schauder decomposition, but concerning a general sequence of bounded linear operators.

Definition 3.3. *A sequence $\{R_k\}_{k=1}^\infty \subset \mathcal{B}(X)$, X Banach, is called a* ***random unconditional*** *if for some $p \in [1,\infty)$ there exists a $C_p > 0$ such that*

$$\left|\sum_{k=1}^n \varepsilon_k R_k x\right|_{L^p(\Omega;X)} \leq C_p \left|\sum_{k=1}^n R_k x\right|_X \tag{3.4}$$

for all $x \in X$ and $n \in \mathbb{Z}_+$.

Remark 3.4. *1. It follows immediately from the Khintchine–Kahane inequality that the condition (3.4) of random unconditionality holds for all $p \in [1,\infty)$, if it holds for one.*

2. From Lemma 3.2(2) it follows that every unconditional Schauder decomposition of X is a random unconditional on X. However, it suffices that D is merely an unconditional Schauder decomposition of $\overline{\operatorname{ran}}(D)$; indeed, then (3.4) *holds for all $x \in \operatorname{ran}(D_k)$, in particular for $x' := \sum_{\ell=1}^n D_k x$, where $x \in X$ is arbitrary, and thus, after simplification, also for x.*

We then give the first result concerning the dual D^* of a Schauder decomposition.

Lemma 3.5. *Let $D = \{D_k\}_{k=1}^\infty$ be an unconditional Schauder decomposition of the Banach space X. Then the dual $D^* := \{D_k^*\}_{k=1}^\infty$ is an unconditional Schauder decomposition of $\overline{\operatorname{ran}}(D^*)$ with the same unconditional constant C_D. The closure here means closure in norm.*

In particular, D^* is a random unconditional on X^*.

Proof. From Lemma 2.12 and Remark 2.13 it follows that the operators $T_\epsilon \in \mathcal{B}(X)$ are uniformly bounded by C_D. Then

$$\begin{aligned}\langle T_\epsilon^* x^*, x\rangle = \langle x^*, M_\epsilon x\rangle = \left\langle x^*, \sum_{k=1}^\infty \epsilon_k D_k x\right\rangle &= \lim_{n\to\infty}\left\langle x^*, \sum_{k=1}^n \epsilon_k D_k x\right\rangle \\ &= \lim_{n\to\infty}\left\langle \sum_{k=1}^n \epsilon_k D_k^* x^*, x\right\rangle =: \left\langle w^*\text{-}\sum_{k=1}^\infty \epsilon_k D_k^* x^*, x\right\rangle,\end{aligned}$$

where the last equality simply defines the weak* series. Thus $T_\epsilon^* = w^*\text{-}\sum_{k=1}^\infty \epsilon_k D_k^*$. Now $D_k^* \in \mathcal{B}(X^*)$ and $\langle D_k^* D_\ell^* x^*, x\rangle = \langle x^*, D_\ell D_k x\rangle = 0$ for $k \neq \ell$ for all $x \in X$; thus $D_k^* D_\ell^* = 0$, $k \neq \ell$. It is then clear that D^* is a Schauder decomposition of $\overline{\operatorname{ran}}(D^*)$. Furthermore we have the simple estimate $|\langle T_\epsilon^* x^*, x\rangle| \leq |x^*|_{X^*} |T_\epsilon x|_X \leq C_D |x^*|_{X^*} |x|_X$; thus $|T_\epsilon^*|_{\mathcal{B}X^*} \leq C_D$ uniformly in ϵ. The unconditionality of the decomposition then follows from Lemma 2.12(4), and the fact that D^* is a random unconditional from Remark 3.4(2). □

Both the assertion of the previous Lemma and its proof were quite trivial, consisting only of some rather obvious identities. If X is reflexive, however, it follows, as will be shown in Corollary 3.6 below, that D^* is a Schauder decomposition of all of X^*. This conclusion is not generally valid if the reflexivity is given up, as the following example demonstrates: Let $X = \ell^1$ and D the Schauder decomposition corresponding to the standard (Schauder) basis $\{e_k\}_{k=1}^\infty$, where the kth coordinate of e_k is 1, all others being 0. The unconditionality of this basis can be demonstrated in various ways, for instance by Lemma 2.12(3) using the fact that $(\ell^1)^*$ can be identified with ℓ^∞ in the obvious way. However, ℓ^∞ is not separable and thus cannot have any countable basis.

Corollary 3.6. *If X is a reflexive Banach space and D is an unconditional Schauder decomposition of X, then D^* is an unconditional Schauder decomposition of X^*.*

Proof. By Lemma 3.5 it suffices to show that the closed subspace $\overline{\operatorname{ran}}(D^*)$ is all of X^*. If this is not the case, then the Hahn–Banach theorem implies the existence of a non-zero $x^{**} \in X^{**}$, which vanishes on all of $\overline{\operatorname{ran}}(D^*)$. Since X is reflexive, $0 = \langle x^{**}, x^* \rangle_{X^*} = \langle x^*, x \rangle_X$ for some $x = X$ and all $x^* \in \overline{\operatorname{ran}}(D^*)$. For arbitrary $x^* \in X^*$, $D_\ell^* x^* \in \operatorname{ran}(D^*)$, and substituting this in place of x^* above we have $0 = \langle x^*, D_\ell x \rangle$ for all $x^* \in X^*$. It follows that $D_\ell x = 0$ for each ℓ; thus $x = \sum_{\ell=1}^\infty D_\ell x_\ell = 0$, but this is a contradiction, since we assumed x^{**} (which was identified with x in the obvious way) to be non-zero. □

Lemma 3.7. *A Schauder decomposition D of a Banach space X is unconditional if and only if both D and D^* are random unconditional.*

Proof. The "only if" part follows directly from Lemma 3.5 and the fact that any unconditional Schauder decomposition is a random unconditional. For the converse, our intention is to apply Remark 2.13(1), but to do so, we must make some preliminary computations. We first observe that

$$\begin{aligned}\left|\left\langle x^*, \sum_{k=1}^n D_k x \right\rangle\right| &= \left|\left\langle x^*, \sum_{k=1}^n \varepsilon_k(\omega) D_k \sum_{\ell=1}^n \varepsilon_\ell(\omega) D_\ell x \right\rangle\right| \\ &= \left|\left\langle \sum_{k=1}^n \varepsilon_k(\omega) D_k^* x^*, \sum_{\ell=1}^n \varepsilon_\ell(\omega) D_\ell x \right\rangle\right| \le \left|\sum_{k=1}^n \varepsilon_k(\omega) D_k^* x^*\right|_{X^*} \left|\sum_{\ell=1}^n \varepsilon_\ell(\omega) D_\ell x\right|_X\end{aligned}$$

We integrate over Ω with respect to $\mathbb{P}$ and invoke the Cauchy–Schwarz–Bunyakovsky inequality to obtain

$$\begin{aligned}\left|\left\langle x^*, \sum_{k=1}^n D_k x \right\rangle\right| &\le \left|\sum_{k=1}^n \varepsilon_k D_k^* x^*\right|_{L^2(\Omega;X^*)} \left|\sum_{\ell=1}^n \varepsilon_\ell D_\ell x\right|_{L^2(\Omega;X)} \\ &\le C_2 \left|\sum_{k=1}^n D_k^* x^*\right|_{X^*} \left|\sum_{\ell=1}^n \varepsilon_\ell D_\ell x\right|_{L^2(\Omega;X)},\end{aligned}$$

where we used in the second inequality the fact that D^* is a random unconditional. Then exploit the assumption that D is a Schauder decomposition via Lemma 2.8 to deduce

$$\left|\left\langle \sum_{k=1}^n D_k^* x^*, x \right\rangle\right| = \left|\left\langle x^*, \sum_{K=1}^n D_k x \right\rangle\right| \le |x^*|_{X^*} \left|\sum_{k=1}^n D_k x\right|_X \le |x^*|_{X^*} \|x\|_X \le K\, |x^*|_{X^*} \, |x|_X\,,$$

for some $K > 0$. The notation of Lemma 2.8 was used in the second to last step, recall that $\|x\|_X := \sup_{n \in \mathbb{Z}_+} |\sum_{k=1}^n D_k x_k|_X$. The last inequality is the essential content of that Lemma. This last computation shows that $|\sum_{k=1}^n D_k^* x^*|_{X^*} \le K\, |x^*|_{X^*}$ and we may use this in the earlier inequality to finally obtain, after choosing x^* of unity norm so as to make the pairing of $\sum_{k=1}^n D_k x$ and x^* equal to the norm of the first quantity,

$$\left|\sum_{k=1}^n D_k x\right|_X \le K C_2 \left|\sum_{\ell=1}^n \varepsilon_\ell D_k x\right|_{L^2(\Omega;X)}.$$

Using this last result, it is straightforward to verify the condition of Remark 2.13(1):

$$\left|\sum_{k=1}^n \epsilon_k D_k x\right|_X \le K C_2 \left|\sum_{k=1}^n \epsilon_k \varepsilon_k D_k x\right|_{L^2(\Omega;X)} = K C_2 \left|\sum_{k=1}^n \varepsilon_k D_k x\right|_{L^2(\Omega;X)} \le K C_2^2 \left|\sum_{k=1}^n D_k x\right|_X.$$

In the last two steps we used the fact that $\epsilon_k \varepsilon_k$ and ε_k have the same joint distribution, and the random unconditionality of D. □

3.3 Equivalence of randomized norms

We already saw in Lemma 3.2 an example of the phenomenon, where the randomization by Rademacher functions makes different L^p norms equivalent. We also indicated the wider generality of this effect and formulated the Khintchine–Kahane inequality, which was applied to deduce the independence of the property of random unconditionality (Definition 3.3) on the exponent p in the definition. There will be other, more significant consequences later on. This section is devoted to justifying this important inequality. In fact, we prove a more general result concerning norms of **Rademacher chaoses**, by which we mean expressions of the form

$$\sum_{k=0}^{n}\lambda^k\sum_{\#I=k}x_I\prod_{i\in I}\varepsilon_i$$
$$=x_\emptyset+\lambda\sum_{i=1}^{n}x_i\varepsilon_i+\ldots+\lambda^r\sum_{1\leq i_1<\ldots<i_r\leq n}x_{i_1,\ldots,i_r}\varepsilon_{i_1}\cdots\varepsilon_{i_r}+\ldots+\lambda^n x_{1,\ldots,n}\varepsilon_1\cdots\varepsilon_n,$$

where $\sum_{\#I=k}$ indicates summation over all $I\subset\{1,\ldots,n\}$ having exactly k elements, and ε_i are Rademacher functions. The x_I (note that they are indexed by sets I) are vectors of a linear normed space X. (One could guess from the form of the Khintchine–Kahane inequality (3.3) that no completeness is involved. Indeed, the inequality is of essentially algebraic (as opposed to analytic) nature as the integrals reduce to the summations (3.1).)

We start with a lemma concerning real variables only; nevertheless, it already contains the essence of the matter.

Lemma 3.8. *For $x,y\in\mathbb{R}$, $1<p\leq q<\infty$ and ε a Rademacher function, we have*

$$\left|x+\sqrt{\frac{p-1}{q-1}}\varepsilon y\right|_{L^q(\Omega)}\leq|x+\varepsilon y|_{L^p(\Omega)}.$$

Proof. We start with some simple reductions. Since $\lambda:=\sqrt{\frac{p-1}{q-1}}\leq 1$, it is clear that the inequality holds if $x=0$ or $y=0$. Otherwise, dividing both sides by x, the inequality reduces to the case $x=1$; i.e., we need to prove

$$\left(\frac{1}{2}\left(|1+\lambda y|^q+|1-\lambda y|^q\right)\right)^{\frac{1}{q}}\leq\left(\frac{1}{2}\left(|1+y|^p+|1-y|^p\right)\right)^{\frac{1}{p}}. \tag{3.5}$$

Consider first the case $1<p\leq q\leq 2$ and $|y|<1$; symmetry of y and $-y$ in the inequality allows us to assume $0<y<1$ without further loss of generality. Now $|\pm\lambda y|<1$, and thus the left-hand side of (3.5) attains the form (omitting the exponent $\frac{1}{q}$)

$$\frac{1}{2}\left((1+\lambda y)^q+(1-\lambda y)^q\right)=\frac{1}{2}\left(\sum_{j=0}^{\infty}\binom{q}{j}\lambda^j y^j+\sum_{j=0}^{\infty}\binom{q}{j}(-\lambda)^j y^j\right)=1+\sum_{k=1}^{\infty}\binom{q}{2k}\lambda^{2k}y^{2k}$$

with an absolutely convergent power series.

An estimate concerning the binomial coefficients in the previous expression is now in order. We claim that $\binom{q}{2k}\frac{p-1}{q-1}\leq\frac{q}{p}\binom{p}{2k}$. For $k=1$, this says that $1\leq 1$. To make an induction it suffices to show that $(q-2k)(q-2k-1)\leq(p-2k)(p-2k-1)$ for $k\geq 1$; the left and right-hand sides of the next member in the sequence of asserted inequalities are obtained from the previous one after multiplying by these quantities (and the common factor $1/(2k+1)(2k+2)$). This last inequality is true, since the derivative $f'(x)=2x-(4k+1)$ of $f(x):=(x-2k)(x-2k-1)$ is negative for $x\leq 2<2k+\frac{1}{2}$; thus $f(q)\leq f(p)$ for $p\leq q\leq 2$.

Using this estimate, it now follows readily that

$$\left(1+\sum_{k=1}^{\infty}\binom{q}{2k}\frac{p-1}{q-1}\lambda^{2(k-1)}y^{2k}\right)^{\frac{p}{q}}\leq\left(1+\sum_{k=1}^{\infty}\frac{q}{p}\binom{p}{2k}y^{2k}\right)^{\frac{p}{q}}\leq 1+\sum_{k=1}^{\infty}\binom{p}{2k}y^{2k},$$

the last step being the easy inequality $(1+x)^r \leq 1+rx$ for $x \geq 0$ and $r = \frac{p}{q} \in (0,1]$. The first step should be clear by the estimate for the binomial coefficient after recalling that $\lambda^2 = (p-1)/(q-1) \in (0,1]$.

Thus, for $1 < p \leq q \leq 2$, we have proved (3.5) for $|y| < 1$, and the case $y = \pm 1$ follows by continuity. If $|y| > 1$, then

$$0 \geq (1-y^2)(1-\lambda^2) = (1+y^2\lambda^2 \pm 2y\lambda) - (\lambda^2 + y^2 \pm 2y\lambda) = |1 \pm y\lambda|^2 - |y \pm \lambda|^2\,;$$

thus $|1 \pm y\lambda| \leq |y|\,|1 \pm \lambda/y|$ and

$$\begin{aligned}\left(\frac{1}{2}(|1+\lambda y|^q + |1-\lambda y|^q)\right)^{\frac{1}{q}} &\leq |y| \left(\frac{1}{2}(|1+\lambda/y|^q + |1-1/y|^q)\right)^{\frac{1}{q}} \\ &\leq |y| \left(\frac{1}{2}(|1+1/y|^p + |1-\lambda/y|^p)\right)^{\frac{1}{p}} = \left(\frac{1}{2}(|y+1|^p + |y-1|^p)\right)^{\frac{1}{p}}.\end{aligned}$$

We used the inequality (3.5) for the case $|y| < 1$ for $1/y$ in the second to last step.

Now (3.5) is proved completely in the case $1 < p \leq q \leq 2$. The case $2 \leq p \leq q < \infty$ will be established by a duality argument. Once this is done, the assertion is completely proved; indeed, the case $1 < p < 2 < q < \infty$ then follows by joining the inequalities for $(p, 2)$ and $(2, q)$.

For the duality argument, observe that the inequality of the lemma asserts the contractivity of the (linear) operator $T : L^q(\{-1,1\}) \to L^p(\{-1,1\})$, which maps a function taking the values $x+y$ and $x-y$ into one taking the values $x+\lambda y$ and $x - \lambda y$. (We consider the set $\{-1,1\}$ with the symmetric probability measure ς, $\varsigma\{\pm 1\} = \frac{1}{2}$.) One can easily verify that the operator T described can be represented as

$$Tf(\eta) = \int_\Omega f(\epsilon)d\varsigma(\epsilon) + \lambda \int_\Omega \epsilon f(\epsilon) d\varsigma(\epsilon) \cdot \eta,$$

where $\Omega := \{-1,1\}$. Furthermore, for $f \in L^q$ (thus $Tf \in L^p$) and $g \in L^{\overline{p}}$ ($\overline{p}$ is the conjugate exponent of p, $\frac{1}{p} + \frac{1}{\overline{p}} = 1$),

$$\begin{aligned}\langle Tf, g\rangle := \int Tf(\eta)g(\eta)d\varsigma(\eta) &= \int \left(\int f(\epsilon)d\varsigma(\epsilon) + \lambda \int f(\epsilon)\epsilon d\varsigma(\epsilon)\cdot \eta\right) g(\eta)d\varsigma(\eta) \\ &= \int \left(\int g(\eta)d\varsigma(\eta) + \lambda \int g(\eta)\eta d\varsigma(\eta) \cdot \epsilon\right) f(\epsilon)d\varsigma(\epsilon) = \langle Tg, f\rangle\,,\end{aligned}$$

i.e., $T^* = T$ (in the sense that the two operators have the same formula given above). The contractivity of T implies contractivity of T^*, and thus T^* is a contraction from $L^{\overline{p}}(\{-1,1\})$ to $L^{\overline{q}}(\{-1,1\})$, where $2 \leq \overline{q} \leq \overline{p} < \infty$ when $1 < p \leq q \leq 2$ (and the conjugate exponent mapping is a bijection between the two ranges). Thus we have proved the assertion for the remaining exponents, and the proof is complete. ☐

Now we do the same in an arbitrary linear normed space X.

Lemma 3.9. *For $x, y \in X$, X a linear normed space, $1 < p \leq q < \infty$ and ε a Rademacher function, we have*

$$\left| x + \sqrt{\frac{p-1}{q-1}}\varepsilon y \right|_{L^q(\Omega;X)} \leq |x + \varepsilon y|_{L^p(\Omega;X)}\,.$$

Proof. We need only minor manipulations to reduce the new assertion to the real valued case already proved; we again denote $\lambda := \sqrt{\frac{p-1}{q-1}}$, and also $u := x + y$, $v := x - y$:

$$\begin{aligned}
&\left(\frac{1}{2}(|x+\lambda y|_X^q + |x-\lambda y|_X^q)\right)^{\frac{1}{q}} \\
&\quad = \left(\frac{1}{2}\left\{\left|\frac{1+\lambda}{2}(x+y) + \frac{1-\lambda}{2}(x-y)\right|_X^q + \left|\frac{1-\lambda}{2}(x+y) + \frac{1+\lambda}{2}(x-y)\right|_X^q\right\}\right)^{\frac{1}{q}} \\
&\quad \le \left(\frac{1}{2}\left\{\left(\frac{1}{2}(|u|_X + |v|_X) + \frac{\lambda}{2}(|u|_X - |v|_X)\right)^q + \left(\frac{1}{2}(|u|_X + |v|_X) - \frac{\lambda}{2}(|u|_X - |v|_X)\right)^q\right\}\right)^{\frac{1}{q}} \\
&\quad \le \left(\frac{1}{2}\left\{\left(\frac{1}{2}(|u|_X + |v|_X) + \frac{1}{2}(|u|_X - |v|_X)\right)^p + \left(\frac{1}{2}(|u|_X + |v|_X) - \frac{1}{2}(|u|_X - |v|_X)\right)^p\right\}\right)^{\frac{1}{p}} \\
&\quad = \left(\frac{1}{2}(|x+y|_X^p + |x-y|_X^p)\right)^{\frac{1}{p}}.
\end{aligned}$$

The first inequality was simply the triangle inequality, the second being the real valued version of the present lemma in Lemma 3.8. □

Next comes the deep result. The proof will essentially be based on the idea of breaking the expectation of a function of independent random variables into expectations with respect to a smaller number of variables at a time. This fundamental idea of conditional expectation will be treated more thoroughly in Section 5.2. Here we only deal with simple functions, so a naïve idea is sufficient: If $\varepsilon := (\varepsilon_i)_{i=1}^{n+1}$, is a finite sequence of Rademacher functions, $\varepsilon' := (\varepsilon_i)_{i=1}^n$ and ψ is a function of $n+1$-sequences, then

$$\begin{aligned}
\mathbb{E}\psi(\varepsilon) &:= \frac{1}{2^{n+1}} \sum_{\epsilon\in\{-1,1\}^{n+1}} \psi(\epsilon) \\
&= \frac{1}{2^n} \sum_{\epsilon'\in\{-1,1\}^n} \left(\frac{1}{2}\sum_{\epsilon_{n+1}\in\{-1,1\}} \psi(\epsilon', \epsilon_{n+1})\right) =: \mathbb{E}\left(\mathbb{E}\left(\psi(\varepsilon', \varepsilon_{n+1})\middle|\,\varepsilon'\right)\right) \\
&= \frac{1}{2} \sum_{\epsilon_{n+1}\in\{-1,1\}} \left(\frac{1}{2^n}\sum_{\epsilon'\in\{-1,1\}^n} \psi(\epsilon', \epsilon_{n+1})\right) =: \mathbb{E}\left(\mathbb{E}\left(\psi(\varepsilon', \varepsilon_{n+1})\middle|\,\varepsilon_{n+1}\right)\right).
\end{aligned}$$

We have here introduced the notion of conditional expectation with the naïve meaning of "taking the expectation with some variables held constant". This is fully sufficient when dealing with simple functions. Recall that the expectation is just the integral over the probability space; this is related to the randomized norm notation adopted earlier by $|\sum_{k=1}^n \varepsilon_k x_k|_{L^p(\Omega;X)} = \left(\mathbb{E}|\sum_{k=1}^n \varepsilon_k x_k|_X^p\right)^{\frac{1}{p}}$.

Now for the chaos inequality:

Proposition 3.10. *In a linear normed space X, the Rademacher chaoses obey the law*

$$\left|\sum_{k=0}^n \left(\frac{p-1}{q-1}\right)^{\frac{k}{2}} \sum_{\#I=k} x_I \prod_{i\in I} \varepsilon_i\right|_{L^q(\Omega;X)} \le \left|\sum_{k=0}^n \sum_{\#I=k} x_I \prod_{i\in I} \varepsilon_i\right|_{L^p(\Omega;X)}$$

for $1 < p \le q < \infty$.

Proof. We have already proved the case $n = 1$ in Lemma 3.9. We assume for induction that the assertion is valid for some $n \ge 1$. The proof consists of showing that the assertion then holds for $n+1$ in place of n. We again denote $\lambda := \sqrt{\frac{p-1}{q-1}}$. It is also worth making the simple observation that each $I \subset \{1, \ldots, n+1\}$ either satisfies $I \subset \{1, \ldots, n\}$, or $I \ni n+1$ and $I\setminus\{n+1\} \subset \{1, \ldots, n\}$. We denote $\mathbb{Z}_n := \{1, \ldots, n\}$. For notational convenience only in this proof, we will use the special expectation symbols $\mathbb{E}' := \mathbb{E}(\,\cdot\,|\,\varepsilon_{n+1})$ and $\mathbb{E}_{n+1} := \mathbb{E}(\,\cdot\,|\,\varepsilon_1, \ldots, \varepsilon_n)$. These symbols emphasize the facts that $\mathbb{E}_{n+1}$ is the expectation when the random variable ε_{n+1} varies and the others are held constant; $\mathbb{E}'$ works the other way round. Now comes the induction step:

$$\left(\mathbb{E}\left|\sum_{k=0}^{n+1}\lambda^k \sum_{\substack{\#I=k\\ I\subset\mathbb{Z}_{n+1}}} x_I\prod_{i\in I}\varepsilon_i\right|_X^q\right)^{\frac{1}{q}} = \left(\mathbb{E}'\left[\mathbb{E}_{n+1}\left|\sum_{k=0}^{n}\lambda^k \sum_{\substack{\#I=k\\ I\subset\mathbb{Z}_{n}}} (x_I+\lambda\varepsilon_{n+1}x_{I\cup\{n+1\}})\prod_{i\in I}\varepsilon_i\right|_X^q\right]\right)^{\frac{1}{q}}$$

$$\leq \left(\mathbb{E}'\left[\mathbb{E}_{n+1}\left|\sum_{k=0}^{n}\lambda^k \sum_{\#I=k} (x_I+\varepsilon_{n+1}x_{I\cup\{n+1\}})\prod_{i\in I}\varepsilon_i\right|_X^p\right]^{\frac{q}{p}}\right)^{\frac{1}{q}}$$

$$\leq \left(\mathbb{E}_{n+1}\left[\mathbb{E}'\left|\sum_{k=0}^{n}\lambda^k \sum_{\#I=k} (x_I+\varepsilon_{n+1}x_{I\cup\{n+1\}})\prod_{i\in I}\varepsilon_i\right|_X^q\right]^{\frac{p}{q}}\right)^{\frac{1}{p}}$$

$$\leq \left(\mathbb{E}_{n+1}\left[\mathbb{E}'\left|\sum_{k=0}^{n} \sum_{\substack{\#I=k\\ I\subset\mathbb{Z}_{n}}} (x_I+\varepsilon_{n+1}x_{I\cup\{n+1\}})\prod_{i\in I}\varepsilon_i\right|_X^p\right]^{\frac{p}{p}}\right)^{\frac{1}{p}} = \left(\mathbb{E}\left|\sum_{k=0}^{n+1} \sum_{\substack{\#I=k\\ I\subset\mathbb{Z}_{n+1}}} x_I\prod_{i\in I}\varepsilon_i\right|_X^p\right)^{\frac{1}{p}}.$$

The first inequality was the case $n = 1$ of the assertion (i.e, Lemma 3.9) and the third the induction assumption. The second inequality used a simple consequence of Minkowski's integral inequality, for $\frac{q}{p} \geq 1$,

$$\left\{\int\left(\int f^p(\eta,\zeta)d\mathfrak{L}_{\varepsilon_{n+1}}(\zeta)\right)^{\frac{q}{p}} d\mathfrak{L}_{\varepsilon_1,\ldots,\varepsilon_n}(\eta)\right\}^{\frac{p}{q}\frac{1}{p}}$$

$$\leq \left\{\int\left(\int f^{p\frac{q}{p}}(\eta,\zeta)d\mathfrak{L}_{\varepsilon_1,\ldots,\varepsilon_n}(\eta)\right)^{\frac{p}{q}} d\mathfrak{L}_{\varepsilon_{n+1}}(\zeta)\right\}^{\frac{1}{p}},$$

where $\mathfrak{L}_{\varepsilon_{n+1}} := \mathbb{P}\circ\varepsilon_{n+1}^{-1}$ is the law of the random variable ε_{n+1}, and the joint law is defined as usual by $\mathfrak{L}_{\varepsilon_1,\ldots,\varepsilon_n} := \mathfrak{L}_{\varepsilon_1}\times\ldots\times\mathfrak{L}_{\varepsilon_n}$. (The fact that the expectations and conditional expectations can be computed by integrating with respect to the joint laws under very general circumstances is a deep result in the theory of product measures. In the present context involving only simple functions, the result is easily verified without resorting to these theorems.)

The proposition is now proved. □

Corollary 3.11. *In a linear normed space X, the following inequalities are valid for*

$$\mathfrak{X} := \sum_{1\leq i_1<\ldots<i_r\leq n} x_{i_1,\ldots,i_r}\varepsilon_{i_1}\cdots\varepsilon_{i_n},$$

where $x_{i_1,\ldots,i_r}\in X$ for $1\leq i_1<\ldots<i_r\leq n$, and ε_i are Rademacher functions, $i = 1,\ldots,n$:

$$|\mathfrak{X}|_{L^q(\Omega;X)} \leq \left(\frac{q-1}{p-1}\right)^{\frac{r}{2}}|\mathfrak{X}|_{L^p(\Omega;X)} \qquad \textit{for} \qquad 1<p\leq q<\infty \tag{3.6}$$

$$|\mathfrak{X}|_{L^2(\Omega;X)} \leq \exp\left(r(\frac{2}{p}-1)\right)|\mathfrak{X}|_{L^p(\Omega;X)} \qquad \textit{for} \qquad 0<p\leq 2 \tag{3.7}$$

Functions of the form of $\mathfrak{X}$ are referred to as **homogeneous Rademacher chaoses**. We should note in (3.7), that the constant 2 has no deeper meaning than the fact that the computations are simplified by this choice. What is relevant to us, is the fact that we obtain the inequality for p in an open interval starting from 0 and extending beyond 1.

Proof. For (3.6), set $x_I := 0$ for $\#I \neq r$ in Proposition 3.10. We will derive (3.7) from (3.6).

We choose a $q > 1$ and compute

$$\mathbb{E}(|\mathfrak{X}|_X^2) = \mathbb{E}(|\mathfrak{X}|_X^{\frac{p}{q}} |\mathfrak{X}|_X^{2-\frac{p}{q}}) \leq (\mathbb{E}(|\mathfrak{X}|_X^p))^{\frac{1}{q}} \left(\mathbb{E}(|\mathfrak{X}|_X^{(2-\frac{p}{q})\overline{q}})\right)^{\frac{1}{\overline{q}}} = (\mathbb{E}(|\mathfrak{X}|_X^p))^{\frac{1}{q}} \left(\mathbb{E}(|\mathfrak{X}|_X^{\frac{2q-p}{q-1}})\right)^{\frac{q-1}{q}}$$

Since $p \leq 2$, $\frac{2q-p}{q-1} \geq \frac{2q-2}{q-1} = 2$, and we can apply (3.6) of the corollary to deduce

$$\left(\mathbb{E}(|\mathfrak{X}|_X^{\frac{2q-p}{q-1}})\right)^{\frac{q-1}{q}} \leq \left(\frac{2q-p}{q-1} - 1\right)^{\frac{r}{2}\frac{2q-p}{q-1}\frac{q-1}{q}} \left(\mathbb{E}(|\mathfrak{X}|_X^2)\right)^{\frac{1}{2}\frac{2q-p}{q-1}\frac{q-1}{q}} .$$

Moving the factors with $\mathbb{E}(|\mathfrak{X}|_X^2)$ to the left-hand side of the inequality and raising both sides to the power of $\frac{q}{p}$, we obtain

$$\left(\mathbb{E}(|\mathfrak{X}|_X^2)\right)^{\frac{1}{2}} \leq \left(1 + \frac{2-p}{q-1}\right)^{\frac{r}{p}(q-\frac{p}{2})} (\mathbb{E}(|\mathfrak{X}|_X^p))^{\frac{1}{p}} .$$

In the limit $q \to \infty$, the coefficient tends to $\exp\left(\frac{r}{p}(2-p)\right)$, but this is just what we wanted to prove. $\square$

Corollary 3.12 (Khintchine–Kahane inequality). *For $0 < p, q < \infty$, there exist finite constants $K_{q,p}$ so that, in every normed linear space X,*

$$\left|\sum_{k=1}^{n} \varepsilon_k x_k\right|_{L^q(\Omega;X)} \leq K_{q,p} \left|\sum_{k=1}^{n} \varepsilon_k x_k\right|_{L^p(\Omega;X)}$$

for all $x_k \in X$, and all Rademacher functions ε_k, $k = 1, \dots, n$.

If X is also complete (i.e., a Banach space), and the series $\sum_{k=1}^{\infty} \varepsilon_k x_k$ converges in one of the L^p norms, then it converges in each of these norms and the inequality above also holds with n replaced by ∞.

This says that in the linear span of constant vector multiples of the Rademacher functions, all L^p norms, $0 < p < \infty$, are equivalent. Strictly speaking, $|\cdot|_{L^p(\Omega;X)}$ with $p < 1$ is not a norm, but the previous statement is nevertheless true with the obvious interpretation.

Proof. For $q \leq p$, Jensen's inequality shows the claim with $K_{q,p} = 1$. For $1 < p \leq q < \infty$, this is a special case of (3.6) in Corollary 3.11 with $r = 1$, and we can take $K_{q,p} = \sqrt{\frac{q-1}{p-1}}$. For $0 < p \leq 1$ and $p < q < \infty$ (or actually $0 < p \leq 2$, but part of this range is already covered), we can estimate the L^q norm by the L^2 norm (with $K_{q,2} = 1$ if $q \leq 2$ and $\sqrt{q-1}$ otherwise) and the L^2 norm by the L^p norm according to (3.7) in Corollary 3.11, with $K_{2,p} = e^{\frac{2}{p}-1}$. Thus $K_{q,p} = K_{q,2}K_{2,p} = (\sqrt{q-1} \vee 1)e^{\frac{2}{p}-1}$ will do for $0 < p \leq 2$ and arbitrary q.

For the assertion concerning Banach spaces, recall that each $L^p(\Omega; X)$, $p \in [1, \infty)$ is Banach, when X is, and $L^p(\Omega; X)$, $p \in (0, 1)$, is a complete metric space with the metric $\varrho(f, g) := |f - g|_{L^p(\Omega;X)}^p$. (These assertions can be shown as in the real case, e.g. [20].) It is immediate from the inequality with finite n, that the sequence of partial sums of $\sum_{k=1}^{\infty} \varepsilon_k x_k$ is Cauchy in either all $L^p(\Omega; X)$ or none. Hence the convergence of the series in some L^p allows us to deduce convergence in all L^p and thus to pass to the limit $n \to \infty$ to deduce the desired inequality for infinite series. $\square$

If the norm of X is induced by an inner product, then the Khintchine–Kahane inequality can be written in the form in which the scalar valued version is traditionally stated.

Corollary 3.13. *For all $p \in (0,\infty)$, there are finite constants a_p, A_p so that, for each inner product space X, all $n \in \mathbb{Z}_+$ and $x_k \in X$, ε_k Rademacher functions for $k = 1,\dots,n$, we have*

$$a_p \left|\sum_{k=1}^n \varepsilon_k x_k\right|_{L^p(\Omega;X)} \leq \left|\sum_{k=1}^n \varepsilon_k x_k\right|_{L^2(\Omega;X)} = \sqrt{\sum_{k=1}^n |x_k|_X^2} \leq A_p \left|\sum_{k=1}^n \varepsilon_k x_k\right|_{L^p(\Omega;X)}.$$

If X is also complete (i.e., a Hilbert space) and if $\sum_{k=1}^\infty |x_k|_X^2 < \infty$, then the previous inequality holds with n replaced by ∞, and all the series are convergent in the corresponding norms.

Proof. Clearly the equality in the middle is all that is new in the first assertion compared to the general form of the Khintchine–Kahane inequality. This equality follows readily:

$$\left|\sum_{k=1}^n \varepsilon_k x_k\right|_{L^2(\Omega;X)} = \int_\Omega \left(\sum_{k=1}^n \varepsilon_k x_k, \sum_{j=1}^n \varepsilon_j x_j\right)_X d\mathbb{P} = \sum_{k,j=1}^n (x_k, x_j)_X \int_\Omega \varepsilon_k \varepsilon_j d\mathbb{P} = \sum_{k=1}^n |x_k|_X^2,$$

where we used the fact that the Rademacher functions are orthonormal.

For the second assertion, observe that $L^2(\Omega;X)$ is also a Hilbert space, when X is, with the inner product $(f,g)_{L^2(\Omega;X)} := \int_\Omega (f(\omega), g(\omega))_X \, d\mathbb{P}(\omega)$. Also observe that $\varepsilon_k x_k |x_k|_X^{-1}$ are orthonormal in $L^2(\Omega;X)$ by the previous computation, and thus

$$\sum_{k=1}^\infty \varepsilon_k x = \sum_{k=1}^\infty |x_k|_X \frac{\varepsilon_k x_k}{|x_k|_X}$$

converges in $L^2(\Omega;X)$ if and only $\sum_{k=1}^\infty |x_k|_X^2$ converges. The rest now follows from the inequality of Khintchine and Kahane for series in Banach spaces. □

3.4 Notes and comments

Section 3.2 comes from Witvliet [28] and Section 3.3 from de la Peña and Giné [4].

The best constants $K_{q,p}$ in the Khintchine–Kahane inequalities are sometimes of interest, and they are known in some cases; the present proof does not give the smallest constants (except, of course, in the rather trivial case $q \leq p$, when $K_{q,p} = 1$). Latała and Oleszkiewicz [13] give an ingenious elementary proof providing the best constant $K_{2,1} = \sqrt{2}$ in the important case of L^2 and L^1; this proof is also found in [4].

There is some variation in the literature concerning the name of the Khintchine–Kahane inequality. Sometimes the scalar valued version of the inequality (often stated in a similar form as Corollary 3.13) is related to the name of Khintchine, who first proved it, whereas the vector valued version is called Kahane's inequality. The classical paper of Kahane [12] containing some inequalities involving randomized norms is often cited in this context; however, the Khintchine–Kahane inequality does not actually appear in this paper. Somewhat easier reasoning can be used to prove the scalar inequality; see e.g. Stein [23].

There is also variation in the way of spelling the name of Khintchine. We have included the maximum number of characters, but in an urgent need to save ink it seems to be possible to leave out the "t", or the "e", or even both.

Chapter 4

R-boundedness

4.1 Introduction

The notion of R-boundedness has proved to be a significant tool in the study of abstract multiplier operators, and it has other far reaching applications falling outside the scope of this work. The definition is given as follows:

Definition 4.1. *A family of bounded linear operators $\mathcal{T} \subset \mathcal{B}(X;Y)$, with X and Y normed linear spaces, is called* ***randomized bounded (R-bounded)*** *if for some $p \in [1,\infty)$ there exists a finite C such that, for all $n \in \mathbb{Z}_+$ and all $T_k \in \mathcal{T}$, $x_k \in X$ and ε_k Rademacher functions, $k = 1,\dots,n$, we have the inequality*

$$\left|\sum_{k=1}^{n} \varepsilon_k T_k x_k\right|_{L^p(\Omega;Y)} \le C \left|\sum_{k=1}^{n} \varepsilon_k x_k\right|_{L^p(\Omega;X)}. \tag{4.1}$$

The smallest C is denoted by $\mathcal{R}_p(\mathcal{T})$ and called the ***R-bound*** *of $\mathcal{T}$ of order p.*

Remark 4.2. *By the Khintchine–Kahane inequality, the definition of R-boundedness is independent of the order p in the sense that any $\mathcal{T} \subset \mathcal{B}(X;Y)$ either satisfies the condition for all $p \in [1,\infty)$ or for none of them. (The R-bounds $\mathcal{R}_p(\mathcal{T})$ may depend on p, though.) In fact, the Khintchine–Kahane inequality shows that we could take different exponents $p, q \in [1,\infty)$ on the two sides of the inequality defining R-boundedness, and the resulting inequality either holds for all pairs (p,q) or for none of them.*

Several properties of R-bounds follow immediately. We first note that these bounds behave like norms:

$$\mathcal{R}_p(\mathcal{T}+\mathcal{S}) \le \mathcal{R}_p(\mathcal{T}) + \mathcal{R}_p(\mathcal{S}), \qquad \mathcal{R}_p(\mathcal{T}\mathcal{S}) \le \mathcal{R}_p(\mathcal{T})\mathcal{R}_p(\mathcal{S}),$$

The first property above follows from the triangle inequality and the second by applying the definition of R-boundedness twice, first to the family $\mathcal{T}$, then to $\mathcal{S}$. Also, every set of one bounded linear operator is R-bounded, and $\mathcal{R}_p\{T\} = |T|_{\mathcal{B}(X;Y)}$, as is easily seen from the definition (4.1) by extracting the norm of T from the left-hand side.

Clearly a subset of an R-bounded set is also R-bounded. It is also useful to observe that we can always assume that $0 \in \mathcal{T}$ without affecting the R-bounds:

Lemma 4.3. *For $\mathcal{T} \subset \mathcal{B}(X;Y)$, $\mathcal{R}_p(\mathcal{T} \cup \{0\}) = \mathcal{R}_p(\mathcal{T})$.*

Proof. Let $T_k \in \mathcal{T} \cup \{0\}$, $k = 1,\dots,n$, and let $J \subset \{1,\dots,n\}$ consists of those k for which $T_k \neq 0$. Then

$$\left|\sum_{k=1}^{n} \varepsilon_k T_k x_k\right|_{L^p(\Omega;Y)} = \left|\sum_{k\in J} \varepsilon_k T_k x_k\right|_{L^p(\Omega;Y)} \le \mathcal{R}_p(\mathcal{T}) \left|\sum_{k\in J} \varepsilon_k x_k\right|_{L^p(\Omega;X)} \le \mathcal{R}_p(\mathcal{T}) \left|\sum_{k=1}^{n} \varepsilon_k x_k\right|_{L^p(\Omega;X)},$$

where the last inequality follows from the monotonicity of the basic sequence $\{\varepsilon_k x_k\}_{k=1}^{\infty}$ (Example 3.1). $\square$

With the convention that $0 \in \mathcal{T}, \mathcal{S}$, we immediately see that $\mathcal{T} \cup \mathcal{S} \subset \mathcal{T} + \mathcal{S}$ is also R-bounded if $\mathcal{T}$ and $\mathcal{S}$ are. Now iteration of the triangle inequality for $\mathcal{R}_p(\mathcal{S} + \mathcal{T})$, with $\mathcal{S}$ and $\mathcal{T}$ singletons, shows that every finite family of bounded linear operators is R-bounded. These properties seem quite reasonable.

To provide some insight into the abstract notion, we give a few examples of the meaning of R-boundedness in some special spaces.

Example 4.4. *Let $\mathcal{T} \subset \mathcal{B}(X;Y)$, X, Y normed linear spaces.*

1. *If $\mathcal{T}$ is R-bounded, then it is uniformly bounded, with*

$$\sup_{T \in \mathcal{T}} |T|_{\mathcal{B}(X;Y)} \le \inf_{p \in [1,\infty)} \mathcal{R}_p(\mathcal{T}).$$

2. *The converse of 1 is true if X and Y are inner product spaces, and in this case $\mathcal{R}_2(\mathcal{T}) = \sup_{T \in \mathcal{T}} |T|_{\mathcal{B}(X;Y)}$.*

Proof. Part 1 is immediate from the definition, taking $n := 1$. Part 2 follows from the Khintchine–Kahane inequality for inner product spaces (Corollary 3.13). □

Example 4.5. *If $X = L^p(\Gamma_1; \mathcal{H}_1)$, $Y = L^q(\Gamma_2; \mathcal{H}_2)$, where Γ_1, Γ_2 are measure spaces equipped with σ-finite measures μ_1, μ_2 (respectively), and $\mathcal{H}_1, \mathcal{H}_2$ are Hilbert spaces, then $\mathcal{T} \subset \mathcal{B}(X;Y)$ is R-bounded if and only if*

$$\left| \left(\sum_{k=1}^{n} |T_k f_k(\cdot)|^2_{\mathcal{H}_2} \right)^{\frac{1}{2}} \right|_{L^q(\Gamma_2)} \le M \left| \left(\sum_{k=1}^{n} |f_k(\cdot)|^2_{\mathcal{H}_1} \right)^{\frac{1}{2}} \right|_{L^p(\Gamma_1)} \tag{4.2}$$

for some finite M, for all $n \in \mathbb{Z}_+$ and $f_k \in X$, $k = 1, \ldots, n$.

Note that the asserted equivalent condition (4.2) can also be formulated by requiring that the operators $T : L^p(\Gamma_1; \ell^2(\mathbb{Z}_+; \mathcal{H}_1)) \to L^q(\Gamma_2; \ell^2(\mathbb{Z}_+; \mathcal{H}_2))$ defined by $Tf := (T_k f_k)_{k=1}^{\infty}$ for $f = (f_k)_{k=1}^{\infty} \in L^p(\Gamma_1; \ell^2(\mathbb{Z}_+; \mathcal{H}_1))$ be bounded for all sequences $\{T_k\}_{k=1}^{\infty} \subset \mathcal{T}$.

Proof. Assume the condition (4.2). Using Fubini's theorem to change the order of integration when desirable (observing in particular that $L^r(\Omega; L^r(\Gamma_i; \mathcal{H}_i)) = L^r(\Gamma_i; L^r(\Omega; \mathcal{H}_i))$), we compute

$$\begin{aligned}
\left| \sum_{k=1}^{n} \epsilon_k T_k f_k \right|_{L^q(\Omega; L^q(\Gamma_2; \mathcal{H}_2))} &= \left(\int_{\Gamma_2} \left| \sum_{k=1}^{n} \epsilon_k T_k f_k \right|^q_{L^q(\Omega; \mathcal{H}_2)} d\mu_2 \right)^{\frac{1}{q}} \\
&\le a_q^{-1} \left(\int_{\Gamma_2} \left(\sum_{k=1}^{n} |T_k f_k|^2_{\mathcal{H}_2} \right)^{\frac{q}{2}} d\mu_2 \right)^{\frac{1}{q}} = a_q^{-1} \left| \left(\sum_{k=1}^{n} |T_k f_k|^2_{\mathcal{H}_2} \right)^{\frac{1}{2}} \right|_{L^q(\Gamma_2)} \\
&\le a_q^{-1} M \left| \left(\sum_{k=1}^{n} |f_k|^2_{\mathcal{H}_1} \right)^{\frac{1}{2}} \right|_{L^p(\Gamma_1)} = a_q^{-1} M \left(\int_{\Gamma_1} \left(\sum_{k=1}^{n} |f_k|^2_{\mathcal{H}_1} \right)^{\frac{p}{2}} d\mu_1 \right)^{\frac{1}{p}} \\
&\le a_q^{-1} M A_p \left(\int_{\Gamma_1} \left| \sum_{k=1}^{n} \epsilon_k f_k \right|^p_{L^p(\Omega; \mathcal{H}_1)} d\mu_1 \right)^{\frac{1}{p}} = \left| \sum_{k=1}^{n} \epsilon_k f_k \right|_{L^p(\Omega; L^p(\Gamma_1; \mathcal{H}_1))}.
\end{aligned}$$

The first and last inequalities used the inner product space version of the Khintchine–Kahane inequality; the second was the assumption of the condition (4.2). The computation shows that this condition implies that $\mathcal{T}$ is R-bounded (using Remark 4.2 on different orders p, q in the definition of R-boundedness). The proof of the converse inequality follows the same pattern. □

The examples show that in special cases R-boundedness reduces to conditions which do not involve any randomization, and in a Hilbert space, this notion does not give anything new. In general, however, the randomization gives significant flexibility due to the equivalence of different L^p norms, as we saw in the previous proof. By now one should appreciate the power of the Khintchine–Kahane inequality, which reduced the proofs of the previous assertions to straightforward computations.

4.2 Elementary properties

Here we give a survey of some simple properties of R-bounds and provide further examples of R-bounded sets of operators. We start with a couple of technical results related to the verification of the R-boundedness of a given set of operators.

Lemma 4.6. *To check the R-boundedness of a family $\mathcal{T} \subset \mathcal{B}(X;Y)$, it is sufficient to verify the inequality (4.1) for all sequences of distinct elements $T_k \in \mathcal{T}$. The best constants are the same.*

It is obvious that the sufficient condition here is also necessary.

Proof. Suppose that the inequality (4.1) holds whenever $T_i \neq T_j$ unless $i = j$. Then consider a general sequence $\{T_j\}_{j=1}^n \subset \mathcal{T}$. Denote by S_k, $k = 1, \ldots, m$, the distinct operators in this sequence, and by I_k the set of those $j \in \{1, \ldots, n\}$ for which $T_j = S_k$. We then compute

$$\left|\sum_{j=1}^n \varepsilon_j T_j x_j\right|_{L^p(\Omega;X)}^p = \left|\sum_{k=1}^m S_k \sum_{j\in I_k} \varepsilon_j x_j\right|_{L^p(\Omega;X)}^p = \left|\sum_{k=1}^m \varepsilon_k'(\omega') S_k \sum_{j\in I_k} \varepsilon_j x_j\right|_{L^p(\Omega;X)}^p.$$

In the last step we inserted in the expression auxiliary Rademacher functions ε_k' on another probability space Ω', recalling that the joint distribution of the Rademacher functions remains invariant under change of signs. We integrate over Ω', and change the order of integration to obtain

$$\begin{aligned}\int_{\Omega'} \left|\sum_{k=1}^m \varepsilon_k'(\omega') S_k \sum_{j\in I_k} \varepsilon_j x_j\right|_{L^p(\Omega;X)}^p d\mathbb{P}'(\omega') &= \int_{\Omega} \left|\sum_{k=1}^m \varepsilon_k' S_k \sum_{j\in I_k} \varepsilon_j(\omega) x_j\right|_{L^p(\Omega';X)}^p d\mathbb{P}(\omega) \\ &\leq C^p \int_{\Omega} \left|\sum_{k=1}^m \varepsilon_k' \sum_{j\in I_k} \varepsilon_j(\omega) x_j\right|_{L^p(\Omega';X)}^p d\mathbb{P}(\omega),\end{aligned}$$

where the assumption was used point-wise for each fixed $\omega' \in \Omega'$ inside the integral. Reverting the steps in which we introduces the auxiliary ε_k', we can manipulate the right-hand side of the last inequality into the form $C^p \left|\sum_{j=1}^n \varepsilon_j x_j\right|_{L^p(\Omega;X)}^p$, and this shows the claim. ∎

Corollary 4.7. *If $\mathcal{T} = \{T^k\}_{k=1}^\infty \subset \mathcal{B}(X;Y)$ is a countable sequence of operators, then it is sufficient to verify the inequality (4.1) for all truncated sequences $\{T^k\}_{k=1}^n$ of the first n members of the sequence.*

It is clear that the R-boundedness of the (countable) set $\mathcal{T}$ is independent of the order in which we enumerate its element. Thus it is interesting that, given any enumeration, the subsets of n first members of the sequence are fully representative of all finite subsets of $\mathcal{T}$ in view of R-boundedness; this is what the assertion above states.

Proof. If S_k, $k = 1, \ldots, n$ are distinct members of $\mathcal{T}$, then there are numbers $m_1 < \ldots < m_n$ such that $\{S_k\}_{k=1}^n = \{T^{m_k}\}_{k=1}^n$. Thus, if the inequality (4.1) holds for all for the truncated sequences as in the assertion, we have, setting $x^{m_k} := x_k$ and $x^j := 0$ for $j \notin \{m_k\}_{k=1}^n$ and letting $\{\varepsilon^j\}_{j=1}^{m_n}$ be another sequence of Rademacher functions,

$$\left|\sum_{k=1}^n \varepsilon_k S_k x_k\right|_{L^p(\Omega;X)} = \left|\sum_{j=1}^{m_n} \varepsilon^j T^j x^j\right|_{L^p(\Omega;X)} \le C\left|\sum_{j=1}^{m_n} \varepsilon^j x^j\right|_{L^p(\Omega;X)} = C\left|\sum_{k=1}^n \varepsilon_k x_k\right|_{L^p(\Omega;X)}.$$

Since this holds for all distinct $S_k \in \mathcal{T}$, $k = 1, \dots, n$, the R-boundedness of $\mathcal{T}$ follows from Lemma 4.6. □

A very useful device in connection with R-bounds is the contraction principle of Kahane:

Lemma 4.8 (Kahane's contraction principle). *For $\alpha_k, \beta_k \in \mathbb{C}$, $|\alpha_k| \le |\beta_k|$, $x_k \in X$, ε_k Rademacher functions, $k = 1, \dots, n$ we have*

$$\left|\sum_{k=1}^n \alpha_k \varepsilon_k x_k\right|_{L^p(\Omega;X)} \le 2\left|\sum_{k=1}^n \beta_k \varepsilon_k x_k\right|_{L^p(\Omega;X)}.$$

The coefficient 2 is not needed if α_k, β_k are real.

Proof. By considering new vectors $y_k := \beta_k x_k$ if necessary, we can always reduce the inequality to the case $\beta_k = 1$, $|\alpha_k| \le 1$, $k = 1, \dots, n$. If the α_k are real, i.e., $\alpha_k \in [-1, 1]$, then Lemma 2.3 applies to give

$$\left|\sum_{k=1}^n \alpha_k \varepsilon_k x_k\right|_{L^p(\Omega;X)} \le \max_{\epsilon \in \{-1,1\}^n} \left|\sum_{k=1}^n \epsilon_k \varepsilon_k x_k\right|_{L^p(\Omega;X)} = \left|\sum_{k=1}^n \varepsilon_k x_k\right|_{L^p(\Omega;X)},$$

where the last equality used the property of the Rademacher functions that $\epsilon_k \varepsilon_k$ and ε_k have the same joint distribution. The assertion for real coefficients is hence established. The complex case follows by applying the first part of the proof to the real and imaginary parts separately. □

The contraction principle has many corollaries. Before proceeding to them, we formulate a generalized notion of R-boundedness, which is useful in some applications:

Definition 4.9. *A sequence $(\mathcal{T}_k)_{k=1}^\infty$ of operator families $\mathcal{T}_k \subset \mathcal{B}(X;Y)$ is called **R-bounded relative to** a sequence $(X_k)_{k=1}^\infty$ of closed subspaces of X if*

$$\left|\sum_{k=1}^n \varepsilon_k T_k x_k\right|_{L^p(\Omega;Y)} \le C\left|\sum_{k=1}^n \varepsilon_k x_k\right|_{L^p(\Omega;X)}$$

for all $n \in \mathbb{Z}_+$, ε_k Rademacher functions, $T_k \in \mathcal{T}_k$ and $x_k \in X_k$, $k = 1, \dots, n$.

Observe that $\mathcal{T}$ is R-bounded if and only if all sequences $(\mathcal{T}_k)_{k=1}^\infty$, $\mathcal{T}_k \subset \mathcal{T}$, are relatively R-bounded for all $(X_k)_{k=1}^\infty$, $X_k \subset X$.

Corollary 4.10. *If $\mathcal{T} \subset \mathcal{B}(X;Y)$ is R-bounded and $\overline{B}(0;r) \subset \mathbb{C}$ is the closed ball of radius r centered at the origin of the complex plane, then*

$$\mathcal{R}(\overline{B}(0;r)\mathcal{T}) \le 2r\mathcal{R}_p(\mathcal{T}).$$

For $[-r, r]$ in place of $\overline{B}(0;r)$, the 2 can be omitted. The same results hold for relatively R-bounded sequences $(\mathcal{T}_k)_{k=1}^\infty$, if we multiply each family in the sequence by $\overline{B}(0;r)$ or $[-r, r]$.

Proof. It is clearly sufficient to consider the relative R-boundedness. Take $T_k \in \mathcal{T}_k$, $\zeta_k \in \overline{B}(0;r)$, $x_k \in X_k$, $k = 1, \dots, n$. Then

$$\left|\sum_{k=1}^n \varepsilon_k \zeta_k T_k x_k\right|_{L^p(\Omega;X)} \le 2\left|\sum_{k=1}^n \varepsilon_k r T_k x_k\right|_{L^p(\Omega;X)} \le 2r \cdot C \cdot \left|\sum_{k=1}^n \varepsilon_k x_k\right|_{L^p(\Omega;X)},$$

where the first inequality was the contraction principle (thus the 2 can be omitted for real ζ_k), and the second simply the definition of the (relative) R-boundedness, C being the relative R-bound. For proper R-boundedness, $C = \mathcal{R}_p(\mathcal{T})$. □

Example 4.11. *Let $\Phi \subset L^\infty(\Gamma)$ be uniformly bounded. Then*

$$\mathcal{R}_p(\{m_\phi : L^p(\Gamma; X) \to L^p(\Gamma; X) : f \mapsto \phi f\}_{\phi\in\Phi}) \leq 2 \sup_{\phi\in\Phi} |\phi|_{L^\infty(\Gamma)}$$

Proof. Using Fubini's theorem and the contraction principle, we obtain

$$\begin{aligned}\left|\sum_{k=1}^n \varepsilon_k m_\phi f\right|_{L^p(\Omega;L^p(\Gamma;X))} &= \left(\int_\Gamma \left|\sum_{k=1}^n \varepsilon_k \phi(\gamma) f(\gamma)\right|^p_{L^p(\Omega;X)} d\mu(\gamma)\right)^{\frac{1}{p}} \\ &\leq 2 \sup_{\phi\in\Phi} |\phi|_{L^\infty(\Gamma)} \left(\int_\Gamma \left|\sum_{k=1}^n \varepsilon_k f(\gamma)\right|^p_{L^p(\Omega;X)} d\mu(\gamma)\right)^{\frac{1}{p}} = 2 \sup_{\phi\in\Phi} |\phi|_{L^\infty(\Gamma)} \left|\sum_{k=1}^n \varepsilon_k f\right|_{L^p(\Omega;L^p(\Gamma;X))},\end{aligned}$$

and this is just what we claimed. □

The following result shows that R-boundedness, even relative, behaves rather well with some common set operations.

Lemma 4.12. *Let $\mathcal{T} \subset \mathcal{B}(X;Y)$ be R-bounded, and $(\mathcal{T}_k)_{k=1}^\infty$ be R-bounded relative to $(X_k)_{k=1}^\infty$. Then the same is true for the following families obtained from the original ones:*

1. *the strong closures $\overline{\mathcal{T}}$ and $(\overline{\mathcal{T}_k})_{k=1}^\infty$,*
2. *the convex hulls* $\operatorname{conv}\mathcal{T}$ *and* $(\operatorname{conv}\mathcal{T}_k)_{k=1}^\infty$, *and*
3. *the (complex)* ***absolute convex hulls***

$$\operatorname{abco}(\mathcal{T}) := \left\{\sum_{j=1}^n \lambda_j T_j : n \in \mathbb{Z}_+; \lambda_j \in \mathbb{C}; T_j \in \mathcal{T}; j = 1, \ldots, n; \sum_{j=1}^n |\lambda_j| = 1\right\},$$

and $(\operatorname{abco}\mathcal{T}_k)_{k=1}^\infty$, *as well as the real absolute convex hull defined similarly, but with $\mathbb{R}$ in place of $\mathbb{C}$.*

The R-bounds remain the same under these set operations, except for the (complex) convex hull, for which the R-bound is at most doubled.

Note that the set operations are applied to each family $\mathcal{T}_k$ separately in the relative case.

Proof. Again, it is sufficient to consider relative R-boundedness.

1. If $T_k \in \overline{\mathcal{T}}_k$, $k = 1, \ldots, n$, where the bar denotes the strong closure, then there are sequences $\{T_k^j\}_{j=1}^\infty \subset \mathcal{T}_k$ such that $T_k^j x \to T_k x$ for each $x \in X$ as $j \to \infty$. Thus, for $x_k \in X_k$,

$$\left|\sum_{k=1}^n \varepsilon_k T_k x_k\right|_{L^p(\Omega;Y)} \leq \left|\sum_{k=1}^n \varepsilon_k T_k^j x_k\right|_{L^p(\Omega;Y)} + \sum_{k=1}^n \left|T_k x_k - T_k^j x_k\right|_Y \leq C \left|\sum_{k=1}^n \varepsilon_k x_k\right|_{L^p(\Omega;X)} + \epsilon,$$

where the ϵ can be forced as small as one likes by choosing sufficiently large j.

2. We use here the fact that $\operatorname{conv}(\mathcal{T}_1) \times \cdots \times \operatorname{conv}(\mathcal{T}_n) = \operatorname{conv}(\mathcal{T}_1 \times \ldots \times \mathcal{T}_n)$ (Lemma 2.4). Thus $(T_k)_{k=1}^n = \sum_{j=1}^N \lambda_j (T_k^j)_{j=1}^n$, i.e., $T_k = \sum_{j=1}^N \lambda_j T_k^j$ with $T_k^j \in \mathcal{T}_k$ and $\lambda_k \geq 0$, $\sum_{k=1}^N \lambda_k = 1$, whenever $T_k \in \operatorname{conv}(\mathcal{T}_k)$, $k = 1, \ldots, n$; thus

$$\begin{aligned}\left|\sum_{k=1}^n \varepsilon_k T_k x_k\right|_{L^p(\Omega;Y)} &= \left|\sum_{k=1}^n \varepsilon_k \sum_{j=1}^N \lambda_j T_k^j x_k\right|_{L^p(\Omega;Y)} \\ &\leq \sum_{j=1}^N \lambda_j \left|\sum_{k=1}^n \varepsilon_k T_k^j x_k\right|_{L^p(\Omega;Y)} \leq \sum_{j=1}^N \lambda_j \cdot C \left|\sum_{k=1}^n \varepsilon_k x_k\right|_{L^p(\Omega;X)},\end{aligned}$$

and this last expression is of the desired form since $\sum_{j=1}^{N}\lambda_j = 1$.

3. For the absolute convex hull, we observe that, whenever $\sum_{j=1}^{n}|\lambda_j| = 1$,

$$\sum_{j=1}^{n}\lambda_j T_j = \sum_{j=1}^{n}|\lambda_j|\left(\frac{\lambda_j}{|\lambda_j|}T_j\right) \in \operatorname{conv}(S(0;1)\mathcal{T}),$$

since $\{|\lambda_j|\}_{j=1}^{n}$ is a proper set of coefficient for a convex combination. ($S(0;1)$ is the unit sphere in $\mathbb{C}$.) Since $S(0;1) \subset \overline{B}(0;1)$, the last assertion follows from Corollary 4.10. $\square$

4.3 Some abstract multiplier theorems

It is time to provide some answers to the natural question: why R-boundedness? We claimed above that the new notion would be useful in the study of multipliers. We now try to verify this claim. In this section, we present a number of results, which follow rather readily from the properties of R-boundedness. We will also see its interplay with the unconditional Schauder decompositions studied in Chapter 2.

In order to study what we have called abstract multiplier operators, we should first define what we mean by this concept. Motivated by the form of multiplier operators acting on Fourier series in Chapter 1, we already investigated in Chapter 2 abstract multiplier operators of the from $T_\lambda x := \sum_{k=1}^{\infty}\lambda_k D_k x$ (where$D = \{D_k\}_{k=1}^{\infty}$ is a Schauder decomposition of the Banach space X). The boundedness of such operators was simply and neatly characterized by Corollary 2.14 for unconditional D. There are two obvious directions in which to generalize matters: either relax the assumptions on D (e.g., give away the unconditionality) or consider more general multipliers. For the second procedure, we introduce abstract multiplier operators of the form

$$Tx := \sum_{k=1}^{\infty} T_k D_k x, \qquad \{T_k\}_{k=1}^{\infty} \subset \mathcal{B}(X;Y). \tag{4.3}$$

To have some multiplier structure, we require that $T_k D_k = \Delta_k T_k D_k$ for all $k \in \mathbb{Z}_+$, where $\Delta = \{\Delta_k\}_{k=1}^{\infty}$ is an unconditional Schauder decomposition of Y. Recalling that the operators D_k, respectively Δ_k, of a Schauder decomposition are projectors onto the closed subspaces $\operatorname{ran}(D_k) \subset X$, respectively $\operatorname{ran}(\Delta_k) \subset Y$, this condition states that T_k maps $\operatorname{ran}(D_k) \subset X$ into $\operatorname{ran}(\Delta_k) \subset Y$.

An important special case is $Y = X$, $\Delta = D$, in which case this says that $\operatorname{ran}(D_k)$ is an invariant subspace of T_k. For this reason, also in the general case, we will refer to multiplier operators T of the form described as **D,Δ-invariant multiplier operators**, and the corresponding sequences $\{T_k\}_{k=1}^{\infty}$ as **D,Δ-invariant sequences** (of operators). For $\Delta = D$, we simply speak of D-invariance. The D-invariance holds, in particular, if D_k commutes with T_k for each $k \in \mathbb{Z}_+$, a slightly stronger requirement.

This generalization of the notion of multipliers seems reasonable, since the starting point in Chapter 1 was the characterization of bounded operators with certain commutativity properties, and since the scalar valued multipliers so far studied certainly commute with “everything”.

The first result providing us with many bounded D-invariant multiplier operators goes as follows:

Theorem 4.13 (Clément et al. 2000). *Let $\{D_k\}_{k=1}^{\infty}$ be an unconditional Schauder decomposition of the Banach space X, and $\{\Delta_k\}_{k=1}^{\infty}$ of Y. Let $\mathcal{T} \subset \mathcal{B}(X;Y)$ be R-bounded. Then for each D,Δ-invariant sequence $\{T_k\}_{k=1}^{\infty} \subset \mathcal{T}$, equation (4.3) defines a bounded linear operator from X to Y. The operators so defined are uniformly bounded with $|T|_{\mathcal{B}(X)} \leq \inf_{p\in[1,\infty)} C_p(\Delta)\mathcal{R}_p(\mathcal{T})C_p(D)$, where $C_p(\Delta), C_p(D)$ are constants as in (3.2) (Lemma 3.2).*

Proof. Using the D, Δ-invariance of $\{T_k\}_{k=1}^{\infty}$, together with Lemma 3.2, we compute

$$\begin{aligned}\left|\sum_{k=m}^{n} T_k D_k x\right|_X &= \left|\sum_{k=m}^{n} \Delta_k T_k D_k x\right|_X \le C_p(\Delta)\left|\sum_{k=m}^{n} \varepsilon_k \Delta_k T_k D_k x\right|_{L^p(\Omega;X)} \\ &= C_p(\Delta)\left|\sum_{k=m}^{n} \varepsilon_k T_k D_k x\right|_{L^p(\Omega;X)} \le C_p(\Delta)\mathcal{R}_p(\mathcal{T})\left|\sum_{k=m}^{n} \varepsilon_k D_k x\right|_{L^p(\Omega;X)} \\ &\le C_p(\Delta)\mathcal{R}_p(\mathcal{T})C_p(D)\left|\sum_{k=m}^{n} D_k x\right|_X .\end{aligned}$$

Since $\sum_{k=1}^{\infty} D_k x_k$ converges for each $x \in X$, the sequence of partial sums is Cauchy, and taking the limit $m, n \to \infty$ in the inequality above shows that the same is true for $\sum_{k=1}^{\infty} T_k D_k x$; thus the operator T in (4.3) is well-defined for each $x \in X$. Setting $m := 0$ and passing n to infinity shows that $|T|_{\mathcal{B}(X)} \le C_p(\Delta)\mathcal{R}_p(\mathcal{T})C_p(D)$. □

The following theorem deals with the other generalization mentioned above, namely relaxing the requirement that the decomposition D be unconditional.

Theorem 4.14 (Marcinkiewicz-type multiplier theorem, Clément et al. 2000). *Let $D = \{D_k\}_{k=1}^{\infty}$ be a Schauder decomposition of the Banach space X, and D' be the blocking corresponding to the sequence $\{n_k\}_{k=1}^{\infty} \subset \mathbb{Z}_+$. Then the following conditions are equivalent:*

1. *The multiplier operators T_λ, defined in (2.3), are uniformly bounded for all $\lambda \in \ell^\infty$ satisfying $|\lambda|_{\ell^\infty} \le 1$ and*

$$\sum_{\ell=n_{k-1}+1}^{n_k-1} |\delta\lambda_{\ell+1}| = \left|\lambda_{n_{k-1}+1} - \lambda_{n_{k-1}+2}\right| + \left|\lambda_{n_{k-1}+2} - \lambda_{n_{k-1}+3}\right| + \ldots + \left|\lambda_{n_k-1} - \lambda_{n_k}\right| \le 1.$$

2. *D' is unconditional and $(\{P_j\}_{j=n_{k-1}+1}^{n_k})_{k=1}^{\infty}$ is R-bounded relative to $(\operatorname{ran} D'_k)_{k=1}^{\infty}$. ($P_j$ denotes denotes the jth partial sum projection of D.)*

Proof. 1 ⇒ 2. Clearly every $\lambda \in \ell^\infty$ bounded in norm by 1 and constant on each block $\{n_{k-1} + 1, \ldots, n_k\}$ but otherwise arbitrary is of the form considered in part 1, and thus T_λ are uniformly bounded for all such λ. The unconditionality of D' then follows from Lemma 2.12(3), since $T_\lambda x = \sum_{k=1}^{\infty} \lambda^k D'_k x$ for such λ, where λ^k is the constant value of λ_j for $n_{k-1} < j \le n_k$.

Let an $n \in \mathbb{Z}_+$ and $\{m_k\}_{k=1}^{\infty} \subset \mathbb{Z}_+$, with $n_{k-1} < m_k \le n_k$ be given. For a fixed $\omega \in \Omega$ and ε_k Rademacher functions, as usual, the operator

$$\sum_{k=1}^{n} \varepsilon_k(\omega) P_{m_k} D'_k = \sum_{k=1}^{n} \varepsilon_k(\omega) \sum_{j=1}^{m_k} D_j \sum_{i=n_{k-1}+1}^{n_k} D_i = \sum_{k=1}^{n} \varepsilon_k(\omega) \sum_{j=n_{k-1}+1}^{m_k} D_j$$

acts as a multiplier operator, the multipliers $\varepsilon_k(\omega)$ of which are bounded in absolute value by 1 and constant on each block $\{n_{k-1}+1, \ldots, n_k\}$. Thus there is a $\lambda(\omega) \in \ell^\infty$ satisfying the conditions in part 1 such that $T_{\lambda(\omega)}$ coincides with the operator described above. By part 1, which is now assumed, these operators are uniformly bounded for all $\omega \in \Omega$; thus

$$\left|\sum_{k=1}^{n} \varepsilon_k(\omega) P_{m_k} D'_k x\right|_X \le M \left|\sum_{k=1}^{n} D'_k x\right|_X ,$$

with M independent of ω, and integrating over Ω yields

$$\left|\sum_{k=1}^{n} \varepsilon_k P_{m_k} D'_k x\right|_{L^p(\Omega;X)} \le M \left|\sum_{k=1}^{n} D'_k x\right|_X \le M C_p(D') \left|\sum_{k=1}^{n} \varepsilon_k D'_k x\right|_{L^p(\Omega;X)}$$

With $x := \sum_{k=1}^{n} x_k$, $x_k \in \operatorname{ran} D'_k$ arbitrary, the asserted relative R-boundedness follows.

2 $\Rightarrow$ 1. We start by manipulating a multiplier operator T_λ, initially defined in terms of the Schauder decomposition D, into a form, in which we can exploit the properties of the blocking D'. Recall that $D_j = P_j - P_{j-1}$, and observe that

$$\sum_{j=n_{k-1}+1}^{n_k} \lambda_j (P_j - P_{j-1}) + \sum_{j=n_{k-1}+1}^{n_k} (\lambda_j - \lambda_{j-1}) P_{j-1} = \sum_{j=n_{k-1}+1}^{n_k} (\lambda_j P_j - \lambda_{j-1} P_{j-1})$$
$$= \lambda_{n_k} P_{n_k} - \lambda_{n_{k-1}} P_{n_{k-1}}.$$

The previous equality obviously remains valid when multiplied from the right by D'_k, and for $n_{k-1} < j \le n_k$, we also have $D_j = D_j D'_k = (P_j - P_{j-1}) D'_k$. Thus

$$\sum_{j=n_{k-1}+1}^{n_k} \lambda_j D_j = \sum_{j=n_{k-1}+1}^{n_k} (\lambda_{j-1} - \lambda_j) P_{j-1} D'_k + \lambda_{n_k} P_{n_k} D'_k - \lambda_{n_{k-1}} P_{n_{k-1}} D'_k.$$

The terms containing $P_{n_{k-1}} D'_k$ vanish, since $D'_k = \sum_{j=n_{k-1}+1}^{n_k} D_j$ and $P_{n_{k-1}} = \sum_{j=1}^{n_{k-1}} D_j$. Furthermore, $P_{n_k} D'_k = D'_k$.

For $x \in \operatorname{ran} D = \operatorname{ran} D'$ we then have, with a finite sum,

$$\sum_{j=1}^{\infty} \lambda_j D_j x = \sum_{k=1}^{\infty} \sum_{j=n_{k-1}+1}^{n_k} \lambda_j D_j x = \sum_{k=1}^{\infty} \lambda_{n_k} D'_k x + \sum_{k=1}^{\infty} \sum_{j=n_{k-1}+1}^{n_k - 1} (\lambda_j - \lambda_{j+1}) P_j D'_k x.$$

The first term is bounded by $C\,|x|_X$ by the first abstract multiplier theorem, Corollary 2.14. For the second term, observe that $\sum_{j=n_{k-1}+1}^{n_k-1} (\lambda_j - \lambda_{j+1}) P_j \in \sum_{j=n_{k-1}+1}^{n_k-1} |\lambda_j - \lambda_{j+1}| \cdot \operatorname{abco}\{P_j\}_{j=n_{k-1}+1}^{n_k} \subset B(0;1) \operatorname{abco}\{P_j\}_{j=n_{k-1}+1}^{n_k}$ for λ as in condition 1. It then follows that

$$\left| \sum_{k=1}^{\infty} \left(\sum_{j=n_{k-1}+1}^{n_k-1} (\lambda_j - \lambda_{j+1}) P_j \right) D'_k x \right|_X \le C_p(D') \left| \sum_{k=1}^{\infty} \varepsilon_k \left(\sum_{j=n_{k-1}+1}^{n_k-1} (\lambda_j - \lambda_{j+1}) P_j \right) D'_k x \right|_{L^p(\Omega;X)}$$
$$\le C_p(D') \cdot 2R \left| \sum_{k=1}^{\infty} \varepsilon_k D'_k x \right|_{L^p(\Omega;X)} \le C_p(D') \cdot 2R \cdot C_p(D') \left| \sum_{k=1}^{\infty} D'_k x \right|_X = 2C_p^2(D') R\,|x|_X,$$

where the first and third inequalities exploited the unconditionality of D' via Lemma 3.2 and the relative R-boundedness in the assumption via Lemmas 4.10 and 4.12(3).

The implication is established. □

4.4 Notes and comments

This chapter is based on treatments in Clément et al. [3], Hieber and Prüss [8], and Witvliet [28]. We have also been inspired by the lectures of, and personal communication with Jan Prüss at Helsinki University of Technology in August 2000.

The notion of relative R-boundedness appears implicitly in [3] and in [28]; the introduction of this concept streamlines the statement of Theorem 4.14. Theorem 4.13 is slightly generalized from the result in [3] or in [8].

R-boundedness gives rise to other related notions which describe various analytic situations by means of randomization. For instance, the requirement of the R-boundedness of the family $\{t(t+A)^{-1}\}_{t>0}$ of resolvent operators, where A is a linear operator, is used in [8] to define an *R-sectorial* operator, a concept which turns out to be useful in the study of parabolic partial differential equations.

Chapter 5

Martingales

5.1 Introduction

Martingales constitute a particular class of stochastic processes, which has found applications in various fields of mathematics. In the present context, they are required to formulate the UMD-property of certain Banach spaces, which turns out to be equivalent to a number of other properties, each important by itself. The definition of martingales goes as follows:

Definition 5.1. *A* ***martingale*** *is a sequence $f = (f_k)_{k=1}^{\infty}$ of random variables (in probabilistic terms, a discrete parameter stochastic process) on a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, which is adapted to an increasing sequence $(\mathfrak{F}_k)_{k=1}^{\infty}$ of sub-σ-algebras of $\mathfrak{F}$, and the* ***difference sequence*** *of which, defined by $\delta f_k := f_k - f_{k-1}$ (with $f_0 := 0$), satisfies the condition $\mathbb{E}(\,\delta f_k | \,\mathfrak{F}_{k-1}) = 0$. By* ***adapted*** *we mean that each f_k is $\mathfrak{F}_k$-measurable.*

The definition of a martingale as given here is the same for vector-valued random variables as for the real case, once the auxiliary concepts appearing in it, in particular the conditional expectations $\mathbb{E}(\,\cdot | \,\mathfrak{F}_k)$ are given proper content. The purpose of the following section is to give insight into this matter.

It follows from the elementary properties of conditional expectation (which will be essentially the same in the vector-valued case) that the martingale condition for the difference sequence $\{\delta f_k\}_{k=1}^{\infty}$ can equivalently be stated as $f_{k-1} = \mathbb{E}(\,f_k | \,\mathfrak{F}_{k-1})$. If, in the real-valued case, the equality in this last equation is replaced by "$\leq$" or "$\geq$", for each k, then f is called a **submartingale** or a **supermartingale**, respectively.

5.2 Conditional expectation

The notion of conditional expectation appears in the very definition of martingales in Section 5.1 and it is one of the most fundamental and far reaching ideas in probability theory. We here give a meaning for this concept in the vector-valued setting. To be precise, we concentrate on separable Banach spaces, for which the integration theory developed in Section A.2 is valid, bearing in mind, however, the extension in Remark A.7.

It is easy to give a definition of the conditional expectation by means of what we want; showing that such an operator exists is more involved:

Definition 5.2. *For $f \in L^1(\mathfrak{F}; X)$ and a σ-algebra $\mathfrak{G} \subset \mathfrak{F}$, the* ***conditional expectation*** *of f with respect to $\mathfrak{G}$, denoted by $\mathbb{E}(\,f | \,\mathfrak{G})$, is a $g \in L^1(\mathfrak{G}; X)$ which satisfies*

$$\int_G g d\mathbb{P} = \int_G f d\mathbb{P} \tag{5.1}$$

for all $G \in \mathfrak{G}$.

Remark 5.3. *An application of Corollary A.5 to $g := g_1 - g_2 \in L^1(\mathfrak{G}; X)$, where g_1, g_2 are two functions satisfying the definition of $\mathbb{E}(\, f | \, \mathfrak{G})$, immediately shows that that $\mathbb{E}(\, f | \, \mathfrak{G})$ is essentially unique.*

In the scalar-valued case the existence of g is a direct consequence of the Radon–Nikodým theorem: $\mu(G) := \int_G f d\mathbb{P}$ defines a measure on $\mathfrak{G}$ which is absolutely continuous with respect to the restriction $\mathbb{P}|_{\mathfrak{G}}$ of $\mathbb{P}$ on $\mathfrak{G}$: $\mu \ll \mathbb{P}|_{\mathfrak{G}}$. It then suffices to take $g := \frac{d\mu}{d\mathbb{P}|_{\mathfrak{G}}}$. (Observe that it is essential in this last expression to have $\mathbb{P}|_{\mathfrak{G}}$ rather than $\mathbb{P}$, since $\frac{d\mu}{d\mathbb{P}} = f$ a.s., and f need not be $\mathfrak{G}$-measurable. See also Section 5.4 for a discussion on an alternative proof.)

As direct a procedure will not work in the vector-valued setting; however, a construction of the conditional expectation essentially similar to that of the integral is possible, starting from the simple functions.

Lemma 5.4. *For $f \in L^1(\mathfrak{F}; X)$ and a σ-algebra $\mathfrak{G} \subset \mathfrak{F}$, the conditional expectation $\mathbb{E}(\, f | \, \mathfrak{G})$ exists and satisfies $|\mathbb{E}(\, f | \, \mathfrak{G})|_X \le \mathbb{E}(\, |f|_X | \, \mathfrak{G})$ (a.s.). The operator $\mathbb{E}(\, \cdot | \, \mathfrak{G})$ is a contractive projector of $L^1(\mathfrak{F}; X)$ onto $L^1(\mathfrak{G}; X)$.*

Note that the right-hand side of the last inequality above is a conditional expectation of the scalar (in fact, positive) random variable $|f|_X$.

Proof. If f is a simple function $f = \sum_{k=1}^n x_k \mathbf{1}_{E_k}$, we can take

$$g := \mathbb{E}(\, f | \, \mathfrak{G}) = \sum_{k=1}^n x_k \mathbb{E}(\, \mathbf{1}_{E_k} | \, \mathfrak{G}),$$

since clearly

$$\int_G g d\mathbb{P} = \sum_{k=1}^n x_k \int_G \mathbb{E}(\, \mathbf{1}_{E_k} | \, \mathfrak{G}) \, d\mathbb{P} = \sum_{k=1}^n x_k \int_G \mathbf{1}_{E_k} d\mathbb{P} = \sum_{k=1}^n x_k \mathbb{P}(G \cap E_k) = \int_G f d\mathbb{P}.$$

(Observe that the conditional expectations appearing in the above equation only involve real random variables, for which the existence was already demonstrated. Also note that the first equality is *not* the definition of the integral for simple functions, since g need not be simple (as $\mathbb{E}(\, \mathbf{1}_{E_k} | \, \mathfrak{G})$ need not be an indicator, or even a finite sum of indicators). However, if the $\mathbb{E}(\, \mathbf{1}_{E_k} | \, \mathfrak{G})$ were such finite sums, then the equality would be a matter of definition, and the general case is obtained by the continuity of $\mathbb{E} = \int_\Omega \cdot d\mathbb{P}$ and the density of simple functions, as in Section A.2.)

We further compute (assuming E_k disjoint)

$$\int_\Omega \left| \sum_{k=1}^n x_k \mathbb{E}\left(\mathbf{1}_{E_k} | \, \mathfrak{G}\right) \right|_X d\mathbb{P} \le \sum_{k=1}^n |x_k|_X \int_\Omega \left|\mathbb{E}\left(\mathbf{1}_{E_k} | \, \mathfrak{G}\right)\right| d\mathbb{P} = \sum_{k=1}^n |x_k|_X \, \mathbb{P}(E_k) = \int_\Omega \left| \sum_{k=1}^n x_k \mathbf{1}_{E_k} \right|_X d\mathbb{P},$$

and this says that $|\mathbb{E}(\, f | \, \mathfrak{G})|_{L^1(\mathfrak{G};X)} \le |f|_{L^1(\mathfrak{F};X)}$ for simple f, by continuity and density for all $f \in L^1(\mathfrak{F}; X)$. (In the second step we implicitly removed the absolute value signs, since the conditional expectation of a non-negative random variable is a.s. non-negative.)

Thus $\mathbb{E}(\, \cdot | \, \mathfrak{G})$ is a bounded (obviously linear) operator from $L^1(\mathfrak{F}; X)$ to $L^1(\mathfrak{G}; X)$ when restricted to simple measurable functions, whence we can uniquely extend it to a bounded linear operator on all of $L^1(\mathfrak{F}; X)$. It is also clear that $\mathbb{E}(\, \cdot | \, \mathfrak{G})$ is the identity (a.s.) when restricted to $L^1(\mathfrak{G}; X)) = \operatorname{ran}(\mathbb{E}(\, \cdot | \, \mathfrak{F}))$. Thus the conditional expectation has the projection property $\mathbb{E}(\, \mathbb{E}(\, \cdot | \, \mathfrak{G}) | \, \mathfrak{G}) = \mathbb{E}(\, \cdot | \, \mathfrak{G})$, and the fact that it is contractive follows by density and continuity from the norm inequality established above for simple random variables. □

Now that we have the conditional expectations, we need some tools to work with them. The first lemma merely reminds us of some results in the scalar case.

Lemma 5.5. *Let $\{f_n\}_{n=1}^\infty \subset L^1(\mathfrak{F})$ and $\mathfrak{G} \subset \mathfrak{F}$ be a sub-σ-algebra. Then the following convergence results hold:*

1. *(Monotone convergence.) If* $0 \le f_n \uparrow f$ *(a.s.),then* $0 \le \mathbb{E}(f_n|\mathfrak{G}) \uparrow \mathbb{E}(f|\mathfrak{G})$ *(a.s.).*

2. *(Fatou's lemma.) If* $0 \le f_n$ *(a.s.), then* $\mathbb{E}(\liminf_{n\to\infty} f_n|\mathfrak{G}) \le \liminf_{n\to\infty} \mathbb{E}(f_n|\mathfrak{G})$ *(a.s.).*

3. *(Dominated convergence.) If* $|f_n| \le g \in L^1(\mathfrak{F})$ *and* $f_n \to f$ *(a.s.), then we have the convergence* $\mathbb{E}(|f_n - f||\mathfrak{G}) \to 0$ *(a.s.) and consequently* $\mathbb{E}(f_n|\mathfrak{G}) \to \mathbb{E}(f|\mathfrak{G})$ *(a.s.).*

Proof. Assume the hypotheses of the monotone convergence theorem. Since we have, by definition, $\int_G \mathbb{E}(f_n|\mathfrak{G})\, d\mathbb{P} = \int_G f_n d\mathbb{P} \ge 0$ for each $G \in \mathfrak{G}$, we know from real analysis that $\mathbb{E}(f_n|\mathfrak{G}) \ge 0$ (a.s.). The same argument with $f_{n+1} - f_n \ge 0$ in place of f_n shows that $\mathbb{E}(f_{n+1}|\mathfrak{G}) \ge \mathbb{E}(f_n|\mathfrak{G})$ (a.s.). Thus there exists $g := \lim_{n\to\infty} \uparrow \mathbb{E}(f_n|\mathfrak{G})$ (a.s.), and the "ordinary" theorem of monotone convergence shows that

$$\int_G g d\mathbb{P} = \lim_{n\to\infty} \uparrow \int_G \mathbb{E}(f_n|\mathfrak{G})\, d\mathbb{P} = \lim_{n\to\infty} \uparrow \int_G f_n d\mathbb{P} = \int_G f d\mathbb{P}$$

for all $G \in \mathfrak{G}$. Thus $g = \mathbb{E}(f|\mathfrak{G})$ (a.s.), and g satisfies the properties asserted.

Fatou's lemma and the dominated convergence theorem now follow as in real analysis [20, 27], with only notational modifications. ☐

Remark 5.6. *The dominated convergence theorem can immediately be extended to Banach spaces: If* $|f_n|_X \le g \in L^1(\mathfrak{F})$ *and* $f_n \to f$ *(a.s.), then Lemma 5.5(3) applied to the scalar-valued functions* $|f_n - f|_X$ *yields* $\mathbb{E}(|f_n - f|_X|\mathfrak{G}) \to 0$ *(a.s.), and thus* $\mathbb{E}(f_n|\mathfrak{G}) \to \mathbb{E}(f|\mathfrak{G})$ *(a.s.).*

The other two convergence results as such have no meaning in a general Banach space, but of course they can be useful in estimating norms of vector-valued random variables.

Next we present a version of Jensen's inequality for vector-valued integrals.

Lemma 5.7 (Jensen's inequality). *For a random variable* $f \in L^1(\mathfrak{F}; X)$*, a* σ*-algebra* $\mathfrak{G} \subset \mathfrak{F}$ *and a continuous convex mapping* $\phi : X \to \mathbb{R}$ *for which* $\phi \circ f \in L^1(\mathfrak{F})$*, we have the inequality (a.s.)*

$$\phi \circ \mathbb{E}(f|\mathfrak{G}) \le \mathbb{E}(\phi \circ f|\mathfrak{G}).$$

Proof. First consider a simple random variable $f = \sum_{k=1}^n x_k \mathbf{1}_{E_k}$, where the measurable sets E_k are chosen in a canonical way so that they are pairwise disjoint and $\cup_{k=1}^n E_k = \Omega$ (possibly with some $x_k = 0$). When this is the case, we have $\sum_{k=1}^n \mathbf{1}_{E_k}(\omega) = 1$ for all $\omega \in \Omega$, and consequently, $\sum_{k=1}^n \mathbb{E}(\mathbf{1}_{E_k}|\mathfrak{G}) = \mathbb{E}(\sum_{k=1}^n \mathbf{1}_{E_k}|\mathfrak{G}) = \mathbb{E}(1|\mathfrak{G}) = 1$ (a.s.). Furthermore, we have $\mathbb{E}(\mathbf{1}_{E_k}|\mathfrak{G}) \ge 0$ (a.s.), since $\mathbf{1}_{E_k} \ge 0$.

We have now shown that $\{\mathbb{E}(\mathbf{1}_{E_k}|\mathfrak{G})(\omega)\}_{k=1}^n$ is a proper set of coefficients for a convex combination, for almost all $\omega \in \Omega$. The desired inequality now follows from the familiar form of Jensen's inequality (which can be proved by induction from the very definition of a convex function):

$$\begin{aligned}\phi(\mathbb{E}(f|\mathfrak{G})(\omega)) = \phi(\sum_{k=1}^n x_k \mathbb{E}(\mathbf{1}_{E_k}|\mathfrak{G})(\omega)) &\le \sum_{k=1}^n \phi(x_k)\mathbb{E}(\mathbf{1}_{E_k}|\mathfrak{G})(\omega) \\ &= \mathbb{E}\left(\sum_{k=1}^n \phi(x_k)\mathbf{1}_{E_k} \middle| \mathfrak{G}\right)(\omega) = \mathbb{E}(\phi \circ f|\mathfrak{G})(\omega).\end{aligned}$$

Now we have the inequality for simple f. For general $f \in L^1(\mathfrak{F}; X)$, we first assume that ϕ, in addition to the assumptions of the lemma, attains a minimum. Then Lemmas A.1 and A.3 provide us with a sequence $\{f_k\}_{k=1}^\infty$ of simple random variables so that $f_k \to f$ pointwise and in $L^1(\mathfrak{F}; X)$, and $\phi(f_k(\omega)) \le \phi(f(\omega)) + \frac{1}{k}$.

Since $\mathbb{E}(\cdot|\mathfrak{G})$ is continuous from $L^1(\mathfrak{F}; X)$ to $L^1(\mathfrak{G}; X)$, the convergence $f_k \to f$ in $L^1(\mathfrak{F}; X)$ implies $\mathbb{E}(f_k|\mathfrak{G}) \to \mathbb{E}(f|\mathfrak{G})$ in $L^1(\mathfrak{G}; X)$, and a subsequence (which we will hereafter concentrate on and still denote by $\{f_k\}_{k=1}^\infty$) converges almost surely. Since ϕ is continuous, we further have $\phi \circ f_k \to \phi \circ f$ (a.s.) and $\phi \circ \mathbb{E}(f_k|\mathfrak{G}) \to \phi \circ \mathbb{E}(f|\mathfrak{G})$ (a.s.).

We then obtain (a.s.)

$$\phi \circ \mathbb{E}(f|\mathfrak{G}) = \lim_{k\to\infty} \phi \circ \mathbb{E}(f_k|\mathfrak{G}) \le \limsup_{k\to\infty} \mathbb{E}(\phi \circ f_k|\mathfrak{G}) \le \limsup_{k\to\infty} \mathbb{E}\left(\phi \circ f + \frac{1}{k}\middle|\mathfrak{G}\right) = \mathbb{E}(\phi \circ f|\mathfrak{G}),$$

where the first inequality was simply Jensen for simple functions (i.e., the first part of the proof) and the second follows from the construction of f_k (and the monotonicity of $\mathbb{E}(\cdot|\mathfrak{G})$, Lemma 5.5(1)). We have now proved the lemma with the additional assumption that ϕ attains a minimum.

We should now get rid of the extra condition on ϕ. If ϕ is any continuous and convex function, so is $\phi \vee t$, $t \in \mathbb{R}$, but this latter one also attains a minimum if we choose $t > \inf \phi$. We then pick a sequence $\{t_n\}_{n=1}^\infty$ of such values so that $t_n \downarrow \inf \phi$ (whether or not $\inf \phi$ is finite). We know from the previous part of the proof that $\phi \vee t_n$ satisfies Jensen's inequality, and thus

$$\phi \circ \mathbb{E}(f|\mathfrak{G}) \le (\phi \vee t_n) \circ \mathbb{E}(f|\mathfrak{G}) \le \mathbb{E}((\phi \vee t_n) \circ f|\mathfrak{G}) \tag{5.2}$$

If $\inf \phi > -\infty$, then we write the right-hand side of (5.2) in the form

$$\mathbb{E}(\phi \circ f|\mathfrak{G}) + \mathbb{E}\left((t_n - \phi)\mathbf{1}_{\{\phi < t_n\}}\middle|\mathfrak{G}\right),$$

which follows from $\phi = \phi \vee t_n + (\phi - t_n)\mathbf{1}_{\{\phi < t_n\}}$ and the linearity of $\mathbb{E}(\cdot|\mathfrak{G})$.

Now $\left|(t_n - \phi)\mathbf{1}_{\{\phi < t_n\}}\right| \le t_1 - \inf \phi$, and $(t_n - \phi)\mathbf{1}_{\{\phi < t_n\}} \to 0$ as $n \to \infty$ pointwise, so that it follows from the dominated convergence theorem that the last term on the right-hand side of (5.2) tends to zero (a.s.) as $n \to \infty$. Since (5.2) holds for all $n \in \mathbb{Z}_+$ and the left-hand side is independent of n, the inequality must also hold in the limit.

Otherwise, $\inf \phi = -\infty$, and we can take our sequence to be $t_n = -n$. If ϕ^+ and ϕ^- are the positive and negative parts of ϕ, then $\phi \vee (-n) = \phi^+ - \phi^- \wedge n$ and we can write the right-hand side of (5.2) as

$$\mathbb{E}\left(\phi^+ \circ f\middle|\mathfrak{G}\right) - \mathbb{E}\left((\phi^- \wedge n) \circ f\middle|\mathfrak{G}\right).$$

As $n \uparrow \infty$, we have $\phi^- \wedge n \uparrow \phi^-$ pointwise; thus it follows from the monotone convergence theorem that the right-hand side of (5.2) tends to $\mathbb{E}(\phi \circ f|\mathfrak{G})$ (a.s.) as $n \uparrow \infty$. Hence the things are settled as asserted also in this case.

Now the proof is complete for all f and ϕ as in the assertion. ☐

Remark 5.8. *The analogous results for "ordinary" expectations (i.e., integrals) $\mathbb{E}$ follow from the corresponding results for conditional expectations by taking $\mathfrak{G} := \{\emptyset, \Omega\}$ (the trivial σ-algebra). Indeed, all $\{\emptyset, \Omega\}$-measurable functions are constants, so $\mathbb{E}(f|\mathfrak{G}) = \mathbb{E}f$, surely (as opposed to almost surely).*

Some corollaries of Jensen's inequality are immediate.

Corollary 5.9. *For σ-algebras $\mathfrak{G} \subset \mathfrak{F}$, the conditional expectation $\mathbb{E}(\cdot|\mathfrak{G})$ is a contractive projection of $L^p(\mathfrak{F}; X)$ onto $L^p(\mathfrak{G}; X)$, $1 \le p \le \infty$.*

Proof. For $p \in [1, \infty)$, the mapping $\phi := |\cdot|_X^p : X \to \mathbb{R}$ is obviously continuous and convex, permitting the computation

$$|\mathbb{E}(f|\mathfrak{G})|_{L^p(\mathfrak{G};X)}^p = \mathbb{E}(|\mathbb{E}(f|\mathfrak{G})|_X^p) \le \mathbb{E}(\mathbb{E}(|f|_X^p|\mathfrak{G})) = \mathbb{E}(|f|_X^p) = |f|_{L^p(\mathfrak{F};X)}^p.$$

For $p = \infty$, use Jensen's inequality to give $|\mathbb{E}(f|\mathfrak{G})|_X \le \mathbb{E}(|f|_X|\mathfrak{G})$. Denoting $g := \mathbb{E}(f|\mathfrak{G})$, take $G := \{|g|_X \ge |f|_{L^\infty(\mathfrak{F};X)} + \frac{1}{n}\} \in \mathfrak{G}$. If $\mathbb{P}(G) > 0$, then $\int_G |g|_X \, d\mathbb{P} > \int_G |f|_X \, d\mathbb{P}$, which contradicts $|g|_X \le \mathbb{E}(|f|_X|\mathfrak{G})$. Letting $n \downarrow 0$, we deduce that $|g|_{L^\infty(\mathfrak{G})} \le |f|_{L^\infty(\mathfrak{F})}$. ☐

Example 5.10. *If $f \in L^p(\Omega; X)^{\mathbb{Z}_+}$ is a martingale, then the difference sequence $\{\delta f_k\}_{k=1}^\infty$ is a monotone basic sequence on $L^p(\Omega; X)$, $1 \le p \le \infty$.*

Proof. Recall that basic sequences and monotonicity where defined following the characterization in Corollary 2.10. Now, for any integers $m \geq n \geq 1$ and any scalars a_k, $k = 1, \ldots, m$,

$$\sum_{k=1}^{n} a_k \delta f_k = \mathbb{E}\left(\sum_{k=1}^{m} a_k \delta f_k \middle| \mathfrak{F}_n \right)$$

and the claim follows from the contractivity of $\mathbb{E}(\,\cdot|\,\mathfrak{F}_n)$ on $L^p(\Omega; X)$. ☐

Corollary 5.11. *If $f = \{f_k\}_{k=1}^{\infty} \in L^1(\Omega; X)^{\mathbb{Z}_+}$ is a martingale, and $\phi : X \to \mathbb{R}$ is continuous and convex, and $\phi \circ f_k$ is integrable for $k \in \mathbb{Z}_+$, then $\phi \circ f := \{\phi \circ f_k\}_{k=1}^{\infty} \in L^1(\Omega)^{\mathbb{Z}_+}$ is a submartingale.*

In particular, taking $\phi = |\cdot|_X$ we see that $\{|f_k(\cdot)|_X\}_{k=1}^{\infty}$ is a non-negative submartingale, whenever f is a martingale.

Proof. Clearly $\phi \circ f_k$ is $\mathfrak{F}_k$-measurable, if f_k is, and $\mathbb{E}(\,\phi \circ f_k|\,\mathfrak{F}_{k-1}) \geq \phi \circ \mathbb{E}(\,f_k|\,\mathfrak{F}_{k-1}) = \phi \circ f_{k-1}$. ☐

The following fundamental property of the conditional expectation is of significant value:

Lemma 5.12. *If $g \in L^p(\mathfrak{G}; \mathcal{B}(X;Y))$, $p \in [1, \infty)$, $f \in L^{\overline{p}}(\mathfrak{F}; X)$, $\mathfrak{G} \subset \mathfrak{F}$, then $\mathbb{E}(\,gf|\,\mathfrak{G}) = g\mathbb{E}(\,f|\,\mathfrak{G})$.*

Observe a number of important special cases: If g is constant, then it is certainly in any L^p on a probability space, so the conditions of the lemma are satisfied. Clearly a scalar-valued g can be interpreted as $\mathcal{B}(X)$-valued by the obvious identification of λ and $\lambda\,\mathrm{id}$. With $Y = \mathbb{C}$, we have a result concerning $g \in L^p(\mathfrak{G}; X^*)$ and $f \in L^{\overline{p}}(\mathfrak{G}; X)$. We can also revert the roles of g and f, since X can always be identified with a subset of $\mathcal{B}(\mathcal{B}(X;Y);Y)$.

Proof. Since g and $\mathbb{E}(\,f|\,\mathfrak{G})$ are $\mathfrak{G}$-measurable, they are limits a.s. of simple $\mathfrak{G}$-measurable functions s_k and t_k, and it is easy to verify that the functions $\omega \mapsto s_k(\omega)t_k(\omega)$ converge a.s. to $\omega \mapsto g(\omega)f(\omega)$, thus this last function is also $\mathfrak{G}$-measurable.

If g is simple, say $g = \sum_{k=1}^{n} A_k \mathbf{1}_{E_k}$, $A_k \in \mathcal{B}(X;Y)$, $E_k \in \mathfrak{G}$ then, for arbitrary $G \in \mathfrak{G}$,

$$\int_G g\mathbb{E}(\,f|\,\mathfrak{G})\,d\mathbb{P} = \sum_{k=1}^{n} A_k \int_{G\cap E_k} \mathbb{E}(\,f|\,\mathfrak{G})\,d\mathbb{P} = \sum_{k=1}^{n} A_k \int_{G\cap E_k} f d\mathbb{P} = \int_G gf d\mathbb{P}.$$

For arbitrary $g \in L^p$, $p < \infty$ take a sequence $\{s_k\}_{k=1}^{\infty} \subset S(\mathfrak{G}; X)$ converging to g a.s. and in L^p, with $|s_k - g|_{\mathcal{B}(X;Y)} \leq |g|_{\mathcal{B}(X;Y)}$ a.s. (Lemma A.3). Then

$$\left| \int_G (s_k - g) f d\mathbb{P} \right|_Y \leq \int_G |s_k - g|_{\mathcal{B}(X;Y)}\, |f|_X \, d\mathbb{P} \leq |s_k - g|_{L^p(\Omega;\mathcal{B}(X;Y))}\, |f|_{L^{\overline{p}}(\Omega;X)} \underset{k\to\infty}{\longrightarrow} 0,$$

and a similar estimate holds with $\mathbb{E}(\,f|\,\mathfrak{G})$ in place of f. Thus the equality $\int_G g\mathbb{E}(\,f|\,\mathfrak{G})\,d\mathbb{P} = \int_G gf d\mathbb{P}$ holds for all $G \in \mathfrak{G}$, and it follows that the $\mathfrak{G}$-measurable function $g\mathbb{E}(\,f|\,\mathfrak{G})$ is the conditional expectation of gf by definition. ☐

Corollary 5.13. *If $f \in L^p(\Omega; X)^{\mathbb{Z}_+}$ is a martingale on X, and $\Lambda \in \mathcal{B}(X;Y)$, then $\Lambda f := (\Lambda f_k)_{k=1}^{\infty} \in L^p(\Omega; Y)^{\mathbb{Z}_+}$ is a martingale on Y.*

Proof. By Lemma 5.12, we have $\mathbb{E}(\,\Lambda f_n|\,\mathfrak{F}_{n-1}) = \Lambda\mathbb{E}(\,f_n|\,\mathfrak{F}_{n-1}) = \Lambda f_{n-1}$. ☐

In a number of special cases, the conditional expectation has an explicit formula, which is sometimes useful. A particularly simple case occurs when the σ-algebra $\mathfrak{G}$ is finite; then there is a unique collection $\mathrm{bs}\,\mathfrak{G}$, the basis of $\mathfrak{G}$, of $G \in \mathfrak{G} \setminus \{\emptyset\}$ such that no proper subset of any such G is in $\mathfrak{G}$. Then

$$\mathbb{E}(\,f|\,\mathfrak{G})\,(\omega) = \sum_{G\in \mathrm{bs}\,\mathfrak{G}} \mathbf{1}_G(\omega) \frac{1}{\mathbb{P}(G)} \int_G f(\omega')d\mathbb{P}(\omega') = \int_\Omega f(\omega') \sum_{G\in \mathrm{bs}\,\mathfrak{G}} \frac{1}{\mathbb{P}(G)} \mathbf{1}_{G^2}(\omega, \omega')d\mathbb{P}(\omega'), \quad (5.3)$$

i.e., $\mathbb{E}(\,\cdot\,|\,\mathfrak{G})$ is an integral operator with symmetric kernel. (The first equality above can easily be verified by checking that the function on the right of this equality satisfies the properties required of $\mathbb{E}(\,f|\,\mathfrak{G})$.)

We conclude this section with a simple geometric characterization of the conditional expectation of an L^2 random variable on a Hilbert space.

Lemma 5.14. *If $\mathcal{H}$ is a Hilbert space, $f \in L^2(\mathfrak{F};\mathcal{H})$ and $\mathfrak{G} \subset \mathfrak{F}$ is a σ-algebra, then $\mathbb{E}(\,f|\,\mathfrak{G})$ is the orthogonal projection of f onto $L^2(\mathfrak{G};\mathcal{H})$, i.e., $f - \mathbb{E}(\,f|\,\mathfrak{G}) \perp L^2(\mathfrak{G};\mathcal{H})$.*

$L^2(\mathfrak{F};\mathcal{H})$ has implicitly been endowed with the inner product

$$(f,g)_{L^2(\mathfrak{F};\mathcal{H})} := \int_\Omega (f(\omega), g(\omega))_{\mathcal{H}}\, d\mathbb{P}(\omega),$$

which makes it a Hilbert space, too.

Proof. Since $L^2(\mathfrak{G};\mathcal{H})$ is complete, it is in particular a closed subspace of $L^2(\mathfrak{F};\mathcal{H})$. Thus there exists an (essentially unique) orthogonal projection of $f \in L^2(\mathfrak{F};\mathcal{H})$ onto $L^2(\mathfrak{G};\mathcal{H})$; we denote it by g.

Now $f - g \perp L^2(\mathfrak{G};\mathcal{H})$; in particular, $f - g \perp x\mathbf{1}_G$ for all $x \in \mathcal{H}$ and $G \in \mathfrak{G}$. Thus

$$0 = (f-g, x\mathbf{1}_G)_{L^2(\mathfrak{G};\mathcal{H})} = \int_G (f(\omega) - g(\omega), x)_{\mathcal{H}}\, d\mathbb{P}(\omega) = \left(\int_G f(\omega) - g(\omega) d\mathbb{P}(\omega), x\right)_{\mathcal{H}}.$$

Since this holds for all $x \in \mathcal{H}$, we conclude that $\int_G (f-g)d\mathbb{P} = 0$ for each $G \in \mathfrak{G}$, and this means that $g = \mathbb{E}(\,f|\,\mathfrak{G})$ (a.s.). □

Corollary 5.15. *If $\mathcal{H}$ is a Hilbert space and $f = \{f_k\}_{k=1}^\infty \in L^2(\Omega;\mathcal{H})^{\mathbb{Z}_+}$ is a martingale adapted to $\{\mathfrak{F}_k\}_{k=1}^\infty$, then the differences δf_k are orthogonal.*

Proof. By Lemma 5.14, $\delta f_k = f_k - f_{k-1} = f_k - \mathbb{E}(\,f_k|\,\mathfrak{F}_{k-1}) \perp L^2(\mathfrak{F}_{k-1};X) \ni \delta f_j$ for $j < k$. □

5.3 Maximal operator and Doob's inequalities

Now that Definition 5.1 of martingales makes sense also in the vector-valued setting, we can explore some further properties of these objects. We are particularly interested in martingales $f = \{f_k\}_{k=1}^\infty$, for which the norm

$$|f|_{\ell^\infty(\mathbb{Z}_+;L^p(\Omega;X))} := \sup_{k\in\mathbb{Z}_+} |f_k|_{L^p(\Omega;X)}$$

is finite. Such martingales are said to be **bounded in** L^p, and they are conveniently characterized by the maximal function, to be defined next.

Definition 5.16. *The* ***maximal operator*** $(\cdot)^*$ *is defined, for martingales $f = \{f_k\}_{k=1}^\infty \in L^1(\Omega;X)^{\mathbb{Z}_+}$, by $f^*(\omega) := \sup_{n\in\mathbb{Z}_+} |f_n(\omega)|_X$. The function f^* is called the* ***maximal function*** *of the martingale f.*

Observe that $(\cdot)^*$ takes $L^1(\Omega;X)^{\mathbb{Z}_+}$ to $[0,\infty]^\Omega$, whatever the X. The maximal operator on martingales possesses characteristics similar to those of other operators in analysis bearing the same name. In martingale theory, the results characterizing this operator are known as Doob's inequalities, the first one of which is given below.

Lemma 5.17 (Doob's L^1 inequality). *The following inequalities hold:*

1. *If $g = \{g_k\}_{k=1}^\infty \in L^1(\Omega)^{\mathbb{Z}_+}$ is a non-negative submartingale adapted to $\{\mathfrak{G}_k\}_{k=1}^\infty$, then*

$$t\mathbb{P}\left(\max_{k\le n} g_k \ge t\right) \le \int_{\{\max_{k\le n} g_k \ge t\}} g_n(\omega) d\mathbb{P}(\omega) \le |g_n|_{L^1(\Omega)}. \tag{5.4}$$

for each $t > 0$.

2. *If $f = \{f_k\}_{k=1}^\infty \in L^1(\Omega;X)^{\mathbb{Z}_+}$ is a martingale, then (5.4) holds with $|f_k(\omega)|_X$ in place of $g_k(\omega)$.*
3. *If $f \in \ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))$ is a martingale, then*
$$t\mathbb{P}(f^* \geq t) \leq |f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}\,.$$

Proof. The sets $G_k := \{\omega \in \Omega : \max_{j<k} g_j(\omega) < t \leq g_k(\omega)\} \in \mathfrak{G}_k$ are disjoint, and $\{\max_{k\leq n} g_k \geq t\} = \cup_{k=1}^n G_k$. Furthermore, $t\mathbb{P}(G_k) \leq \int_{G_k} g_k d\mathbb{P} \leq \int_{G_k} g_n d\mathbb{P}$, since $g_k \leq \mathbb{E}(\,g_n|\,\mathfrak{G}_k)$ for $k < n$ follows by iteration from the definition of submartingale, which states this inequality for $k = n-1$. The assertion 1 follows after summing these inequalities for $k = 1,\dots,n$.

The assertion 2 is a direct consequence, since $\{|f_k(\cdot)|_X\}_{k=1}^\infty \in L^1(\Omega)^{\mathbb{Z}_+}$ is a nonnegative submartingale by Corollary 5.11.

For assertion 3, we use part 2 to deduce $t\mathbb{P}(\sup_{k\leq n} \geq t) \leq |f_n|_{L^1(\Omega;X)} \leq |f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}$, and then take the limit as $n \to \infty$. $\square$

Lemma 5.17(3) is a familiar weak-type inequality for the maximal operator. One might guess that, for $p > 1$, we should have the corresponding strong type inequality, i.e., that the maximal operator is bounded from $\ell^\infty(\mathbb{Z}_+;L^p(\Omega;X))$ to $L^p(\Omega)$. This is indeed the case. We require one preliminary lemma.

Lemma 5.18. *If the random variables $f, g \geq 0$ satisfy $t\mathbb{P}(f \geq t) \leq \int_{f\geq t} g d\mathbb{P}$ for all $t > 0$, then $|f|_{L^p(\Omega)} \leq \overline{p}\,|g|_{L^p(\Omega)}$ for all $p \in (1,\infty]$.*

Proof. Take first $p < \infty$. This is a straightforward computation using distribution functions, with one application of Hölder's inequality.

$$\begin{aligned}
|f|^p_{L^p(\Omega;X)} &= \int_0^\infty pt^{p-1}\mathbb{P}(f \geq t)dt \leq \int_0^\infty pt^{p-2}\left(\int_{f\geq t} g(\omega)d\mathbb{P}(\omega)\right)dt \\
&= \int_\Omega d\mathbb{P}(\omega)\int_0^{f(\omega)} pt^{p-2}g(\omega)dt = \int_\Omega \frac{p}{p-1} f^{p-1}(\omega)g(\omega)d\mathbb{P}(\omega) \leq \overline{p}\left(\mathbb{E}f^{(p-1)\overline{p}}\right)^{\frac{1}{\overline{p}}}\left(\mathbb{E}g^p\right)^{\frac{1}{p}} \\
&= \overline{p}\,|f|^{\frac{1}{\overline{p}}}_{L^p(\Omega)}\,|g|^{\frac{1}{p}}_{L^p(\Omega)}\,.
\end{aligned}$$

The identity $(p-1)\overline{p} = p$ was occasionally used, and of course Fubini's theorem to change the order of integration, all integrands being non-negative. If $|f|_{L^p(\Omega)} < \infty$, the assertion follows after division by this quantity raised to the power of $\frac{1}{\overline{p}}$. Otherwise, the assertion is true for $f \wedge n \in L^p(\Omega)$, and the conclusion follows after taking the limit $n \to \infty$ and applying the monotone convergence theorem.

For $p = \infty$, assume $|f|_{L^\infty(\Omega)} > t$, in which case $\mathbb{P}(f \geq t) > 0$. If $s := |g|_{L^\infty(\Omega)} < t$, then $\int_{\{f\geq t\}} g d\mathbb{P} \leq s\mathbb{P}(f \geq t) < t\mathbb{P}(f \geq t)$, a contradiction. If $|f|_{L^\infty(\Omega)} < \infty$, the previous argument shows that $|g|_{L^\infty(\Omega)} \geq |f|_{L^\infty(\Omega)} - \epsilon$ for all $\epsilon > 0$, whence $|g|_{L^\infty(\Omega)} \geq |f|_{L^\infty(\Omega)}$. If $|f|_{L^\infty(\Omega)} = \infty$, we deduce $|g|_{L^\infty(\Omega)} \geq t$ for all t, thus $|g|_{L^\infty(\Omega)} = \infty$, too. $\square$

Lemma 5.19 (Doob's L^p inequality). *If $f \in \ell^\infty(\mathbb{Z}_+;L^p(\Omega;X))$, $1 < p < \infty$, is a martingale, or a non-negative submartingale with $X = \mathbb{R}$, then*

$$|f|_{\ell^\infty(\mathbb{Z}_+;L^p(\Omega;X))} \leq |f^*|_{L^p(\Omega)} \leq \overline{p}\,|f|_{\ell^\infty(\mathbb{Z}_+;L^p(\Omega;X))}$$

The first inequality holds also for $p = 1$.

Proof. The first inequality follows directly by integrating both sides of $|f_k(\omega)|_X \leq f^*(\omega)$ in the power of p over Ω and taking the supremum on the left-hand side. For the second inequality, let first f be a non-negative sub-martingale. By Lemma 5.17(1), $t\mathbb{P}(\max_{k\leq n} f_k \geq t) \leq \int_{\{\max_{k\leq n} f_k\geq t\}} f_n d\mathbb{P}$. It then follows from Lemma 5.18 that $|\max_{k\leq n} f_k|_{L^p(\Omega)} \leq \overline{p}\,|f_n|_{L^p(\Omega)} \leq \overline{p}\,|f|_{\ell^\infty(\mathbb{Z}_+;L^p(\Omega))}$, and the assertion follows by taking the limit as $n \to \infty$. The assertion for $f \in \ell^\infty(\mathbb{Z}_+;L^p(\Omega;X))$ follows from the result for non-negative sub-martingales applied to $\{|f_k(\cdot)|_X\}_{k=1}^\infty$, recalling Corollary 5.11. $\square$

5.4 Notes and comments

The construction of the conditional expectation follows Neveu [16]; Doob's inequalities are taken from Williams [27].

The proof of the vector-valued Jensen's inequality here is possibly new. Less powerful machinery can be used to establish its corollaries, see e.g. Diestel and Uhl [6], page 123. The proof differs from the hyperplane argument usually used in proving the scalar-version of this inequality. A main ingredient of our reasoning is the approximation result in Lemma A.1, which is motivated by a similar result in Neveu [16].

Lemma 5.14 is essentially an adaptation of the scalar-valued result in Williams [27]. There the construction of $\mathbb{E}(f|\mathfrak{G})$ for $f \in L^2(\mathfrak{F})$ as the orthogonal projection is exploited to show the existence of the conditional expectation without relying on the Radon–Nikodým theorem. (Williams actually proves the Radon–Nikodým theorem using the conditional expectations via martingales.) Of course, a further approximation argument is required for $f \in L^1(\mathfrak{F})$.

In our (or Neveu's [16]) proof of Lemma 5.4 we only applied the scalar-valued existence result for random variables which were indicators of measurable sets. Since each indicator is in $L^2(\mathfrak{F})$, the existence of conditional expectations is covered by the L^2 result of Lemma 5.14. Combining this with Lemma 5.4, we have another existence proof, which does not need the Radon–Nikodým theorem.

Chapter 6

UMD-Spaces

6.1 Introduction

It is now time to formulate the celebrated UMD property. Exploring its meaning will employ our efforts in this chapter.

Definition 6.1. *A Banach space X is said to have* ***unconditional martingale differences*** *in L^p, $p \in (1,\infty)$, for short, the property* ***UMD-****p, if the inequality*

$$\left|\sum_{k=1}^{n} \epsilon_k \delta f_k\right|_{L^p(\Omega;X)} \leq M_p \left|\sum_{k=1}^{n} \delta f_k\right|_{L^p(\Omega;X)} = M_p \left|f_n\right|_{L^p(\Omega;X)} \tag{6.1}$$

is satisfied, for some constant M_p, by each martingale $f \in L^p(\Omega;X)^{\mathbb{Z}_+}$ and each $\epsilon \in \{-1,1\}^{\mathbb{Z}_+}$.

The equality on the right is of course a tautology, and not part of the definition of UMD. (Observe that the ϵ_k here are simply scalars, not random variables. Recall from Section 3.2 that we use ϵ_k for scalars equal to ± 1 and ε_k for symmetric random variables with this range.)

It is often useful to observe that the UMD-inequality is automatically two-sided: this follows by substituting $\epsilon_k \delta f_k$ in place of δf_k (it is easy to see that this is another martingale difference sequence) and using $\epsilon_k^2 = 1$.

The UMD-condition turns out to be equivalent to several important properties of certain Banach spaces. The relation of some of these is not at all obvious at the first sight, but several equivalent forms to (6.1) of minor depth follow more readily. For instance, we could instead of (6.1) use the condition

$$\left|\sum_{k=1}^{n} \epsilon_k \delta g_k\right|_{L^p(\Omega;X)} \leq M_p \left|\widetilde{g}\right|_{L^p(\Omega;X)} \qquad \text{for} \quad g_k = \mathbb{E}(\,\widetilde{g}|\,\mathfrak{F}_k)\,,$$

which of course follows from (6.1), since $|g_n|_{L^p(\Omega;X)} = |\mathbb{E}(\,\widetilde{g}|\,\mathfrak{F}_k)|_{L^p(\Omega;X)} \leq |\widetilde{g}|_{L^p(\Omega;X)}$ by the contractivity of $\mathbb{E}(\,\cdot|\,\mathfrak{F}_k)$ (Corollary 5.9). To see the converse, take $\widetilde{g} := f_n$.

The smallest constant M_p in (6.1) for given p and a Banach space X is denoted by $M_p(X)$. If X does not have the property UMD-p, then $M_p(X) = \infty$. Note that $M_p(X)$ must be independent of n, i.e., we require in the property UMD-p that (6.1) hold for all n. The smallest constant for which (6.1) holds for a fixed n is denoted by $M_p^n(X)$. Then obviously $M_p^n(X) \uparrow M_p(X)$ as $n \to \infty$. (We can always take some $\delta f_k = 0$ to deduce $M_p^m(X) \leq M_p^n(X)$ for $m < n$.)

Lemma 6.2. *The constants $M_p^n(X)$ are finite; $M_p^n(X) \leq 2n$.*

Proof. This is a simple and crude estimate:

$$\left|\sum_{k=1}^{n} \epsilon_k (f_k - f_{k-1})\right|_{L^p(\Omega;X)} \leq \sum_{k=1}^{n} \left(|f_k|_{L^p(\Omega;X)} + |f_{k-1}|_{L^p(\Omega;X)}\right),$$

and each of the $2n$ terms can be estimated by $|f_k|_{L^p(\Omega;X)} = |\mathbb{E}(\, f_n |\, \mathfrak{F}_k)|_{L^p(\Omega;X)} \leq |f_n|_{L^p(\Omega;X)}$. Thus (6.1) always holds for fixed n with $M_p = 2n$. □

To immediately see that UMD is not a fancy characterization of the null set, we observe the following simple example:

Example 6.3. *Every Hilbert space has UMD-2.*

In particular, $\mathbb{C}$ and $\mathbb{R}$ with absolute value norm have UMD-2.

Proof. By Corollary 5.15, the martingale differences δf_k are orthogonal. It follows from the Pythagoras theorem that the square of the left-hand side of the inequality in (6.1) with $p = 2$ is just $\sum_{k=1}^n |\delta f_k|_X^2$, but so is the square of the right-hand side, if we take $M_2 = 1$. □

In the sequel, we will be working a lot with the definition of UMD-p. To facilitate the work, it will be convenient to make some of observations and notational conventions. So far we have regarded a martingale f as a countably infinite sequence $\{f_k\}_{k=1}^\infty$ adapted to an infinite sequence $\{\mathfrak{F}_k\}_{k=1}^\infty$ of σ-algebras. However, the definition of UMD-p only involves a finite number of random variables f_k, $k = 1, \ldots, n$, and the corresponding σ-algebras. Below, we will also call $\{f_k\}_{k=1}^n$ a martingale adapted to $\{\mathfrak{F}_k\}_{k=1}^n$, if the usual martingale properties are satisfied for $k = 1, \ldots, n$. Such a finite martingale can always be viewed as a sequence of the n first members of an infinite martingale $\{f_k\}_{k=1}^\infty$; indeed, take $\mathfrak{F}_k := \mathfrak{F}_n$ and $f_k := f_n$ for $k > n$.

6.2 Martingale transforms

The UMD-p condition asserts a certain type of uniform boundedness of the operators $f \mapsto \epsilon \star f$ defined by $(\epsilon \star f)_n := \sum_{k=1}^n \epsilon_k \delta f_k$. A direct generalization of these operators leads to the definition of martingale transforms.

Definition 6.4. *Let $f = \{f_k\}_{k=1}^\infty \in L^1(\Omega; X)^{\mathbb{Z}_+}$ be a martingale adapted to $\{\mathfrak{F}_k\}_{k=1}^\infty$. A sequence $v = \{v_k\}_{k=1}^\infty \in \ell^\infty(\mathbb{Z}_+; L^\infty(\Omega))$ is called $\{\mathfrak{F}_k\}_{k=1}^\infty$-**predictable** if each v_k is $\mathfrak{F}_{k-1}$-measurable (and v_1 is $\mathfrak{F}_1$-measurable). For such f and v, the **martingale transform** is $v \star f := \{(v \star f)_k\}_{k=1}^\infty \in L^1(\Omega; X)^{\mathbb{Z}_+}$ given by*

$$(v \star f)_k := \sum_{j=1}^k v_k \delta f_k.$$

It follows that $v \star f$ is also a martingale adapted to $\{\mathfrak{F}_k\}_{k=1}^\infty$; indeed,

$$\mathbb{E}(\, \delta(v \star f)_k |\, \mathfrak{F}_{k-1}) = \mathbb{E}(\, v_k \delta f_k |\, \mathfrak{F}_{k-1}) = v_k \mathbb{E}(\, \delta f_k |\, \mathfrak{F}_{k-1}) = 0.$$

It is also easy to see that $v \star \cdot$ takes L^p martingales into L^p martingales. The sequence of σ-algebras with respect to which a predictable sequence is predictable will usually be clear from the context and will not be referred to explicitly.

In analogy to the UMD-p, it is natural to ask when, if ever, do the operators $v \star \cdot$, satisfy a uniform bound

$$|(v \star f)_n|_{L^p(\Omega;X)} \leq \widetilde{M}_p \, |f_n|_{L^p(\Omega;X)}$$

for all v with $|v|_{\ell^\infty(\mathbb{Z}_+; L^\infty(\Omega))} \leq 1$ (say). (The UMD-p condition is clearly a special case with $v = \epsilon = \{\epsilon_k\}_{k=1}^\infty$; a sequence of constants is certainly predictable.)

The answer, which is not at all obvious, is that the above condition is equivalent to UMD-p. Furthermore, the best constants $M_p(X)$ and $\widetilde{M}_p(X)$ are the same. We now set out to verify this claim. We will refer to this new condition as the property **MT-p**, shorthand for the martingale transform property of order p. If MT-p, as stated above, is satisfied by X, then clearly

$$|(v \star f)_n|_{L^p(\Omega;X)} \leq \widetilde{M}_p \, |v|_{\ell^\infty(\mathbb{Z}_+; L^\infty(\Omega))} \, |f_n|_{L^p(\Omega;X)} \, ,$$

for all predictable $v \in \ell^\infty(\mathbb{Z}_+; L^\infty(\Omega))$, and conversely. For simplicity of notion we can consider, without loss of generality, v with norm bounded by 1. We take the freedom not to always state this explicitly below.

The original conditions UMD-p and MT-p involving arbitrary martingales on arbitrary probability spaces appear overwhelming to work with. Therefore, we first wish to make some reductions. These are established in a series of steps, in each of which we show that a formally weaker condition implies something stronger. The converse implications will be obvious, and we do not always state them explicitly. We begin by showing that it is sufficient to consider **nicely divisible** probability spaces. These are defined by requiring that each set A of positive probability $\mathbb{P}(A)$ has a subset of probability $c\mathbb{P}(A)$ for every $c \in (0,1)$.

Lemma 6.5. *If a Banach space X satisfies the property UMD-p or MT-p for all martingales f on a nicely divisible probability space, then it satisfies the same property for all martingales on any probability space, with the same constant.*

Once this lemma is proved, the assumption that the probability space with which we work is nicely divisible will be assumed throughout this section and the following without explicit reference. When this assumption holds, we can cut sets into appropriate pieces just the way we like. This is clearly a nice property.

Proof. Let $(\Omega, \mathfrak{F}, \mathbb{P})$ be an arbitrary (not necessarily nicely divisible) probability space, and let $\{\mathfrak{F}_k\}_{k=1}^n$ be an increasing sequence of sub-σ-algebras of $\mathfrak{F}$. Consider the product space $(\Omega \times [0,1], \mathfrak{F} \times \mathfrak{M}, \mathbb{P} \times m)$ and the σ-algebras $\mathfrak{G}_k := \{F \times [0,1] : F \in \mathfrak{F}_k\} \subset \mathfrak{F} \times \mathfrak{M}$.

The reason for this construction is the fact that the product space is nicely divisible, whether or not $(\Omega, \mathfrak{F}, \mathbb{P})$ is. Indeed, for $A \in \mathfrak{F} \times \mathfrak{M}$, we know that

$$\mathbb{P} \times m(A) = \int_0^1 \mathbb{P}(\{\omega : (\omega, t) \in A\}) dm(t).$$

Now $g(x) := \int_0^x \mathbb{P}(\{\omega : (\omega, t) \in A\}) dm(t)$ is an increasing continuous function of x with $g(0) = 0$, $g(1) = \mathbb{P} \times m(A)$. Thus g attains every positive $r < \mathbb{P} \times m(A)$ for some $x \in (0,1)$. Now $A \supset A_x := A \cap \{(\omega, t) : t < x\} \in \mathfrak{F} \times \mathfrak{M}$, and $m(A_x) = g(x) = r$, for any given positive $r < \mathbb{P} \times m(A)$. Since the choice of A was quite arbitrary, we see that $(\Omega \times [0,1], \mathfrak{F} \times \mathfrak{M}, \mathbb{P} \times m)$ is nicely divisible.

Now, if $f = \{f_k\}_{k=1}^n$ is a martingale adapted to $\{F_k\}_{k=1}^n$, we define $g_k(\omega, t) := f_k(\omega)$, $\omega \in \Omega$, $t \in [0,1]$. To see that $\{g_k\}_{k=1}^n$ is also a martingale, it suffices to show that $g_k = \mathbb{E}(\, g_n | \, \mathfrak{G}_k)$. This is true, since, for $G = F \times [0,1] \in \mathfrak{G}_k$ $(F \in \mathfrak{F}_k)$,

$$\begin{aligned}
\int_{F \times [0,1]} g_k(\omega, t) d(\mathbb{P} \times m)(\omega, t) &= \int_F \int_0^1 g_k(\omega, t) dm(t) d\mathbb{P}(\omega) = \int_F f_k(\omega) d\mathbb{P}(\omega) \\
&= \int_F f_n(\omega) d\mathbb{P}(\omega) = \int_F \int_0^1 g_n(\omega, t) dm(t) d\mathbb{P}(\omega) = \int_{F \times [0,1]} g_k(\omega, t) d(\mathbb{P} \times m)(\omega, t).
\end{aligned}$$

The vector valued Fubini's theorem was applied in the first and last steps.

Similarly, if $v = \{v_k\}_{k=1}^n \in L^\infty(\Omega)^n$ is $\{\mathfrak{F}_k\}_{k=1}^n$-predictable and $w = \{w_k\}_{k=1}^n$ is defined by $w_k(\omega, t) := v_k(\omega)$, then w is clearly $\{\mathfrak{G}_k\}_{k=1}^n$-predictable.

If we denote by $\widetilde{m}_p^n(X)$ the constant defined like $\widetilde{M}_p^n(X)$, but requiring that the condition MT-p (similarly with UMD-p) need only hold for martingales on nicely divisible probability spaces, it is now straightforward to compute

$$\begin{aligned}
\int_\Omega \left| \sum_{k=1}^n v_k \delta f_k \right|_X^p &= \int_{\Omega \times [0,1]} \left| \sum_{k=1}^n w_k \delta g_k \right|_X^p d(\mathbb{P} \times m) \\
&\leq \left(\widetilde{m}_p^n(X) \right)^p \int_{\Omega \times [0,1]} \left| \sum_{k=1}^n \delta g_k \right|_X^p d(\mathbb{P} \times m) = \widetilde{m}_p^n(X)^p \int_\Omega \left| \sum_{k=1}^n \delta f_k \right|_X^p d\mathbb{P}.
\end{aligned}$$

(Here we only needed a scalar version of Fubini's theorem.) The special case $v = w = \epsilon$ is certainly included in the above computation. We thus see that $\widetilde{M}_p^n(X) \leq \widetilde{m}_p^n(X)$, $M_p^n(X) \leq m_p^n(X)$. The reversed inequalities are obvious. $\square$

Our next goal is reduction to martingales adapted to a sequence of finite subalgebras. Recall that a finite algebra is automatically a σ-algebra.

Lemma 6.6. *The property UMD-p or MT-p is satisfied by X, with the same constant, if it holds for all martingales adapted to finite subalgebras.*

Proof. Let $\{f_k\}_{k=1}^n \in L^p(\Omega; X)^n$ be a martingale adapted to $\{\mathfrak{F}_k\}_{k=1}^n$. The proof essentially relies on the density of simple functions on $L^p(\mathfrak{F}_k; X)$ and the fact that the σ-algebra generated by a simple function, i.e., the smallest σ-algebra on which it is measurable, is finite.

For each k, we choose a simple $s_k \in L^p(\mathfrak{F}_k; X)$ such that $|s_k - \delta f_k|_{L^p(\Omega;X)} < \epsilon$, and a simple $t_k \in L^\infty(\mathfrak{F}_{k-1})$ such that $|t_k - v_k|_{L^\infty(\Omega)} < \epsilon$ and $|t_k|_{L^\infty(\Omega)} \leq |v_k|_{L^\infty(\Omega)}$ (where $(v_k)_{k=1}^\infty$ is a given predictable sequence). (Note that although our general density result (Lemma A.3) does not cover L^∞, it is certainly possible to choose the desired t_k in the scalar valued case; for instance, we could take $t_k := \sum_i x_i \mathbf{1}_{v_k^{-1}(x_i, x_i+\epsilon)}$, where the x_i are points placed at intervals of length ϵ from one another so as to cover the interval between $\pm |v_k|_{L^\infty(\Omega)}$.) Let $\mathfrak{G}_k$ be the algebra generated by $s_1, \ldots, s_k, t_1, \ldots, t_{k+1}$, i.e., by the E_i in the canonical representations $\sum_{i=1}^m x_i \mathbf{1}_{E_i}$ of the above mentioned s_j and t_j. It is then clear that $\{\mathfrak{G}_k\}_{k=1}^n$ is an increasing sequence of finite algebras and that $\{t_k\}_{k=1}^n$ is $\{\mathfrak{G}_k\}_{k=1}^n$-predictable. Since s_k is $\mathfrak{F}_k$-measurable, $\mathfrak{G}_k \subset \mathfrak{F}_k$ and consequently $\mathbb{E}(\,\delta f_{k+1} |\, \mathfrak{G}_k) = \mathbb{E}(\,\mathbb{E}(\,\delta f_{k+1} |\, \mathfrak{F}_k) |\, \mathfrak{G}_k) = 0$.

We define a new martingale g by $\delta g_k := \mathbb{E}(\,\delta f_k |\, \mathfrak{G}_k)$. To see that this is a proper difference sequence of a martingale adapted to $\{\mathfrak{G}_k\}_{k=1}^n$, compute

$$\mathbb{E}(\,\delta g_k |\, \mathfrak{G}_{k-1}) = \mathbb{E}(\,\mathbb{E}(\,\delta f_k |\, \mathfrak{G}_k) |\, \mathfrak{G}_{k-1}) = \mathbb{E}(\,f_k |\, \mathfrak{G}_{k-1}) = 0.$$

Now the rest is just approximation:

$$\begin{aligned}\left|\sum_{k=1}^n v_k \delta f_k\right|_{L^p(\Omega;X)} \leq & \left|\sum_{k=1}^n t_k \delta g_k\right|_{L^p(\Omega;X)} \\ & + \sum_{k=1}^n \left(|t_k(\delta f_k - s_k)|_{L^p(\Omega;X)} + |t_k(s_k - \delta g_k)|_{L^p(\Omega;X)} + |(t_k - v_k)\delta f_k|_{L^p(\Omega;X)}\right)\end{aligned}$$

Here $|\delta f_k - s_k|_{L^p(\Omega;X)} < \epsilon$ and $|t_k - v_k|_{L^\infty(\Omega)} < \epsilon$ by the choice of the s_k and t_k, and

$$|s_k - \delta g_k|_{L^p(\Omega;X)} = |\mathbb{E}(\,s_k - \delta f_k |\, \mathfrak{G}_k)|_{L^p(\Omega;X)} \leq |\delta f_k - s_k|_{L^p(\Omega;X)} < \epsilon,$$

since $\mathbb{E}(\,\cdot |\, \mathfrak{G}_k)$ is contractive on $L^p(\Omega; X)$. Then it is clear how to estimate the L^p norms of products of L^∞ and L^p functions. If we again use $\widetilde{m}_p^n$ as a temporary notation, this time for the best constant in MT-p for martingales f adapted to finite algebras, we have

$$\begin{aligned}\left|\sum_{k=1}^n v_k \delta f_k\right|_{L^p(\Omega;X)} &\leq \left|\sum_{k=1}^n v_k \delta g_k\right|_{L^p(\Omega;X)} + n\left(2 + |f|_{L^p(\Omega;X)}\right)\epsilon \\ &\leq \widetilde{m}_p^n(X) \left|\sum_{k=1}^n \delta g_k\right|_{L^p(\Omega;X)} + n\left(2 + |f|_{L^p(\Omega;X)}\right)\epsilon \\ \leq \widetilde{m}_p^n(X) & \left(\left|\sum_{k=1}^n \delta f_k\right|_{L^p(\Omega;X)} + \sum_{k=1}^n \left(|\delta g_k - s_k|_{L^p(\Omega;X)} + |s_k - \delta f_k|_{L^p(\Omega;X)}\right)\right) + n\left(2 + |f|_{L^p(\Omega;X)}\right)\epsilon \\ &\leq \widetilde{m}_p^n(X) \left|\sum_{k=1}^n \delta f_k\right|_{L^p(\Omega;X)} + n\left(2\widetilde{m}_p^n(X) + 2 + |f|_{L^p(\Omega;X)}\right)\epsilon.\end{aligned}$$

(Recall that we restrict ourselves (without loss of generality) to the case $|v_k|_{L^\infty(\Omega)} \le 1$.) Since the inequality obtained holds for all $\epsilon > 0$, as well as for all martingales $\{f_k\}_{k=1}^n$, we deduce that $\widetilde{M}_p^n(X) \le \widetilde{m}_p^n(X)$, and also $M_p^n(X) \le m_p^n(X)$ as a special case of the same computations. $\square$

From finite algebras we proceed to so called Haar systems of algebras, where the size of the algebras in the sequence $\{\mathfrak{F}_k\}_{k=1}^\infty$ grows in a controlled manner. More precisely, we have the definition below. Recall that a **basis** of a finite algebra $\mathfrak{F}$ on Ω is a partitioning of Ω into disjoint sets $F_i \in \mathfrak{F}$, $i = 1, \dots, m$, which generate $\mathfrak{F}$, i.e., each $F \in \mathfrak{F}$ is a union $\cup_{i\in I} F_i$ for some $I \subset \{1, \dots, n\}$. It is easy to see that the basis of a finite algebra is unique and, in fact, uniquely determines the algebra. It is thus justified to speak of *the* basis and denote it by $\operatorname{bs}\mathfrak{F}$. The sets F_i are sometimes called the **atoms** of $\mathfrak{F}$.

Definition 6.7. *An increasing sequence $\{\mathfrak{H}_k\}_{k=0}^n$ of finite algebras is called a **Haar system** if $\mathfrak{H}_k$ has a basis consisting of $k+1$ sets of positive probability.*

On a nicely divisible probability space, the two conditions in the definition, the size of each $\mathfrak{H}_k$, and the requirement that the sequence be increasing, are easily seen to be equivalent to a more constructive definition of a Haar system: $\mathfrak{H}_0$ is the trivial algebra $\{\emptyset, \Omega\}$ with basis $\{\Omega\}$. If $\mathfrak{H}_k$ (or equivalently, $\operatorname{bs}\mathfrak{H}_k$) is constructed, then the basis of $\mathfrak{H}_{k+1}$ is obtained by taking some $H \in \operatorname{bs}\mathfrak{H}_k$ and dividing it into two parts H_1 and H_2 of positive probability each. Then $\operatorname{bs}\mathfrak{H}_{k+1} := \{H_1, H_2\} \cup \operatorname{bs}\mathfrak{H}_k \setminus \{H\}$.

Now we can formulate the next reduction.

Lemma 6.8. *In the definition of UMD-p or MT-p, it is sufficient to consider all martingales adapted to Haar systems.*

Proof. From the previous results we already know that martingales f on a nicely divisible probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ adapted to finite algebras $\{\mathfrak{F}_k\}_{k=1}^n$ are fully representative of all martingales in view of the property UMD-p.

For a given sequence $\{\mathfrak{F}_k\}_{k=1}^\infty$ of finite algebras, we will construct an auxiliary Haar system $\{\mathfrak{H}_k\}_{k=0}^{N_n}$ so that $\mathfrak{H}_0 \subset \mathfrak{H}_1 \subset \dots \subset \mathfrak{H}_{N_1} = \mathfrak{F}_1 \subset \mathfrak{H}_{N_1+1} \subset \dots \subset \mathfrak{H}_{N_n} = \mathfrak{F}_n$. If we can do this, then the lemma is easily proved: If $f = \{f_k\}_{k=1}^\infty$ is a martingale adapted to $\{\mathfrak{F}_k\}_{k=1}^n$, we define a new martingale h by $h_r := \mathbb{E}(f_n|\mathfrak{H}_r)$. Then clearly $h_{N_k} = f_k$ and $\delta f_k = f_k - f_{k-1} = \sum_{N_{k-1}<r\le N_k} \delta h_r$. Also, if $v = \{v_k\}_{k=1}^n$ is $\{\mathfrak{F}_k\}_{k=1}^\infty$-predictable, then w defined by $w_{N_{k-1}+1} := \dots := w_{N_k} := v_k \in L^\infty(\mathfrak{F}_{k-1}) = L^\infty(\mathfrak{H}_{N_{k-1}})$ is $\{\mathfrak{H}_r\}_{r=1}^{N_n}$-predictable. Thus

$$\left|\sum_{k=1}^n v_k \delta f_k\right|_{L^p(\Omega;X)} = \left|\sum_{r=1}^{N_n} w_r \delta h_r\right|_{L^p(\Omega;X)} \le \widetilde{m}_p(X) \left|\sum_{r=1}^{N_n} \delta h_r\right|_{L^p(\Omega;X)} = \widetilde{m}_p(X) \left|\sum_{k=1}^{n} \delta f_k\right|_{L^p(\Omega;X)},$$

where we once again used $\widetilde{m}_p(X)$ analogously with the proofs of the previous lemmas.

To complete the proof, it suffices to show that the auxiliary Haar system can be constructed. This is done inductively as follows: Let $\mathfrak{H}_0 := \mathfrak{F}_0 := \{\emptyset, \Omega\} \subset \mathfrak{F}_1$. Then assume for induction that we have constructed a Haar system $\{\mathfrak{H}_k\}_{k=1}^r$ so that

$$\mathfrak{H}_0 \subset \dots \subset \mathfrak{H}_{N_i} = \mathfrak{F}_i \subset \dots \subset \mathfrak{H}_r \subsetneq \mathfrak{F}_{i+1} \tag{6.2}$$

for some i. For the construction of $\mathfrak{H}_{r+1}$, choose $F \in \operatorname{bs}\mathfrak{F}_{i+1} \setminus \operatorname{bs}\mathfrak{H}_r$. Such an F must exist, since otherwise $\mathfrak{H}_r \subsetneq \mathfrak{F}_{i+1}$ could not hold. Now each element of $\mathfrak{F}_{i+1} \supset \mathfrak{H}_r$, in particular, each $H \in \operatorname{bs}\mathfrak{H}_r$, is a union of some atoms of $\mathfrak{F}_{i+1}$. On the other hand, $\operatorname{bs}\mathfrak{H}_r$ covers all of Ω. Thus $F \subsetneq H$ for some $H \in \mathfrak{H}_r$. Now define $\operatorname{bs}\mathfrak{H}_{r+1} := \operatorname{bs}\mathfrak{H}_r \cup \{F, H \setminus F\} \setminus \{H\}$. Here $H \setminus F \neq \emptyset$ is a union of some atoms of $\mathfrak{F}_{i+1}$, thus of positive probability. Hence $\{\mathfrak{H}_k\}_{k=1}^{r+1}$ is a Haar system. Also, $\mathfrak{H}_{r+1}$ satisfies either (6.2) with $r+1$ in place of r, or $\mathfrak{H}_{r+1} = \mathfrak{F}_{i+1}$, but if the latter condition holds, then the first one holds for a larger i, unless we already had $i+1 = n$, but then the construction is complete; so is the proof. $\square$

After legitimating these restrictions on the class of martingales we need to consider in the definitions of the properties UMD-p and MT-p, the verification of the equivalence of these two conditions requires almost no effort at all. This was not at all obvious in the beginning, so we nevertheless state the promised result.

Proposition 6.9. *The conditions UMD-p and MT-p are equivalent, with the same constants.*

Thus, we can (and will) henceforth abandon the symbol $\widetilde{M}_p(X)$ and only use $M_p(X)$.

Proof. According to Lemma 6.8, both of these properties reduce to martingales on Haar systems, so all we need to do is to deduce MT-p for Haar systems from UMD-p for Haar systems. We need one simple observation: If $h = \{h_k\}_{k=1}^n$ is a martingale adapted to a Haar system $\{\mathfrak{H}_k\}_{k=1}^n$ and $v = \{v_k\}_{k=1}^\infty$ is a predictable sequence, then $v_k \delta h_k = \lambda_k \delta h_k$ for some $\lambda_k \in \mathbb{R}$, $|\lambda_k| \leq |v_k|_{L^\infty(\Omega)} \leq 1$. Indeed, if $\operatorname{bs}\mathfrak{H}_k$ is obtained from $\operatorname{bs}\mathfrak{H}_{k-1}$ by splitting $H \in \operatorname{bs}\mathfrak{H}_{k-1}$ into H_1 and H_2, then it is clear that the values of f_k and f_{k-1} only differ on H, and the $\mathfrak{H}_{k-1}$-measurable v_k attains a constant value λ_k on this set, which is an atom of $\mathfrak{H}_{k-1}$. Then we compute

$$\left|\sum_{k=1}^n v_k \delta f_k\right|_{L^p(\Omega;X)} = \left|\sum_{k=1}^n \lambda_k \delta f_k\right|_{L^p(\Omega;X)} \leq \max_{\epsilon\in\{-1,1\}^n} \left|\sum_{k=1}^n \epsilon_k \delta f_k\right|_{L^p(\Omega;X)} \leq M_p(X) \left|\sum_{k=1}^n \delta f_k\right|_{L^p(\Omega;X)},$$

where $M_p(X)$ denotes the constant in the condition UMD-p; by this computation this same constant is also appropriate in MT-p. (We used Lemma 2.3 to obtain the first inequality above.) □

6.3 Reduction to Paley–Walsh martingales

In the previous section we already saw that a rather restricted class of martingales is fully representative in view of verifying whether or not a given Banach space X has the property UMD-p (or MT-p). This reduction can be continued even further. In this section we show that the UMD-property is equivalent to merely requiring that the condition (6.1) hold for one special class of martingales, namely those bearing the names of Paley and Walsh. (By Proposition 6.9, the same is true for MT-p. Owing to this lemma, we can concentrate only on the property UMD-p in the following results, without any loss of generality.) The Paley–Walsh martingales are defined as follows:

Definition 6.10. *Let $\mathfrak{D}_k$, $k \in \mathbb{Z}_+$, be the finite algebra generated the partition of $[0,1]$ into 2^k intervals of equal length. The sequence $\{\mathfrak{D}_k\}_{k=1}^\infty$ is called the* **Paley–Walsh system**, *and a martingale $f \in L^1([0,1];X)^{\mathbb{Z}_+}$ adapted to $\{\mathfrak{D}_k\}_{k=1}^\infty$ is called a* **Paley–Walsh martingale**.

We always consider the interval $[0,1]$ with the Lebesgue measure m, which is clearly a probability measure on this space. The collection of Lebesgue measurable sets is denoted by $\mathfrak{M}$. For definiteness, in the definition above, we could take $\mathfrak{D}_k := \{[0,2^{-k}), [2^{-k}, 2\cdot 2^{-k}), \ldots, [1-2^{-k}, 1]\}$, say, but since the set of points $\{r \cdot 2^{-k} \in [0,1] : r, k \in \mathbb{Z}_+\}$ is only countable and thus of measure zero, the way we define the openness or closedness of the intervals is quite irrelevant.

The next step toward Paley–Walsh martingales is to replace the Haar systems with arbitrary bases by **dyadic systems**, where the ratio $\frac{\mathbb{P}(H_1)}{\mathbb{P}(H)}$ of the probability of the new basis element $H_1 \in \mathfrak{H}_{r+1}$ to to the probability of $H \in \mathfrak{H}_r$, $H \supset H_1$, is always of the form $r2^{-m}$, $m \in \mathbb{Z}_+$, $r \in \{1, \ldots, 2^m - 1\}$. Since such dyadic fractions are dense in $(0,1)$, one might guess that the reduction to dyadic Haar systems is a matter of approximation.

Lemma 6.11. *The inequality (6.1) holds for all f adapted to Haar systems if it holds for all f adapted to dyadic Haar systems. The best constants $M_p(X)$ are the same.*

Proof. For a Haar system $\{\mathfrak{H}_k\}_{k=1}^n$, let $\operatorname{bs}\mathfrak{H}_n = \{H_j\}_{j=0}^n$. On a nicely divisible probability space, we can clearly construct a partitioning $\{G_j\}_{j=0}^n$ of Ω, so that the probability of each G_j is an integral multiple of 2^{-m} for some $m \in \mathbb{Z}_+$ and $\mathbb{P}(H_j \Delta G_j) < \epsilon$ for a given $\epsilon > 0$. (We first fix ϵ and then a suitably large m. Δ denotes the symmetric difference $A\Delta B := (A\setminus B)\cup(B\setminus A)$.) We then define $\mathfrak{G}_k$, $k=1,\ldots,n$, by letting $\mathfrak{G}_k$ consist of $\cup_{i\in I}G_i$ for $I\subset\{0,\ldots,n\}$ such that $\cup_{i\in I}H_i\in\mathfrak{H}_k$. Then clearly $\{\mathfrak{G}_k\}_{k=1}^n$ is a dyadic Haar system.

Our intention is to approximate a martingale $h=\{h_k\}_{k=1}^n$ adapted to the original Haar system $\{\mathfrak{H}_k\}_{k=1}^n$ (thus $h_k=\mathbb{E}(\,h_n|\,\mathfrak{H}_k))$ by a martingale g, defined by $g_k := \mathbb{E}(\,h_n|\,\mathfrak{G}_k)$, which is adapted to the dyadic Haar system $\{\mathfrak{G}_k\}_{k=1}^n$. If we can show that each h_k can be approximated by the corresponding g_k arbitrarily well in the L^p norm, then the desired result follows by similar computations as in the earlier lemmas:

$$\begin{aligned}\left|\sum_{k=1}^n \epsilon_k\delta h_k\right|_{L^p(\Omega;X)} &\leq \left|\sum_{k=1}^n \epsilon_k\delta g_k\right|_{L^p(\Omega;X)} + \sum_{k=1}^n |\delta h_k-\delta g_k|_{\leq}\, m_p^n(X)\left|\sum_{k=1}^n \delta g_k\right|_{L^p(\Omega;X)} + 2n\eta \\ &\leq m_p^n(X)\left|\sum_{k=1}^n \delta h_k\right|_{L^p(\Omega;X)} + 2n\left(m_p^n(X)+1\right)\eta,\end{aligned}$$

where $\eta\geq\max_k|h_k-g_k|_X$ is a small positive parameter under our control (hence we now pass it to zero), and $m_p^n(X)$ is used as in previous lemmas.

To complete the proof, we only need to show the desired approximation properties of g. Let us, for the shake of convenience, denote $f:=h_n$. The conditional expectation with respect to a finite algebra has a simple explicit form, and thus we can write

$$h_k=\mathbb{E}(\,f|\,\mathfrak{H}_k)=\sum_{j=0}^k f_{H_j^k}\mathbf{1}_{H_j^k}, \qquad g_k=\mathbb{E}(\,f|\,\mathfrak{G}_k)=\sum_{j=0}^k f_{G_j^k}\mathbf{1}_{G_j^k},$$

where $\operatorname{bs}\mathfrak{H}_k := \{H_j^k\}_{j=0}^k$, $\operatorname{bs}\mathfrak{G}_k := \{G_j^k\}_{j=0}^k$, and we have adopted the shorthand notation $f_A := \frac{1}{\mathbb{P}(A)}\int_A fd\mathbb{P}$ for the average of f on a set A of positive probability.

The goal is to show that the difference of h_k and g_k in L^p can be controlled by the parameter ϵ introduced above. We first show that f_A and f_B are close, if A and B differ only little. To this end, we estimate the terms on the right-hand side of the identity

$$f_A-f_B=\frac{1}{\mathbb{P}(A)}\left(\int_A fd\mathbb{P}-\int_B fd\mathbb{P}\right)+\frac{\mathbb{P}(B)-\mathbb{P}(A)}{\mathbb{P}(A)\mathbb{P}(B)}\int_B fd\mathbb{P}.$$

For the first term we have

$$\left|\int_A fd\mathbb{P}-\int_B fd\mathbb{P}\right|_X \leq \int_{A\Delta B}|f|_X\,d\mathbb{P}\leq |f|_{L^p(\Omega;X)}\left(\mathbb{P}(A\Delta B)\right)^{\frac{1}{q}}\leq |f|_{L^p(\Omega;X)}\,\epsilon^{\frac{1}{q}},$$

where Hölder's inequality was applied, and $\frac{1}{q}>0$, since $p>1$. Since $|\mathbb{P}(A)-\mathbb{P}(B)|\leq\mathbb{P}(A\Delta B)$, it follows that $|f_A-f_B|_X\leq\mathbb{P}(A)^{-1}|f|_{L^p(\Omega;X)}\,\epsilon^{\frac{1}{q}}+(\mathbb{P}(A)\mathbb{P}(B))^{-1}|f|_{L^p(\Omega;X)}\,\epsilon$. We can still estimate $\mathbb{P}(B)^{-1}\leq(\mathbb{P}(A)-\epsilon)^{-1}$ to get an estimate only in terms of A, f and ϵ, and this is a decreasing function $r(\epsilon)$ of $\epsilon>0$ small enough for fixed f and A, and $r(\epsilon)\downarrow 0$ as $\epsilon\downarrow 0$.

Finally, $|f_A\mathbf{1}_A-f_B\mathbf{1}_B|_X \leq |f_A-f_B|_X\,\mathbf{1}_{A\cap B}+|f_A|_X\vee|f_B|_X\,\mathbf{1}_{A\Delta B}$. The first term was already bounded in terms of ϵ above, and the L^p norm of the second term is bounded by $|f|_{L^p(\Omega;X)}(\mathbb{P}(A)^{-1}\vee\mathbb{P}(B)^{-1})\mathbb{P}(A\Delta B)^{\frac{1}{p}}$, and this also decreases to 0 as $\epsilon\downarrow 0$. Since h_k and g_k are combinations of finitely many functions of the form $f_A\mathbf{1}_A$, we see that the L^p norm of their difference is at our disposal. Since there are only finitely many $k=1,\ldots,n$, it is clear that we can even control all the differences $|h_k-g_k|_{L^p(\Omega;X)}$ simultaneously by adjusting the parameter ϵ. □

To facilitate the statement of the following lemma, which will essentially complete our mission in this section, we make some observations concerning functions f measurable with respect to a finite algebra $\mathfrak{F}$ on Ω_1. Since f attains a constant value on each $F\in\operatorname{bs}\mathfrak{F}$, it can be identified with a function with domain $\operatorname{bs}\mathfrak{F}$ in an obvious manner. Below, we will consider a **probability**

preserving Boolean isomorphism $b : \mathfrak{F} \to \mathfrak{G}$, where $\mathfrak{G}$ is another finite algebra on another probability space Ω_2. (Probability preserving simply means $\mathbb{P}_1(F) = \mathbb{P}_2(b(F))$ for all $F \in \mathfrak{F}$, where $\mathbb{P}_1$ and $\mathbb{P}_2$ are probabilities on the two spaces; a necessary and sufficient condition for this is that $\mathbb{P}_1(F) = \mathbb{P}_2(b(F))$ for all $F \in \operatorname{bs}\mathfrak{F}$. A Boolean homomorphism is a mapping that commutes with finite set operations, and an isomorphism is in addition a bijection, as usual.) When we have such a b, and an $\mathfrak{F}$-measurable $f : \Omega_1 \to X$ is identified with a function (still denoted by f) taking $\operatorname{bs}\mathfrak{F}$ into X, then we can define g on $\operatorname{bs}\mathfrak{G}$ by $g := f \circ b^{-1}$, and this can again be identified with a $\mathfrak{G}$-measurable function on Ω_2. It is clear that g defined this way has the same distribution as f, and also that joint distributions of sets of random variables are invariant under the mapping $f \mapsto f \circ b^{-1}$.

Note that "$f \circ b^{-1}$" is actually misuse of notation, when $f : \Omega_1 \to X$ and $b : \mathfrak{F} \to \mathfrak{G}$. Without the identifications of domains used above, the definition of g would be the following: $g(\omega_2) := f(\omega_1)$ whenever $\omega_2 \in b(F)$ and $\omega_1 \in F \in \operatorname{bs}\mathfrak{F}$. This is somewhat cumbersome, and we hence prefer the shorthand notion introduced, which is literally incorrect, but logically sound, when interpreted as explained.

Lemma 6.12. *If $\{\mathfrak{H}_k\}_{k=1}^n$ is a dyadic Haar system on $(\Omega, \mathfrak{F}, \mathbb{P})$, and $\{\mathfrak{D}_k\}_{k=1}^\infty$ is the Paley–Walsh system, then there exists a dyadic Haar system $\{\mathfrak{G}_k\}_{k=1}^n$ on $([0,1], \mathfrak{M}, m)$ such that*

1. *there is a probability preserving Boolean isomorphism $b : \mathfrak{H}_n \to \mathfrak{G}_n$, and*
2. *there are numbers N_k, $k = 1, \dots, n$, such that $\mathfrak{G}_k \subset \mathfrak{D}_{N_k}$ and*
$$\mathbb{E}(\, f |\, \mathfrak{H}_k) = \mathbb{E}\big(\, f \circ b^{-1} \big|\, \mathfrak{G}_k\big) = \mathbb{E}\big(\, f \circ b^{-1} \big|\, \mathfrak{D}_{N_k}\big) \tag{6.3}$$
for all $f \in L^1(\mathfrak{H}_n; X)$.

Note that the conditions imply that b is a probability preserving Boolean homomorphism from $\mathfrak{H}_n$ into $\mathfrak{D}_{N_n}$.

Proof. The construction is similar in spirit to that one in the reduction from systems of finite algebras to Haar systems in Lemma 6.8. We can always set $\mathfrak{D}_0 := \mathfrak{G}_0 := \{\emptyset, [0,1]\}$, $\mathfrak{H}_0 = \{\emptyset, \Omega\}$. Suppose $\mathfrak{G}_1, \dots, \mathfrak{G}_k$ are constructed, so that the conditions of the lemma are satisfied. Note that then the atoms of $\mathfrak{G}_k$ are unions of intervals whose endpoints are integral multiples of 2^{-N_k}. (All this is obvious for the initial step $k = 0$.) If $\operatorname{bs}\mathfrak{H}_{k+1}$ is obtained from $\operatorname{bs}\mathfrak{H}_k$ by splitting $H \in \operatorname{bs}\mathfrak{H}_{k+1}$ into H_1 and H_2 with $\frac{\mathbb{P}(H_1)}{\mathbb{P}(H)} = r2^{-m}$, then we construct $\mathfrak{G}_{k+1}$ as follows: Set $N_{k+1} := N_k + m$. $b(H) \in \mathfrak{G}_k$ is a union of a finite number of intervals with end points $i2^{-N_k}$ and $(i+1)2^{-N_k}$, with integral i. For each such interval, let G_1 consist of the r first subintervals of length $2^{-N_{k+1}}$, and let G_2 consist of the remaining $2^m - r$ subintervals. Then let $\operatorname{bs}\mathfrak{G}_{k+1} := \operatorname{bs}\mathfrak{G}_k \cup \{G_1, G_2\} \setminus \{b(H)\}$ and extend b to a probability preserving Boolean isomorphism from $\mathfrak{H}_{k+1}$ to $\mathfrak{G}_{k+1}$ by defining $b(H_i) := G_i$, $i = 1, 2$. (The rest of the values of b are uniquely determined by the Boolean homomorphism property.)

This construction clearly gives $\{\mathfrak{G}_k\}_{k=1}^n$ such that the property 1 holds. To finish the proof, we must show that the property 2, too, is satisfied. Observe that the first equality in (6.3) is immediate from the construction of $\mathfrak{G}_k = b(\mathfrak{H}_k)$ and $f \circ b^{-1}$. All that remains is the second equality, and we can further denote $g := f \circ b^{-1} \in L^1(\mathfrak{G}_n; X)$ for convenience.

We first show that $\mathbb{E}(\, g |\, \mathfrak{G}_{k-1}) = \mathbb{E}(\, g |\, \mathfrak{D}_{k-1})$ if $g \in L^1(\mathfrak{G}_k; X)$. Indeed, since $\{\mathfrak{G}_k\}_{k=1}^n$ is a Haar system, g has the form $g = g_0 + x_1 \mathbf{1}_{G_1} + x_2 \mathbf{1}_{G_2}$, where $g_0 \in L^1(\mathfrak{G}_{k-1}; X)$ and $G_1, G_2 \in \operatorname{bs}\mathfrak{G}_k \setminus \mathfrak{G}_{k-1}$, $G_1 \cup G_2 \in \operatorname{bs}\mathfrak{G}_{k-1}$ and $g_0 = 0$ on $G_1 \cup G_2$. Then obviously
$$\mathbb{E}(\, g |\, \mathfrak{G}_{k-1}) = g_0 + \frac{m(G_1)x_1 + m(G_2)x_2}{m(G_1 \cup G_2)} \mathbf{1}_{G_1 \cup G_2}.$$

Since $\mathfrak{G}_{k-1} \subset \mathfrak{D}_{N_{k-1}}$, we also have $\mathbb{E}\big(\, g_0 |\, \mathfrak{D}_{N_{k-1}}\big) = g_0$. Furthermore, in the construction of $\mathfrak{G}_k$ from $\mathfrak{G}_{k-1}$ above, each of the intervals $\Delta \in \operatorname{bs}\mathfrak{D}_{N_{k-1}}$, $\Delta \subset G_1 \cup G_2$, was divided between G_1 and

G_2 in the ratio $m(G_1) : m(G_2)$. Thus, for $t \in \Delta \subset G_1 \cup G_2$, $\mathbb{E}\left(x_1 \mathbf{1}_{G_1} + x_2 \mathbf{1}_{G_2} \middle| \mathfrak{D}_{N_{k-1}} \right)(t) = \frac{m(G_1)x_1 + m(G_2)x_2}{m(G_1 \cup G_2)}$. But this exactly agrees with $\mathbb{E}(\, g | \mathfrak{G}_{k-1})$ above.

Now the asserted property 2 follows by induction. Obviously $\mathbb{E}(\, g | \mathfrak{G}_n) = \mathbb{E}(\, g | \mathfrak{D}_n)$ for $g \in L^1(\mathfrak{G}_n; X) \subset L^1(\mathfrak{D}_n; X)$. Assume then that the property 2 is verified for some $k \leq n$, and deduce

$$\mathbb{E}(\, g | \mathfrak{G}_{k-1}) = \mathbb{E}(\, \mathbb{E}(\, g | \mathfrak{G}_k) | \mathfrak{G}_{k-1}) = \mathbb{E}(\, \mathbb{E}(\, g | \mathfrak{D}_k) | \mathfrak{D}_{k-1}) = \mathbb{E}(\, g | \mathfrak{D}_{k-1})\,.$$

In the second step we used the induction assumption, and the previous part of the proof applied to $\mathbb{E}(\, g | \mathfrak{G}_k) \in L^1(\mathfrak{G}_k; X)$.

The proof is complete. □

Finally, it is time to collect the pieces together.

Lemma 6.13. *A Banach space X satisfies the property UMD-p if and only if it satisfies that property for all Paley–Walsh martingales. The best constants $M_p(X)$ are the same.*

Proof. By Lemma 6.11, it is sufficient to check the condition UMD-p for all martingales adapted to dyadic Haar systems. If $f = \{f_k\}_{k=1}^n \in L^1(\Omega; X)^n$ is one such martingale adapted $\{\mathfrak{F}_k\}_{k=1}^n$, Lemma 6.12 provides us with numbers N_k, $k = 1, \ldots, n$ and a probability preserving Boolean homomorphism $b : \mathfrak{F}_n \to \mathfrak{D}_{N_n}$ such that the following computation holds ($f_n \circ b^{-1}$ is to be interpreted as discussed before the statement of Lemma 6.12):

$$\begin{aligned}
&\int_\Omega \left| \sum_{k=1}^n \epsilon_k (\mathbb{E}(\, f_n | \mathfrak{F}_k) - \mathbb{E}(\, f_n | \mathfrak{F}_{k-1})) \right|_X^p d\mathbb{P} \\
&\quad = \int_0^1 \left| \sum_{k=1}^n \epsilon_k (\mathbb{E}\left(\, f_n \circ b^{-1} \middle| \mathfrak{D}_{N_k} \right) - \mathbb{E}\left(\, f_n \circ b^{-1} \middle| \mathfrak{D}_{N_{k-1}} \right)) \right|_X^p dm \\
&\quad = \int_0^1 \left| \sum_{j=1}^{N_n} \epsilon_k (\mathbb{E}\left(\, f_n \circ b^{-1} \middle| \mathfrak{D}_j \right) - \mathbb{E}\left(\, f_n \circ b^{-1} \middle| \mathfrak{D}_{j-1} \right)) \right|_X^p dm \\
&\quad \leq m_p(X) \int_0^1 \left| \sum_{j=1}^{N_n} (\mathbb{E}\left(\, f_n \circ b^{-1} \middle| \mathfrak{D}_j \right) - \mathbb{E}\left(\, f_n \circ b^{-1} \middle| \mathfrak{D}_{j-1} \right)) \right|_X^p dm \\
&\quad = m_p(X) \int_0^1 \left| \sum_{k=1}^n (\mathbb{E}\left(\, f_n \circ b^{-1} \middle| \mathfrak{D}_{N_k} \right) - \mathbb{E}\left(\, f_n \circ b^{-1} \middle| \mathfrak{D}_{N_{k-1}} \right)) \right|_X^p dm \\
&\quad = m_p(X) \int_\Omega \left| \sum_{k=1}^n (\mathbb{E}(\, f_n | \mathfrak{F}_k) - \mathbb{E}(\, f_n | \mathfrak{F}_{k-1})) \right|_X^p d\mathbb{P}.
\end{aligned}$$

Here $m_p(X)$ denoted the smallest constant for which the property UMD-p holds when restricted to Paley–Walsh martingales, and the computation above shows that $M_p(X) \leq m_p(X)$. □

Remark 6.14. *The Paley–Walsh system is not a Haar system. By repeating the argument in the proof of Lemma 6.8, it nevertheless follows that we can even restrict ourselves to martingales adapted to the* ***standard Haar system*** *$\{\mathfrak{H}_k\}_{k=1}^\infty$ with*

$$\text{bs}\, \mathfrak{H}_{2^m+r} := \{[0, 2^{-(m+1)}), \ldots, [(2r-1)2^{-(m+1)}, 2r2^{-(m+1)}), [r2^{-m}, (r+1)2^{-m}), \ldots, [1-2^{-m}, 1]\},$$

for $m \in \mathbb{Z}_+$, $0 \leq r < 2^m$. Clearly $\mathfrak{H}_{2^m} = \mathfrak{D}_m$, where $\{\mathfrak{D}_k\}_{k=1}^\infty$ is the Paley–Walsh system.

6.4 Gundy decomposition and weak-UMD property

So far we have considered the UMD property for martingales on L^p with $p > 1$. Using the notion of martingale transforms, this condition can be given the compact formulation

$$|(\epsilon \star f)_n|_{L^p(\Omega;X)} \leq M_p(X)\, |f_n|_{L^p(\Omega;X)}\,.$$

This can be put in yet another form; taking the supremum over $n \in \mathbb{Z}_+$, it is clear that the above condition implies

$$|\epsilon \star f|_{\ell^\infty(\mathbb{Z}_+;L^p(\Omega;X))} \leq M_p(X)\,|f|_{\ell^\infty(\mathbb{Z}_+;L^p(\Omega;X))}\,.$$

Conversely, if the latter inequality holds for all martingales, it holds in particular for $\{f_k\}_{k=1}^n$, and we deduce the earlier inequality. (Recall that $|f_k|_{L^p(\Omega;X)} = |\mathbb{E}(\,f_n|\,\mathfrak{F}_k)|_{L^p(\Omega;X)} \leq |f_n|_{L^p(\Omega;X)}$ for $k<n$.)

It turns out to be useful to formulate on L^1 a similar weak-type condition.

Definition 6.15. *A Banach space X is said to have the property **weak-UMD** if the inequality*

$$\lambda \mathbb{P}((\epsilon \star f)^* > \lambda) \leq M_w\,|f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}$$

*holds for some constant M_w, for all martingales $f \in L^1(\Omega;X)^{\mathbb{Z}_+}$ and all sequences $\epsilon = \{\epsilon_k\}_{k=1}^\infty \subset \{-1,1\}^{\mathbb{Z}_+}$. The property **weak-MT** is defined similarly.*

Recall that $(\cdot)^*$ denotes the maximal operator, defined in Definition 5.16.

Our goal in this section is to show that each of the conditions MT-p, $1<p<\infty$, implies the weak-MT condition, thus justifying the name. (This result then implies that UMD-p implies weak-UMD, since MT-p follows from UMD-p by Proposition 6.9 and weak-UMD is clearly a special case of weak-MT. Observe, however, that even though we know that UMD-p and MT-p are equivalent, we cannot immediately say the same about the corresponding weak-type conditions.)

In fact, the converse statements are also true and will be examined later on. (And once we do this, it immediately follows (UMD-p $\Rightarrow$ weak-UMD $\Rightarrow$ UMD-q) that all the uncountably many UMD-p conditions are in fact equivalent, something that is so far not at all obvious. The deduction of this fact will, as indicated, go through the weak-type condition; this should give sufficient motivation for the definition of the new property weak-UMD.) For the moment, we concentrate on the first mentioned implication MT-p $\Rightarrow$ weak-MT. The essential tool to establish this will be the Gundy decomposition, which allows us to break arbitrary martingales into pieces with convenient properties. Some preliminary results are first in order.

Lemma 6.16. *If $f \in \ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))$ is a martingale adapted to $\{\mathfrak{F}_k\}_{k=1}^\infty$ and τ is a stopping time, then*

$$\int_{\{\tau<\infty\}} |f_\tau|_X\,d\mathbb{P} \leq |f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}\,.$$

Proof. Using the basic properties of martingales and stopping times we compute

$$\begin{aligned}\sum_{k=1}^n \int_{\{\tau=k\}} |f_k|_X\,d\mathbb{P} \leq \sum_{k=1}^n \int_{\{\tau=k\}} \mathbb{E}(\,|f_n|_X|\,\mathfrak{F}_k)\,d\mathbb{P} &= \sum_{k=1}^n \int_{\{\tau=k\}} |f_n|_X\,d\mathbb{P} \\ &\leq |f_n|_{L^1(\Omega;X)} \leq |f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}\,.\end{aligned}$$

The first step follows from the fact that $\{|f_k(\cdot)|_X\}_{k=1}^\infty$ is a submartingale whenever f is a martingale, the second is a consequence of the definition of conditional expectation and the fact that $\{\tau=k\} \in \mathfrak{F}_k$ for a stopping time τ, and the remaining steps are quite obvious, Taking the limit $n\to\infty$ and observing that $\cup_{k=1}^\infty\{\tau=k\} = \{\tau<\infty\}$ we arrive at the claim of the lemma. □

Lemma 6.17. *Let $f \in \ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))$ be a martingale adapted to $\{\mathfrak{F}_k\}_{k=1}^\infty$ and $\lambda>0$. Let the stopping time τ be defined by*

$$\tau(\omega) := \inf\{n \geq 1 : |f_n(\omega)|_X > \lambda\}$$

and let σ be another stopping time. Then $\left|f_{n\wedge\sigma\wedge(\tau-1)}\right|_{L^1(\Omega;X)} \leq 2\,|f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}$ and

$$\left|\sum_{k=1}^{n\wedge\sigma-1} \mathbb{E}\left(\,\mathbf{1}_{\{\tau=k+1\}}\delta f_{k+1}\,\middle|\,\mathfrak{F}_k\right)\right|_{L^1(\Omega;X)} \leq 2\,|f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}\,.$$

Proof. For the first assertion, we first observe that, on the set $\{\tau < \infty\}$, we have the estimate $\left|f_{n\wedge\sigma\wedge(\tau-1)}\right|_X \leq \lambda < |f_\tau|_X$ by the very definition of τ. On $\{\tau = \infty\}$, we clearly have $n \wedge \sigma \wedge \tau = n \wedge \sigma$. Combining these two simple observations and integrating over Ω we find that

$$\left|f_{n\wedge\sigma\wedge(\tau-1)}\right|_{L^1(\Omega;X)} \leq \int_{\{\tau<\infty\}} |f_\tau|_X \, d\mathbb{P} + \int_{\{\tau=\infty\}} |f_{n\wedge\sigma}|_X \, d\mathbb{P} \leq |f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))} + |f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}.$$

The last step used Lemma 6.16: the first integral is explicitly of the form considered in that lemma, whereas the second involves the stopping time $n \wedge \sigma$, which is everywhere finite, in fact, bounded by n.

For the second assertion, we clearly have

$$\left|\sum_{k=1}^{n\wedge\sigma-1} \mathbb{E}\left(\mathbf{1}_{\{\tau=k+1\}}\delta f_{k+1} \,\middle|\, \mathfrak{F}_k\right)\right|_X \leq \sum_{k=1}^{n\wedge\sigma-1} \mathbb{E}\left(\mathbf{1}_{\{\tau=k+1\}} \left|\delta f_{k+1}\right|_X \,\middle|\, \mathfrak{F}_k\right) X$$

by Jensen's inequality applied to the norm function. Integrating over Ω and estimating $n \wedge \sigma \leq n$ we obtain

$$\begin{aligned}
\left|\sum_{k=1}^{n\wedge\sigma-1} \mathbb{E}\left(\mathbf{1}_{\{\tau=k+1\}}\delta f_{k+1} \,\middle|\, \mathfrak{F}_k\right)\right|_{L^1(\Omega;X)} &\leq \sum_{k=1}^{n-1} \int_{\{\tau=k+1\}} |\delta f_{k+1}|_X \, d\mathbb{P} \\
&\leq \sum_{k=1}^{n-1} 2 \int_{\{\tau=k+1\}} |f_{k+1}|_X \, d\mathbb{P} \leq 2 \sum_{k=2}^{n} \int_{\{\tau=k\}} \mathbb{E}\left(|f_n|_X \,\middle|\, \mathfrak{F}_k\right) d\mathbb{P} \\
&= 2 \sum_{k=2}^{n} \int_{\{\tau=k\}} |f_n|_X \, d\mathbb{P} \leq 2\,|f_n|_{L^1(\Omega;X)} \leq 2\,|f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}.
\end{aligned}$$

The first step simply used the earlier inequality and the fact that the integral of a conditional expectation of a random variable over Ω coincides with the integral of the original random variable. In the second step we estimated $|\delta f_{k+1}|_X$ by the triangle inequality, and observed that $|f_k(\omega)|_X \leq \lambda < |f_{k+1}(\omega)|_X$ for $\omega \in \{\tau = k+1\}$. The third step used the fact that the norm sequence of a martingale is a submartingale, and we also changed the indexing in the summation. The fourth step employed the definition of conditional expectation on $\{\tau = k\} \in \mathfrak{F}_k$, and the rest is obvious. Both assertions have now been verified. □

Now we come to the Gundy decomposition.

Lemma 6.18 (Gundy decomposition). *Let $f \in \ell^\infty(\mathbb{Z}_+; L^1(\Omega;X))$ be a martingale adapted to $\{\mathfrak{F}_k\}_{k=1}^\infty$ and $\lambda > 0$. Then there exists a decomposition $f = g + h + b$, where $g, h, b \in L^1(\Omega;X)^{\mathbb{Z}_+}$ are martingales, which satisfy the following estimates:*

1. $|g|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))} \leq 4\,|f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}$ *and* $|g|_{\ell^\infty(\mathbb{Z}_+;L^\infty(\Omega;X))} \leq 2\lambda$,
2. $\sum_{k=1}^\infty |\delta h_k|_{L^1(\Omega;X)} \leq 4\,|f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}$, *and*
3. $\lambda \mathbb{P}(b^* > 0) \leq 3\,|f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}$.

The numerical constants appearing in this lemma are not necessarily the best possible, and their value is actually quite irrelevant for our purposes. The given constants are the ones that naturally follow from the method of proof.

Proof. The proof will in fact present a construction of the desired decomposition and then show that it satisfies the asserted properties. We define the auxiliary stopping time random variables τ and σ: τ is defined as in Lemma 6.17 and σ is given by

$$\sigma(\omega) := \inf\{n \geq 1 : \sum_{k=1}^{n} \mathbb{E}\left(|\delta f_{k+1}|_X \, \mathbf{1}_{\{\tau=k+1\}} \,\middle|\, \mathfrak{F}_k\right) > \lambda\}.$$

Observe that $\sum_{k=1}^{n} \mathbb{E}\left(|\delta f_{k+1}|_X \mathbf{1}_{\{\tau=k+1\}} \middle| \mathfrak{F}_k\right)$ is $\mathfrak{F}_n$-measurable, so that $\{\sigma \le n\} \in \mathfrak{F}_n$ and σ actually is a stopping time. We will soon exploit the fact that $\{\sigma \ge n\} = \{\sigma \le n-1\}^c \in \mathfrak{F}_{n-1}$. The usefulness of the stopping times τ and σ lies in the fact that "prior to" these moments of (discrete) time, the martingale f remains appropriately bounded.

We now define the martingales g, h and b by giving the corresponding difference sequences, starting from g:

$$\begin{aligned}
\delta g_1 &:= \mathbf{1}_{\{\tau>1\}}\delta f_1 \\
\delta g_k &:= \mathbf{1}_{\{\sigma\ge k\}}\left(\mathbf{1}_{\{\tau>k\}}\delta f_k - \mathbb{E}\left(\mathbf{1}_{\{\tau>k\}}\delta f_k \middle| \mathfrak{F}_{k-1}\right)\right) \qquad k>1.
\end{aligned}$$

Note that we could also have $\mathbf{1}_{\{\sigma\ge 1\}}$ as an additional factor in δg_1, but this is unnecessary, since $\sigma \ge 1$ on all of Ω. $\{\delta g_k\}_{k=1}^{\infty}$ is a proper martingale difference sequence, as is easily verified by considering the conditional expectation $\mathbb{E}\left(\delta g_k \middle| \mathfrak{F}_{k-1}\right)$ and "taking out" the $\mathfrak{F}_{k-1}$-measurable $\mathbf{1}_{\{\sigma\ge k\}}$. Observe how we have used the stopping time τ to ensure that f_k and f_{k-1} appearing in the formula for δg_k are bounded in norm by λ, since the indicators differ from zero only on $\{\tau > k\}$. g represents the part of f that is bounded in L^∞ relative to the parameter λ, as we wanted. More precisely, we have

$$\begin{aligned}
g_n = \sum_{k=1}^{n} \delta g_k = \sum_{k=1}^{n} \mathbf{1}_{\{\sigma\ge k\}}\mathbf{1}_{\{\tau>k\}}\delta f_k - \sum_{k=2}^{n} \mathbf{1}_{\{\sigma\ge k\}}\mathbb{E}\left(\mathbf{1}_{\{\tau>k\}}\delta f_k \middle| \mathfrak{F}_{k-1}\right) \\
= \sum_{k=1}^{n\wedge\sigma\wedge(\tau-1)} \delta f_k + \sum_{k=2}^{n\wedge\sigma} \mathbb{E}\left(\mathbf{1}_{\{\tau=k\}}\delta f_k \middle| \mathfrak{F}_{k-1}\right),
\end{aligned}$$

where in the last step we converted the indicator involving $\{\tau > k\}$ into one with $\{\tau = k\}$ by the identity

$$\mathbb{E}\left(\mathbf{1}_{\{\tau=k\}}\delta f_k \middle| \mathfrak{F}_{k-1}\right) + \mathbb{E}\left(\mathbf{1}_{\{\tau>k\}}\delta f_k \middle| \mathfrak{F}_{k-1}\right) = \mathbb{E}\left(\mathbf{1}_{\{\tau\ge k\}}\delta f_k \middle| \mathfrak{F}_{k-1}\right) = \mathbf{1}_{\{\tau\ge k\}}\mathbb{E}\left(\delta f_k \middle| \mathfrak{F}_{k-1}\right) = 0, \quad (6.4)$$

which holds since $\mathbf{1}_{\{\tau\ge k\}}$ is $\mathfrak{F}_{k-1}$-measurable. $|g_n|_X$ can now be estimated, using the triangle inequality, by the sum of $\left|f_{n\wedge\sigma\wedge(\tau-1)}\right|_X$ and $\left|\sum_{k=1}^{n\wedge\sigma-1} \mathbb{E}\left(\mathbf{1}_{\{\tau=k+1\}}\delta f_{k+1} \middle| \mathfrak{F}_k\right)\right|_X$, and by the definitions of the stopping times τ and σ, both of these are bounded by λ; thus g satisfies the asserted L^∞ bound. Furthermore, estimating the L^1 norm of g_n similarly and recalling Lemma 6.17 we also deduce the desired L^1 bound.

While g was the part of f corresponding to the "moments of time" $k < \tau$, when f remains bounded by the parameter λ, we define h to be the part of f related to the time τ, when f just exceeds this bound:

$$\begin{aligned}
\delta h_1 &:= \mathbf{1}_{\{\tau=1\}}\delta f_1 \\
\delta h_k &:= \mathbf{1}_{\{\sigma\ge k\}}\left(\mathbf{1}_{\{\tau=k\}}\delta f_k - \mathbb{E}\left(\mathbf{1}_{\{\tau=k\}}\delta f_k \middle| \mathfrak{F}_{k-1}\right)\right) \qquad k>1.
\end{aligned}$$

We add g and h making use of the equation (6.4) to find that $\delta g_k + \delta h_k = \mathbf{1}_{\{\sigma\ge k\}}\mathbf{1}_{\{\tau\ge k\}}\delta f_k = \mathbf{1}_{\{\sigma\wedge\tau\ge k\}}\delta f_k$. Thus, if we want to define the remaining martingale b so as to produce a proper decomposition of f, we must set

$$\delta b_k := \delta f_k - \delta g_k - \delta h_k = \mathbf{1}_{\{\sigma\wedge\tau< k\}}\delta f_k.$$

Observe that $b_1 = 0$.

It remains to show that h and b satisfy the asserted estimates. For h we have:

$$\begin{aligned}
\sum_{k=1}^{\infty} |\delta h_k|_{L^1(\Omega;X)} \le \sum_{k=1}^{\infty} \int_\Omega \left(\mathbf{1}_{\{\tau=k\}} |\delta f_k|_X + \mathbb{E}\left(\mathbf{1}_{\{\tau=k\}} |\delta f_k|_X \middle| \mathfrak{F}_{k-1}\right)\right) d\mathbb{P} \\
\le \sum_{k=1}^{\infty} \int_{\{\tau=k\}} 4\,|f_k|_X \, d\mathbb{P} = 4\int_{\{\tau<\infty\}} |f_\tau|_X \, d\mathbb{P} \le 4\,|f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}\,.
\end{aligned}$$

The first inequality was a simple norm estimate (where we can interpret the conditional expectation with respect to $\mathfrak{F}_0$ to be zero). In the second step we have used the identity $\int_\Omega \mathbb{E}(\,\cdot \,|\, \mathfrak{F}_k)\, d\mathbb{P} = \int_\Omega \cdot d\mathbb{P}$ to remove the conditional expectation, and we have also estimated $|\delta f_k|_X \leq 2\,|f_k|_X$, since on $\{\tau = k\}$ we have $|f_{k-1}|_X \leq \lambda < |f_k|_X$. The last step is a direct application of Lemma 6.16.

Finally, we must estimate the probability of the event $\{b^* > 0\}$. Since

$$b_n = \sum_{k=1}^{n} \delta b_k = \sum_{k=1}^{n} \mathbf{1}_{\{\sigma\wedge\tau<k\}} \delta f_k = \sum_{k=\sigma\wedge\tau+1}^{n} \delta f_k = 0 \qquad \text{for} \qquad n \leq \sigma \wedge \tau,$$

it is clear that b and thus b^* can only differ from zero if $\tau \wedge \sigma < \infty$. Thus $\mathbb{P}(b^* > 0) \leq \mathbb{P}(\tau < \infty) + \mathbb{P}(\sigma < \infty)$, and $\mathbb{P}(\tau < \infty) = \mathbb{P}(f^* > \lambda) \leq \frac{1}{\lambda}\,|f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}$ gives an estimate of the desired form. Similarly, using the definition of σ, we find that

$$\{\sigma < \infty\} = \left\{ \sum_{k=1}^{\infty} \mathbb{E}\left(\mathbf{1}_{\{\tau=k+1\}}\, |\delta f_{k+1}|_X \,\middle|\, \mathfrak{F}_k \right) > \lambda \right\}, \tag{6.5}$$

and to estimate the probability of a set on which a function exceeds a given value it is sufficient to estimate the L^1 norm, as is well known. Thus we compute

$$\begin{aligned} \sum_{k=1}^{\infty} \int_\Omega \mathbb{E}\left(\mathbf{1}_{\{\tau=k+1\}}\, |\delta f_{k+1}|_X \,\middle|\, \mathfrak{F}_k \right) d\mathbb{P} &= \sum_{k=1}^{\infty} \int_{\{\tau=k+1\}} |\delta f_{k+1}|_X \, d\mathbb{P} \\ &\leq 2 \int_{\{\tau<\infty\}} |f_\tau|_X \, d\mathbb{P} \leq 2\, |f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}\,, \end{aligned}$$

where in the second step we once again estimated $|f_k|_X \leq \lambda < |f_{k+1}|_X$ on $\{\tau = k+1\}$, and the last step was again Lemma 6.16. It then follows, recalling (6.5), that $\mathbb{P}(\sigma < \infty) \leq \frac{1}{\lambda} 2\,|f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}$, and this combined with the estimate for $\mathbb{P}(\tau < \infty)$ gives the asserted bound for $\mathbb{P}(b^* > 0)$. All the asserted properties of the Gundy decomposition have now been verified. □

The Gundy decomposition in our tool box, the derivation of the property weak-MT from MT-p is a straightforward computation.

Proposition 6.19. *Each of the properties MT-p implies weak-MT.*

Proof. Let X satisfy the condition MT-p for a certain p, which will be fixed from now on. Also fix $\lambda > 0$. Let $f \in \ell^\infty(\mathbb{Z}_+; L^1(\Omega;X))$ be a martingale, and v be a predictable sequence, $|v|_{\ell^\infty(\mathbb{Z}_+;L^\infty(\Omega;X))} \leq 1$. We must prove that f satisfies the inequality $\lambda\mathbb{P}((v \star f)^* > \lambda) \leq M_w\, |f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}$ for some M_w independent of f, λ and v. We do this by establishing separately a similar inequality for each of the three martingales in the Gundy decomposition of f. More precisely, let $f = g+h+b$ be the Gundy decomposition of f relative to λ. Then $v\star f = v\star g+v\star h+v\star b$ and

$$\mathbb{P}((v \star f)^* > \lambda) \leq \mathbb{P}((v \star g)^* > \frac{\lambda}{3}) + \mathbb{P}((v \star h)^* > \frac{\lambda}{3}) + \mathbb{P}((v \star b)^* > \frac{\lambda}{3}).$$

The last of these is easily estimated by observing that $(v \star b)^*$ can only differ from zero if b^* does, and thus

$$\mathbb{P}((v \star b)^* > \frac{\lambda}{3}) \leq \mathbb{P}(b^* > 0) \leq \frac{3}{\lambda}\, |f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}$$

by the properties of the Gundy decomposition.

It is also immediate that

$$\mathbb{P}((v \star h)^* > \frac{\lambda}{3}) \leq \frac{3}{\lambda}\, |v \star h|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}$$

by Doob's inequality, and

$$|(v \star h)_n|_{L^1(\Omega;X)} \leq \sum_{k=1}^{n} |v_k|_{L^\infty(\Omega;X)}\, |\delta h_k|_{L^1(\Omega;X)} \leq \sum_{k=1}^{\infty} |\delta h_k|_{L^1(\Omega;X)} \leq 4\, |f|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}$$

(again by the properties of the Gundy decomposition), and taking the supremum over $n \in \mathbb{Z}_+$ we obtain the same inequality for $|v \star h|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}$. Combining this with the previous inequality yields a bound of the desired form.

So far we have not used the assumption MT-p, but only the properties of the martingales h and b in the Gundy decomposition. One might guess that the real work is to be done on $v \star g$. The basic difficulty with the proof is clearly the fact that we should deduce a result covering all L^1-bounded martingales from an assumption concerning only L^p martingales for some $p > 1$. Now the g in the Gundy decomposition is not only bounded in L^1 but also in L^∞, and the boundedness in L^p $1 < p < \infty$ follows by interpolation. This is the idea of the proof, now for the detail:

$$\begin{aligned}
\mathbb{P}\left((v \star g)^* > \frac{3}{\lambda}\right) &\le \frac{3^p}{\lambda^p} \left|(v \star g)^*\right|^p_{L^p(\Omega;X)} \le \frac{3^p}{\lambda^p}\overline{p}^p \left|v \star g\right|^p_{\ell^\infty(\mathbb{Z}_+;L^p(\Omega;X))} \\
&\le \frac{3^p\overline{p}^p}{\lambda^p} M_p(X)^p \left|g\right|^p_{\ell^\infty(\mathbb{Z}_+;L^p(\Omega;X))} \le \frac{3^p\overline{p}^p}{\lambda^p} M_p(X)^p \left|g\right|^{p-1}_{\ell^\infty(\mathbb{Z}_+;L^\infty(\Omega;X))} \left|g\right|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))} \\
&\le \frac{3^p\overline{p}^p}{\lambda^p} M_p(X)^p \left(4^{p-1}\lambda^{p-1}\right)\left(4\left|f\right|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}\right) = \frac{1}{\lambda} 12^p \overline{p}^p M_p(X)^p \left|f\right|_{\ell^\infty(\mathbb{Z}_+;L^1(\Omega;X))}.
\end{aligned}$$

The first step is a simple standard estimate and the second is Doob's L^p inequality. The third step employs the assumption of UMD-p, and the fourth inequality is obtained by taking the supremum over n on both sides of the inequality

$$\int_\Omega |g_n|^p_X \, d\mathbb{P} = \int_\Omega |g_n|^{p-1}_X |g_n|_X \, d\mathbb{P} \le |g_n|^{p-1}_{L^\infty(\Omega;X)} \int_\Omega |g_n|_X \, d\mathbb{P}.$$

The fifth step uses the properties of the martingale g in the Gundy decomposition, and the last step is just rearrangement. The proof is complete. ☐

6.5 Weak equals strong

As a last task in this chapter, we will show the converse of Proposition 6.19. Once this is done, we have rather good picture of various characterizations of the UMD property; in fact, we have the following:

Theorem 6.20 (Burkholder 1981). *In a Banach space X, the following conditions are equivalent:*

1. *X has UMD-p for all $p \in (1, \infty)$.*
2. *X has MT-p for all $p \in (1, \infty)$.*
3. *X has UMD-p for some $p \in (1, \infty)$.*
4. *X has MT-p for some $p \in (1, \infty)$.*
5. *X has weak-UMD.*
6. *X has weak-MT.*

Proof. By Proposition 6.9 the conditions 1 and 2, respectively 3 and 4, are equivalent.

$1, 2 \Rightarrow 3, 4$ is obvious.

$4 \Rightarrow 6$ was just proved in Proposition 6.19.

$6 \Rightarrow 5$ is obvious.

$5 \Rightarrow 1$ will be proved in this section, Proposition 6.24. ☐

So far we have talked of a Banach space X having the property UMD-p etc., and not really given a meaning for the concept UMD-space appearing in the heading of this chapter. We have done so in order not to place any one of the conditions in Theorem 6.20 above others, but to emphasize the equivalence of all of them. Thus we state:

Definition 6.21. *A Banach space X which has any one, and thus all, of the properties enumerated in Theorem 6.20 is called a* ***UMD-space****.*

Having celebrated a result whose proof we have not yet finished, we now calm down and return to work. We first require a couple of lemmas.

Lemma 6.22. *Let $\beta, p > 1$ and $\gamma, \delta > 0$ so that $\beta^p\gamma < 1$. If the positive measurable functions f and g satisfy*

$$\mathbb{P}(g > \beta t, f \leq \delta t) \leq \gamma \mathbb{P}(g > t)$$

for all $t > 0$, then $|g|_{L^p(\Omega)} \leq (1 - \beta^p\gamma)^{-\frac{1}{p}} \frac{\beta}{\delta} |f|_{L^p(\Omega)}$.

Proof. If g satisfies the conditions, so does $g \wedge n$. (This essentially depends on the fact that $\beta > 1$.) Thus, if the lemma is verified for $g \in L^p(\Omega)$, the general case follows by the monotone convergence theorem. For $g \in L^p(\Omega; X)$, this is a straightforward computation using distribution functions. □

Lemma 6.23. *Let X be a Banach space with weak-UMD and $f = (f_k)_{k=1}^\infty$ a martingale. If there is a predictable sequence $w = (w_k)_{k=1}^\infty$ which dominates the differences of f, i.e., $|\delta f_k|_X \leq w_k$ (a.s.), then*

$$\mathbb{P}((\epsilon \star f)^* > 2\lambda, f^* \vee w^* \leq \delta\lambda) \leq \frac{3\delta}{1-\delta} M_w(X) \mathbb{P}((\epsilon \star f)^* > \lambda)$$

for all $\lambda > 0$, $\delta \in (0,1)$ and all sequences of signs $\epsilon \in \{-1,1\}^{\mathbb{Z}_+}$.

Proof. Fix $\lambda > 0$ and $\epsilon \in \{-1,1\}^{\mathbb{Z}_+}$, and define the auxiliary stopping time random variables

$$\begin{aligned}\tau_j(\omega) &:= \inf\{k : |(\epsilon \star f)_k(\omega)|_X > j\lambda\} \qquad j = 1,2,\\ \Delta(\omega) &:= \inf\{k : |f_k(\omega)|_X \vee w_{k+1}(\omega) > \delta\lambda\}.\end{aligned}$$

Also denote

$$T_k := \{\tau_1 < k \leq \tau_2 \wedge \Delta\} = \{\omega : \lambda < \max_{j \leq k-1} |(\epsilon \star f)_j(\omega)|_X \leq 2\lambda, \max_{j \leq k-1} \left(|f_j(\omega)|_X \vee w_{j+1}(\omega)\right) \leq \delta\lambda\}.$$

Then $T_k \in \mathfrak{F}_{k-1}$ and $u := (u_k)_{k=1}^\infty := (\mathbf{1}_{T_k})_{k=1}^\infty$ is a predictable sequence.

We first claim that

$$\{(\epsilon \star f)^* > 2\lambda, f^* \vee w^* \leq \delta\lambda\} \subset \{(u \star \epsilon \star f)^* > (1-\delta)\lambda\}. \tag{6.6}$$

Observe that the left-hand side appears in the assertion of the lemma, whereas the right-hand side is something that we can estimate by the assumption of weak-UMD; indeed

$$\mathbb{P}((\epsilon \star u \star f)^* > (1-\delta)\lambda) \leq \frac{M_w(X)}{(1-\delta)\lambda} |u \star f|_{\ell^\infty(\mathbb{Z}_+; L^1(\Omega;X))}. \tag{6.7}$$

We have written $u \star \epsilon \star f$ for $u \star (\epsilon \star f)$ (and similarly $\epsilon \star u \star f$), whose nth member is given by

$$(u \star (\epsilon \star f))_n = \sum_{k=1}^n u_k \delta(\epsilon \star f)_k = \sum_{k=1}^n u_k \epsilon_k \delta f_k.$$

This last form shows that $u \star (\epsilon \star f) = \epsilon \star (u \star f)$, making the estimate (6.7) useful once the claim (6.6) is proved. We have left out the parentheses for convenience.

To prove (6.6), we will in fact show that $(u \star \epsilon \star f)_{\tau_2} > (1-\delta)\lambda$ for each point on the left-hand side of (6.6); observe that $(\epsilon \star f)^* > 2\lambda$ implies $\tau_2 < \infty$. On the other hand, $f^* \vee w^* \leq \delta\lambda$ implies $\Delta = \infty$.

It is always true that $\tau_1 \leq \tau_2$. On the set on the left-hand side of (6.6), we have $\tau_1 < \tau_2$ strictly; indeed, from $(\epsilon \star f)_{\tau_2} = (\epsilon \star f)_{\tau_2 - 1} + \epsilon_{\tau_2} \delta f_{\tau_2}$ we get, for each point on the left-hand side of (6.6),

$$|(\epsilon \star f)_{\tau_2 - 1}|_X \geq |(\epsilon \star f)_{\tau_2}|_X - |\delta f_{\tau_2}|_X > 2\lambda - |w_{\tau_2}|_X \geq (2-\delta)\lambda > \lambda,$$

and this means $\tau_1 \leq \tau_2 - 1$ by the very definition of τ_1.

We then compute (still considering a point in the left-hand side of (6.6), whence $\Delta = \infty$)

$$(u \star \epsilon \star f)_{\tau_2} = \sum_{k=1}^{\tau_2} \mathbf{1}_{\tau_1 < k \leq \tau_2 \wedge \Delta} \delta(\epsilon \star f)_k = \sum_{\tau_1 < k \leq \tau_2} \delta(\epsilon \star f)_k = (\epsilon \star f)_{\tau_2} - (\epsilon \star f)_{\tau_1},$$

and thus

$$|(u \star \epsilon \star f)_{\tau_2}|_X \geq |(\epsilon \star f)_{\tau_2}|_X - |(\epsilon \star f)_{\tau_1}|_X > 2\lambda - |(\epsilon \star f)_{\tau_1 - 1}|_X - |\delta f_{\tau_1}|_X \geq 2\lambda - \lambda - w_{\tau_1} \geq \lambda - \delta\lambda.$$

This verifies the claim (6.6).

Looking at the inequality (6.7), it now seems natural to start working on $u \star f$. There are two exclusive possibilities (for each point ω of the probability space Ω): either $\tau_1 \geq \tau_2 \wedge \Delta$, or $\tau_1 < \tau_2 \wedge \Delta$. In the first case, we clearly have $\omega \notin T_k = \{\tau_1 < k \leq \tau_2 \wedge \Delta\}$ for any k, thus $u_k(\omega) = \mathbf{1}_{T_k}(\omega) = 0$, and $(u \star f)_k(\omega) = 0$ for all $k \in \mathbb{Z}_+$. In the second case, $\omega \in T_k$ for some (possibly many) $k \in \mathbb{Z}_+$. For such ω and k, $\tau_1 < k \leq \tau_2 \wedge \Delta$, we have

$$(u \star f)_k = \sum_{j=1}^{k} \mathbf{1}_{\{\tau_1 < j \leq \tau_2 \wedge \Delta\}} \delta f_j = \sum_{j=\tau_1+1}^{k} \delta f_j = f_k - f_{\tau_1}.$$

A similar computation for $k \leq \tau_1$ gives $(u \star f)_k = 0$, and for $k > \tau_2 \wedge \Delta$ we have $(u \star f)_k = f_{\tau_2 \wedge \Delta} - f_{\tau_1}$.

Thus $|(u \star f)_k|_X \leq |f_{k \wedge \tau_2 \wedge \Delta}|_X + |f_{\tau_1}|_X$. Now we restrict the considerations to the set $\{\tau_1 < \tau_2 \wedge \Delta\}$ (recall that outside this set, $u \star f = 0$); thus $|f_{\tau_1}|_X \leq \delta\lambda$ by the definition of Δ. Furthermore,

$$|f_{k \wedge \tau_2 \wedge \Delta}|_X \leq |f_{k \wedge \tau_2 \wedge \Delta - 1}|_X + |\delta f_{k \wedge \tau_2 \wedge \Delta}|_X \leq \delta\lambda + w_{k \wedge \tau_2 \wedge \Delta} \leq 2\delta\lambda,$$

where we again used the definition of the stopping time Δ. These estimates combine to give $|(u \star f)_k|_X \leq 3\delta\lambda$ for all $k \in \mathbb{Z}_+$ on $\{\tau_1 < \tau_2 \wedge \Delta\}$, and $u \star f = 0$ outside this set, in particular, $u \star f = 0$ on $\{\tau_1 = \infty\} = \{(\epsilon \star f)^* > \lambda\}$. Thus we conclude that

$$|(u \star f)^*|_X \leq 3\delta\lambda \mathbf{1}_{\{(\epsilon \star f)^* > \lambda\}}$$

and thus

$$|u \star f|_{\ell^\infty(\mathbb{Z}_+; L^1(\Omega; X))} \leq |(u \star f)^*|_{L^1(\Omega; X)} \leq 3\delta\lambda \mathbb{P}((\epsilon \star f)^* > \lambda).$$

(The first step used the easy part of Doob's inequality.)

To complete the proof, it suffices to combine (6.6), (6.7) and the inequality above with the assumption of weak-UMD to yield

$$\mathbb{P}((\epsilon \star f)^* > 2\lambda, f^* \vee w^* \leq \delta\lambda) \leq \frac{M_w(X)}{(1-\delta)\lambda} |u \star f|_{\ell^\infty(\mathbb{Z}_+; L^1(\Omega; X))} \leq \frac{3\delta}{1-\delta} M_w(X) \mathbb{P}((\epsilon \star f)^* > \lambda).$$

The proof is finished. □

Now we obtain the desired result.

Proposition 6.24. *If a Banach space X has weak-UMD, then it has UMD-p for all $p \in (1, \infty)$.*

Proof. By Remark 6.14, it suffices to prove the property UMD-p for all martingales adapted to the standard Haar system. If f is such a martingale, then δf_k is non-zero only on one interval $I_k = I_k^1 \cup I_k^2$, $m(I_k^1) = m(I_k^2)$, $I_k^1, I_k^2 \in \operatorname{bs} \mathfrak{H}_k$, $I_k \in \operatorname{bs} \mathfrak{H}_{k-1}$. From the condition $\mathbb{E}\delta f_k = 0$ it now follows that $\delta f_k = x_k \mathbf{1}_{I_k^1} - x_k \mathbf{1}_{I_k^2}$ for some $x_k \in X$. Define $w_k := |\delta f_k|_X = |x_k|_X \mathbf{1}_{I_k} \in L^\infty(\mathfrak{H}_{k-1})$. Then obviously $|\delta f_k|_X \leq w_k$, and $w = \{w_k\}_{k=1}^\infty$ is predictable. Furthermore, $w_k \leq |f_k|_X + |f_{k-1}|_X \leq 2f^*$; thus $w^* \leq 2f^*$.

Now we are in a position to apply Lemmas 6.22 and 6.23. The latter one guarantees that

$$\mathbb{P}((\epsilon \star f)^* > 2t, f^* \vee w^* \leq \delta t) \leq \frac{3\delta}{1-\delta} M_w(X) \mathbb{P}((\epsilon \star f)^* > t)$$

for all $\delta \in (0,1)$ and all $t > 0$. For a fixed $p > 1$, choose $\delta = \delta(p) \in (0,1)$ small enough so that $2^p \frac{3\delta}{1-\delta} M_w(X) < 1$. Then Lemma 6.22 applied to the positive measurable functions $(\epsilon \star f)^*$ and $f^* \vee w^*$ says that

$$|(\epsilon \star f)^*|_{L^p(\Omega)} \leq \left(1 - 2^p \frac{3\delta}{1-\delta} M_w(X)\right)^{-\frac{1}{p}} \frac{2}{\delta} |f^* \vee w^*|_{L^p(\Omega)} =: M_p' |f^* \vee w^*|_{L^p(\Omega)}.$$

Bearing in mind that $w^* \leq 2f^*$, we further compute

$$|(\epsilon \star f)^*|_{L^p(\Omega)} \leq M_p' \left(|f^*|_{L^p(\Omega)} + |w^*|_{L^p(\Omega)}\right) \leq 3M_p' |f^*|_{L^p(\Omega)},$$

and finally

$$|\epsilon \star f|_{\ell^\infty(\mathbb{Z}_+;L^p(\Omega;X))} \leq |(\epsilon \star f)^*|_{L^p(\Omega)} \leq 3M_p'\overline{p}\, |f|_{\ell^\infty(\mathbb{Z}_+;L^p(\Omega;X))},$$

where we again used Doob's inequalities. The assertion now follows with $M_p(X) \leq 3M_p'\overline{p}$. $\square$

Now that this proposition, and thus Theorem 6.20, is proved we can we can legitimately explore some immediate consequences. First of all, we can restate Example 6.3 at the beginning of the chapter:

Example 6.25. *Every Hilbert space is UMD; in particular, $\mathbb{C}$ and $\mathbb{R}$ are UMD.*

We will see in Chapter 7 that there are other UMD-spaces, too. At this point, we present another close relative of the UMD-inequality, the square function estimate, which is satisfied by every L^p-martingale on a Hilbert space, as the example will show.

Example 6.26. *For a martingale $f = (f_k)_{k=1}^\infty \subset \ell^\infty(\mathbb{Z}_+; L^p(\Omega;\mathcal{H}))$, $p \in (1,\infty)$, $\mathcal{H}$ Hilbert, the* ***square function*** *Sf, defined by*

$$Sf(\omega) := \sqrt{\sum_{k=1}^\infty |\delta f_k(\omega)|_{\mathcal{H}}^2},$$

satisfies

$$s_p\, |f|_{\ell^\infty(\mathbb{Z}_+;L^p(\Omega;\mathcal{H}))} \leq |Sf|_{L^p(\Omega)} \leq S_p\, |f|_{\ell^\infty(\mathbb{Z}_+;L^p(\Omega;\mathcal{H}))}$$

Proof. The result follows quite immediately from the UMD-inequality, which is valid, since every Hilbert space is UMD, and the Hilbert space version of the Khintchine–Kahane inequality. Indeed, let ε_k be Rademacher functions on Ω, as usual, and ε_k' be Rademacher functions on another probability space Ω' with probability measure $\mathbb{P}'$. First observe that

$$\left|\sqrt{\sum_{k=1}^n |\delta f_k(\cdot)|_{\mathcal{H}}^2}\right|_{L^p(\Omega)}^p = \int_\Omega \left(\sum_{k=1}^n |\delta f_k(\omega)|_{\mathcal{H}}^2\right)^{\frac{p}{2}} d\mathbb{P}(\omega) = \int_\Omega \left|\sum_{k=1}^n \varepsilon_k' \delta f_k(\omega)\right|_{L^2(\Omega';\mathcal{H})}^p d\mathbb{P}(\omega),$$

which is easily seen by writing the norm on the right-hand side in terms of the inner product of $\mathcal{H}$ and using the orthonormality of the Rademacher functions.

By the Khintchine–Kahane inequality (applied pointwise inside the integral), this quantity can be estimated by

$$\begin{aligned}\leq A_p^p \int_\Omega \left|\sum_{k=1}^n \varepsilon_k' \delta f_k(\omega)\right|_{L^p(\Omega';\mathcal{H})}^p d\mathbb{P}(\omega) &= A_p^p \int_{\Omega'} \left|\sum_{k=1}^n \varepsilon_k'(\omega') \delta f_k\right|_{L^p(\Omega;\mathcal{H})}^p d\mathbb{P}'(\omega') \\ &\leq A_p^p M_p^p(\mathcal{H}) \int_{\Omega'} \left|\sum_{k=1}^n \delta f_k\right|_{L^p(\Omega;\mathcal{H})}^p d\mathbb{P}'(\omega') = A_p^p M_p^p(\mathcal{H}) \left|\sum_{k=1}^n \delta f_k\right|_{L^p(\Omega;\mathcal{H})}^p,\end{aligned}$$

where the first equality follows from Fubini's theorem. The reverse inequality follows similarly, since both inequalities above can be reversed: for the Khintchine–Kahane estimate this is in the very statement of Corollary 3.13, and for the UMD-estimate we observe that $\epsilon \star (\epsilon \star f) = f$ (since $(\pm 1)^2 = 1$), from which it is obvious how to reverse the inequality $|\epsilon \star f|_{L^p(\Omega;X)} \leq |f|_{L^p(\Omega;X)}$.

We now have the inequalities

$$a_p M_p^{-1}(\mathcal{H})\,|f_n|_{L^p(\Omega;\mathcal{H})} \leq \left|\sqrt{\sum_{k=1}^{n} |\delta f_k(\cdot)|_{\mathcal{H}}}\right|_{L^p(\Omega)} \leq A_p M_p(\mathcal{H})\,|f_n|_{L^p(\Omega;\mathcal{H})}\,.$$

The assertion follows by taking the supremum over $n \in \mathbb{Z}_+$. □

The previous proof showed a beautiful interplay of the UMD-condition and the inequality of Khintchine and Kahane, another place where the exponent p becomes irrelevant. We emphasize the obvious fact that neither of the two results, Burkholder's theorem (if we restrict for a while to the equivalence of the various UMD-p) or the Khintchine–Kahane inequality is covered by the other; they are about two different phenomena: in the Khintchine–Kahane inequality, the random variables are the scalar valued Rademacher functions, which are multiplied by constant vectors; in Burkholder's theorem, the random variables are the vector valued martingales, which are multiplied by constant signs. It is nevertheless interesting to note the similarity of the results, although there is no clear connection.

6.6 Notes and comments

This chapter follows the seminar notes of de Pagter [5]. There exist more direct proofs of the UMD-condition on $\mathbb{R}$ or $\mathbb{C}$, but since we work in the vector-valued setting, it is natural to derive these results as special cases of the general theory.

The name of the UMD-property is related to the fact that, by Lemma 2.12(4) and Remark 7.1 in the chapter, the UMD-inequality in fact requires the unconditionality of the Schauder decomposition

$$\{\mathbb{E}(\,\cdot\,|\,\mathfrak{F}_k) - \mathbb{E}(\,\cdot\,|\,\mathfrak{F}_{k-1})\}_{k=1}^{\infty}$$

of $L^p(\mathfrak{F};X)$, whenever $\mathfrak{F}_k \uparrow \mathfrak{F}$.

The original result on the unconditionality of the martingale differences in the scalar-valued setting is also due to Burkholder (in 1966; see Diestel and Uhl [6], page 143). The vector-valued results where not yet known when [6] was written, but the potential significance of such extensions was nevertheless noted there.

It is easy to see that the square function estimate, if satisfied by the martingales on a space Y, again implies the UMD-condition; indeed,

$$|\epsilon \star f|_{\ell^\infty(\mathbb{Z}_+;L^p(\Omega;Y))} \leq s_p^{-1}\,|S(\epsilon \star f)|_{L^p(\Omega;Y)} = s_p^{-1}\,|Sf|_{L^p(\Omega;Y)} \leq s_p^{-1} S_p\,|f|_{\ell^\infty(\mathbb{Z}_+;L^p(\Omega;Y))}\,,$$

for it is clear that the square functions of f and $\epsilon \star f$ coincide. In the scalar-valued, or even in a Hilbert-space setting, it is sometimes more convenient to work with the square functions instead of the UMD-inequality (6.1); however, while the UMD-property is satisfied by a large class of important spaces, as we will see in Chapter 7, the square function estimate actually characterizes spaces isomorphic to a Hilbert space. This is a result of Kwapien (1972; see Rubio de Francia [18]).

One can show that the UMD-constants of every Hilbert space are the same, usually stated as $M_p(\mathcal{H}) = M_p(\mathbb{C})$; see Section A.9.

Chapter 7

Properties of UMD-Spaces

7.1 Introduction

In this chapter, we discover some of the basic properties of the UMD-spaces, i.e., some consequences of the uniform boundedness of martingale transforms in a Banach space. These properties are essential tools in the analysis of multipliers in UMD-spaces in the following chapters. We also obtain new examples of UMD-spaces, as well as counter-examples showing that not every Banach space is UMD.

More precisely, our goal is to establish the following results: If X is UMD, so are X^* and $L^p(\Gamma; X)$, where Γ is any σ-finite measure space, and $p \in (1, \infty)$. Furthermore, every UMD-space is reflexive.

7.2 New UMD-spaces from known ones

The purpose of this section is to demonstrate that some close relatives of a UMD-space share the UMD-property. The results follow from general principles of extensions of linear operators (presented in the Appendix, Section A.4), once we observe the following useful formulation of the UMD-condition in terms of boundedness of operators:

Remark 7.1. *X has UMD-p if and only if, for every probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, all operators*

$$\sum_{k=1}^{n} \epsilon_k \left(\mathbb{E}(\,\cdot\,|\,\mathfrak{F}_k) - \mathbb{E}(\,\cdot\,|\,\mathfrak{F}_{k-1}) \right), \tag{7.1}$$

where $\epsilon_k = \pm 1$, $(\mathfrak{F}_k)_{k=1}^{\infty}$ is an increasing sequence of sub-σ-algebras of $\mathfrak{F}$ and $\mathbb{E}(\,\cdot\,|\,\mathfrak{F}_0)$ is interpreted as the zero operator, are uniformly bounded on $L^p(\mathfrak{F}; X)$.

Now we can verify the UMD-property of a large number of spaces.

Lemma 7.2. *If X is UMD and Γ is a measure space equipped with a σ-finite measure μ, then $L^p(\Gamma; X)$, $p \in (1, \infty)$, are also UMD, and $M_p(L^p(\Gamma; X)) = M_p(X)$.*

Proof. By Remark 7.1, the property UMD-p is about uniform boundedness of the operators (7.1) consisting of linear combinations of conditional expectations. Now the X-valued and $L^p(\Gamma; X)$-valued conditional expectations are vector-valued extensions of the scalar-valued conditional expectation (see Section A.4). Furthermore, the L^p-extension result (Lemma A.15) says that the $L^p(\Gamma; X)$-valued extension is no larger in norm than the X-valued extension. The uniform boundedness in $L^p(\Omega; L^p(\Gamma; X))$ of the operators (7.1) thus follows from the uniform boundedness in $L^p(\Omega; X)$, with the same bound. $\square$

Lemma 7.3. *If X is UMD, then X^* is also UMD.*

Proof. For scalar-valued functions, we have $L^p(\Omega)^* = L^{\overline{p}}(\Omega)$ for $p \in [1, \infty)$. Furthermore, since

$$\int_\Omega g\mathbb{E}(\, h|\, \mathfrak{F})\, d\mathbb{P} = \int_\Omega \mathbb{E}(\, gh|\, \mathfrak{F})\, d\mathbb{P} = \int_\Omega \mathbb{E}(\, g|\, \mathfrak{F})\, hd\mathbb{P}$$

for $h \in L^p(\Omega)$, $g \in L^{\overline{p}}(\Omega)$ by Lemma 5.12, we see that the dual operator of $\mathbb{E}(\,\cdot|\,\mathfrak{F}) \in \mathcal{B}(L^p(\Omega))$ is $\mathbb{E}(\,\cdot|\,\mathfrak{F}) \in \mathcal{B}(L^{\overline{p}}(\Omega))$. Since $\mathbb{E}(\,\cdot|\,\mathfrak{F}) \in \mathcal{B}(L^p(\Omega; X))$ is the X-valued extension of $\mathbb{E}(\,\cdot|\,\mathfrak{F}) \in \mathcal{B}(L^p(\Omega))$, and $\mathbb{E}(\,\cdot|\,\mathfrak{F}) \in \mathcal{B}(L^{\overline{p}}(\Omega; X^*))$ is the X^*-valued extension of $\mathbb{E}(\,\cdot|\,\mathfrak{F}) \in \mathcal{B}(L^{\overline{p}}(\Omega))$, and the same is true of the operators

$$\sum_{k=1}^{n} \epsilon_k \left(\mathbb{E}(\,\cdot|\,\mathfrak{F}_k) - \mathbb{E}(\,\cdot|\,\mathfrak{F}_{k-1})\right) \tag{7.2}$$

by linearity, Lemma A.16 guarantees that the operator (7.2) acting on $L^{\overline{p}}(\Omega; X^*)$ is no larger in norm than the corresponding operator on $L^p(\Omega; X)$. Since the UMD-condition is the requirement of the uniform boundedness of these operators, we find that X^* is also UMD whenever X is. In fact, $M_{\overline{p}}(X^*) \leq M_p(X)$. □

7.3 Unbounded martingale transforms

To show that UMD-spaces are reflexive, i.e., that non-reflexivity makes the UMD-condition impossible, we must understand why a certain Banach space may fail to be UMD. It is the purpose of this section to construct a prototype of a martingale failing the UMD-inequality, and to see in what kind of spaces we can have such a martingale.

Lemma 7.4. *Let ε_j, $j \in \mathbb{Z}_+$, be Rademacher functions on Ω, and let $\mathfrak{F}_k$ be the finite algebra generated by $\{\varepsilon_j\}_{j=1}^k$ (i.e., the smallest algebra on which these functions are measurable). Let the stopping times $\tau_k : \Omega \to \mathbb{Z}_+ \cup \{\infty\}$, $k \geq 1$, be defined by*

$$\tau_k := \inf\{n \in \mathbb{Z}_+ : \sum_{j=1}^{n} \varepsilon_j^+ \geq k\}, \tag{7.3}$$

and $\tau_0 := 0$. Then

1. *$\tau_k < \infty$ a.s.,*
2. *$\{\tau_k - \tau_{k-1}\}_{k=1}^\infty$ is a sequence of independent identically distributed random variables,*
3. *$\epsilon_n = 1$ at a given sample point $\omega \in \Omega$ if and only if $n = \tau_k$ for some k at that point,*
4. *$\cup_{k=1}^n \{\tau_{k-1} < n \leq \tau_k\} = \Omega$,*
5. *$\{\tau_{k-1} < n \leq \tau_k\} \in \mathfrak{F}_{n-1}$, and*
6. *given any $m \in \mathbb{Z}_+$, it is almost certain that, for some $k \in \mathbb{Z}_+$, we have $\tau_k - \tau_{k-1} > m$.*

Proof. The random variables $\varepsilon_j^+ := \varepsilon_j \vee 0$ are independent, identically distributed and symmetrically $\{0, 1\}$-valued. It is illustrative to think of $n \in \mathbb{Z}_+$ as (discrete) time; then $\sum_{j=1}^n \varepsilon_j^+$ gives the number of 1's in n independent experiments with equally likely outcomes 0 and 1, and τ_k is the time (i.e., number of experiments) it takes to get k 1's.

The probability to have r positive results in n experiments is clearly $\binom{n}{r}2^{-n} \leq \frac{n^r}{r!}2^{-n}$, and this certainly tends to 0 for fixed r as $n \to \infty$. Thus it is clear that the probability that we only have less than k positive outcomes in a series of infinite experiments is zero.

Then $\tau_k - \tau_{k-1}$ $(k > 1)$ is well defined, and represents the time it takes to get the kth 1 after the $(k-1)$th. Due to the independence of the experiments, this is the same as the time, counted from the beginning of the process, it takes to get the first 1; i.e. $\tau_k - \tau_{k-1}$ has the same distribution as $\tau_1 = \tau_1 - \tau_0$. The fact that the various $\tau_k - \tau_{k-1}$ are independent for different values of k also follows from the independence of the ε_j^+.

Item 3 is immediate from the definition. The assertions 4 and 5 are also consequences of simple observations: Since $\tau_k = \sum_{j=1}^{\tau_k} 1 \geq \sum_{j=1}^{\tau_k} \varepsilon_j^+ = k$, we have $\cup_{k=1}^{n}\{\tau_{k-1} < n \leq \tau_k\} = \{n \leq \tau_n\} = \Omega$. From the definition of τ_k, we have $\{\tau_{k-1} < n\} = \{\sum_{j=1}^{n-1} \varepsilon_j^+ \geq k-1\} \in \mathfrak{F}_{n-1}$ and $\{n \leq \tau_k\} = \{\sum_{j=1}^{n-1} \varepsilon_j^+ < k\} \in \mathfrak{F}_{n-1}$.

For 6, note first that

$$\mathbb{P}(\tau_k - \tau_{k-1} > m) = \mathbb{P}(\tau_1 > m) = \mathbb{P}\left(\sum_{j=1}^{m} \varepsilon_j^+ = 0\right) = \mathbb{P}(\varepsilon_1^+ = \ldots = \varepsilon_m^+ = 0) = \frac{1}{2^m}.$$

Thus $\sum_{k=1}^{\infty} \mathbb{P}(\tau_k - \tau_{k-1} > m) = \infty$, and it follows from the Borel–Cantelli lemma [27] that infinitely many of the independent events $\{\tau_k - \tau_{k-1} > m\}$ take place with probability one. Thus, almost certainly, at least one such event occurs. $\square$

We next exploit the stopping times τ_k to construct a convenient martingale in view of some counterexamples.

Lemma 7.5. *The martingale $f \in L^1(\Omega; c_0)^{\mathbb{Z}_+}$, adapted to $(\mathfrak{F}_n)_{n=1}^{\infty}$, whose difference sequence is defined by*

$$\delta f_n := (-1)^n \varepsilon_n \sum_{k=1}^{\infty} \mathbf{1}_{\{\tau_{k-1} < n \leq \tau_k\}} e_k,$$

where $e_k := (\delta_{k,j})_{j=1}^{\infty} \in c_0$, satisfies $|f|_{\ell^\infty(\mathbb{Z}_+; L^\infty(\Omega; c_0))} \leq 2$.

Observe that only the n first terms in the summation are non-zero by Lemma 7.4(4); thus $f_n(\omega) \in c_0 = \{\lambda \in \ell^\infty : \lim_{i\to\infty} \lambda_i = 0\}$ as asserted.

By the monotonicity of L^p norms on probability spaces, it follows that f is also bounded in $\ell^\infty(\mathbb{Z}_+; L^p(\Omega; c_0))$ with the same bound 2. This boundedness essentially depends on the factor $(-1)^n$, which becomes apparent in the proof.

Proof. The kth component of $f_\ell = \sum_{n=1}^{\ell} \delta f_n$ is given by

$$\sum_{n=1}^{\ell} (-1)^n \varepsilon_n \mathbf{1}_{\{\tau_{k-1} < n \leq \tau_k\}} = \sum_{\tau_{k-1} < n \leq \tau_k \wedge \ell} (-1)^n \varepsilon_n.$$

By Lemma 7.4(3), $\varepsilon_n = -1$ for all the terms with $\tau_{k-1} < n < \tau_k$ above. Since these are multiplied by alternating sings $(-1)^n$, it follows that $\sum_{\tau_{k-1} < n \leq (\tau_k - 1)\wedge\ell} (-1)^n \varepsilon_n \in \{0, \pm 1\}$. If $\ell < \tau_k$, then this is all there is in the sum, and in the opposite case we only get one more term $(-1)^{\tau_k} \epsilon_{\tau_k}$ of absolute value 1. Thus, by the triangle inequality, the absolute value of the kth component of f_ℓ is at most 2, and the assertion follows, since this is true for all $k, \ell \in \mathbb{Z}_+$. $\square$

While the f above was such a well-behaving martingale, the transform $\epsilon \star f$, with $\epsilon := ((-1)^n)_{n=1}^{\infty}$, exhibits quite different properties. Indeed, the kth component of $(\epsilon \star f)_\ell$ is given by

$$\sum_{n=1}^{\ell} \varepsilon_n \mathbf{1}_{\{\tau_{k-1} < n \leq \tau_k\}} = \sum_{\tau_{k-1} < n \leq \tau_k \wedge n} \varepsilon_n = \sum_{\tau_{k-1} < n < \tau_k \wedge n} (-1) + \varepsilon_{n \wedge \tau_k} = (-1)(\tau_k - \tau_{k-1} - 1) \pm 1;$$

the absolute value of this no less than $\tau_k - \tau_{k-1} - 2$; thus on $\{\tau_k - \tau_{k-1} > m + 1\}$, we have $|(\epsilon \star f)_\ell|_{c_0} \geq m$. Since at least one of the differences $\tau_k - \tau_{k-1}$ exceeds $m+1$ a.s. (by Lemma 7.4(6)), it follows that $|(\epsilon \star f)_\ell|_{c_0} \geq m$ (a.s.) for all $\ell \in \mathbb{Z}_+$.

Fatou's lemma then yields, for $p \in (1, \infty)$,

$$\liminf_{\ell\to\infty} \int_\Omega |(\epsilon \star f)_\ell(\omega)|_{c_0}^p \, dt \geq \int_\Omega \liminf_{\ell\to\infty} |(\epsilon \star f)_\ell(\omega)|_{c_0}^p \, dt \geq m^p.$$

This being valid for arbitrary $m \in \mathbb{Z}_+$, it follows that $\lim_{\ell\to\infty} |(\epsilon \star f)_\ell|_{L^p(\Omega; c_0)} = \infty$.

The following conclusions are now immediate.

Lemma 7.6. *The following limit holds:* $\lim_{n\to\infty} M_p(\mathbb{C}^n, |\cdot|_\infty) = \infty$.
Consequently, c_0 is not UMD.

Proof. Observe that the subspace of c_0 spanned by $\{e_i\}_{i=1}^n$ is naturally identified with $(\mathbb{C}^n, |\cdot|_\infty)$. Furthermore, for any Banach space X, we have, by definition,

$$M_p(X) \geq \frac{|(\epsilon \star f)_n|_{L^p(\Omega;X)}}{|f_n|_{L^p(\Omega;X)}},$$

whenever f is an X-valued L^p-martingale. The assertions then follow, taking as f the martingale f constructed in Lemma 7.5, and observing that $\{f_k\}_{k=1}^n$ takes values in $\operatorname{span}\{e_i\}_{i=1}^n$, so it can also be interpreted as a $\mathbb{C}^n$-valued martingale. □

The sequence space c_0 gives one example of a non-UMD Banach space. To find others, we should know how the UMD-constants of embedded spaces are related. This is done in the following:

Lemma 7.7. *Let X, Y be Banach spaces and assume that $\Lambda \in \mathcal{B}(Y;X)$ has a bounded inverse $\Lambda^{-1} \in \mathcal{B}(\operatorname{ran}\Lambda; X)$. Then*

$$M_p(Y) \leq \left|\Lambda^{-1}\right|_{\mathcal{B}(\operatorname{ran}\Lambda;Y)} M_p(X)\, |\Lambda|_{\mathcal{B}(Y;X)}\,.$$

Proof. Given any martingale $f \in L^p(\Omega;Y)^{\mathbb{Z}_+}$, we have, using the fact that $\Lambda f \in L^p(\Omega;X)^{\mathbb{Z}_+}$ is also a martingale (Lemma 5.13),

$$\begin{aligned}
|(\epsilon\star f)_n|_{L^p(\Omega;Y)} &= \left|\Lambda^{-1}(\epsilon\star\Lambda f)_n\right|_{L^p(\Omega;Y)} \leq \left|\Lambda^{-1}\right|_{\mathcal{B}(\operatorname{ran}\Lambda;Y)} |(\epsilon\star\Lambda f)_n|_{L^p(\Omega;Y)} \\
&\leq \left|\Lambda^{-1}\right|_{\mathcal{B}(X;Y)} M_p(X)\, |\Lambda f_n|_{L^p(\Omega;X)} \leq \left|\Lambda^{-1}\right|_{\mathcal{B}(\operatorname{ran}\Lambda;Y)} M_p(X)\, |\Lambda|_{\mathcal{B}(Y;X)}\, |f_n|_{L^p(\Omega;X)}\,.
\end{aligned}$$

□

7.4 Characterization of reflexivity

We will give here some equivalent conditions to the reflexivity of a Banach space, which are easier to relate to the UMD-concept. We start with an auxiliary criterion concerning solvability of abstract linear equations. Observe that reflexivity is about solving, given any $x^{**} \in X^{**}$, for $x \in X$ the uncountable number of equations $\langle x^*, x\rangle = \langle x^*, x^{**}\rangle$, $x^* \in X^*$. The next criterion only involves a finite system.

Lemma 7.8 (Helly's condition). *Let X be a complex normed linear space, $x_i^* \in X^*$, $c_i \in \mathbb{C}$, $i = 1, \dots, n$, $M > 0$. Then the system of equations $\langle x_i^*, x\rangle = c_i$, $i = 1, \dots, n$, has, for every $\epsilon > 0$, a solution x (possibly depending on ϵ) with $|x|_X \leq M + \epsilon$, if and only if*

$$\left|\sum_{i=1}^n a_i c_i\right| \leq M \left|\sum_{i=1}^n a_i x_i^*\right|_{X^*} \qquad \forall a_i \in \mathbb{C}, \quad i = 1, \dots, n. \tag{7.4}$$

The same is true with $\mathbb{C}$ replaced by $\mathbb{R}$ at every occurrence, if X is a real normed linear space.

In the proof, we only consider the complex case. A proof for real scalars is the same, with $\mathbb{C}$ simply replaced by $\mathbb{R}$, as in the assertion.

Observe, for later use, that we can always normalize one of the a_i, say a_j, to unity and vary the others, yielding an equivalent Helly's condition. Indeed, all possible combinations with $a_j \neq 0$ are obtained by multiplying the normalized inequality by $|a_j|$ and then varying the other scalars, if necessary. The case $a_j = 0$ can be dealt with by taking Ka_i in place of a_i, diving by K, and considering the limit as $K \to \infty$.

A constant $M > 0$ for which Helly's condition holds, will be called **Helly's constant**.

Proof. The necessity of (7.4) for the solvability of the system is immediate, since then

$$(M+\epsilon)\left|\sum_{i=1}^{n} a_i x_i^*\right|_{X^*} \geq \left|\left\langle \sum_{i=1}^{n} a_i x_i^*, x\right\rangle\right| = \left|\sum_{i=1}^{n} a_i c_i\right|,$$

and (7.4) follows as $\epsilon \to 0$.

Assume then that (7.4) is satisfied, and also for the moment that the x_i^* are linearly independent. Consider the mapping $T : X \to \mathbb{C}^n : x \mapsto (\langle x_i^*, x\rangle)_{i=1}^n$, which is obviously linear and continuous (since each of the components are). Then the image under T of the convex set $\overline{B}(x; M+\epsilon) \subset X$ is a convex non-empty subset S of $\mathbb{C}^n$.

We now make the counterassumption that our linear system is not solvable for certain $\epsilon > 0$. This is equivalent to saying that the point $c = (c_i)_{i=1}^n \in \mathbb{C}^n$ is not in the convex set S. But then we can find a separating hyperplane through c having the whole of S on its one side, i.e., for certain $p \neq 0$ (tangent to the hyperplane) we have $\Re p \cdot y \leq \Re p \cdot c$ for all $y \in S$. Using the definition of S, this means that $\Re \sum_{i=1}^n p_i \langle x_i^*, x\rangle \leq \Re \sum_{i=1}^n p_i c_i$ for all $x \in \overline{B}(0; M+\epsilon) \subset X$. Since this holds equally well with ζx, $|\zeta| = 1$, in place of x, we actually have an inequality for the absolute values. But then

$$(M+\epsilon)\left|\sum_{i=1}^{n} p_i x_i^*\right|_{X^*} = \sup_{x\in\overline{B}(x;M+\epsilon)} \left|\left\langle \sum_{i=1}^{n} p_i x_i^*, x\right\rangle\right| \leq \left|\sum_{i=1}^{n} p_i c_i\right| \leq M \left|\sum_{i=1}^{n} p_i x_i^*\right|_{X^*}.$$

This can only be the case if $\sum_{i=1}^n p_i x_i^* = 0$, which is impossible, since $p \neq 0$ and the x_i^* are independent. This completes the proof with the additional assumption of independence of the x_i^*.

For general x_i^*, $i = 1, \ldots, n$, we can first extract a maximal linearly independent subcollection. Setting some a_i to 0 in (7.4), it is obvious that this subcollection satisfies Helly's condition as well. The previous part of the proof then gives us an $x \in X$, so that the equations $\langle x_i^*, x\rangle = c_i$ are satisfied for x_i in the independent collection chosen. Due to dependence, these conditions already determine $\langle x_i^*, x\rangle$ for the rest of the x_i^*. To see that this yields correct values, express any one of the remaining dependent functionals as $x_i^* = \sum_j a_j x_j^*$, in which case $\langle x_i^*, x\rangle = \sum_j a_j c_j$, and $\left|c_i - \sum_j a_j c_j\right| \leq \left|x_i^* - \sum_j a_j x_j^*\right|_{X^*} = 0$, so $\langle x_i^*, x\rangle = c_i$ is satisfied for all $i = 1, \ldots, n$. □

With the help of Helly's condition, we now establish the following characterizations of reflexivity:

Lemma 7.9 (James 1964). *For a Banach space X, the following conditions are equivalent:*

1. *X is not reflexive.*
2. *For some $\vartheta \in (0,1)$, there are $x_j \in \overline{B}(0;1) \subset X$ and $x_i^* \in \overline{B}(0;1) \subset X^*$, $i, j \in \mathbb{Z}_+$, such that*

$$\langle x_i^*, x_j\rangle = \begin{cases} \vartheta & \text{if } i \leq j \\ 0 & \text{if } i > j \end{cases}. \tag{7.5}$$

3. *For some $\vartheta > 0$, there exists a bounded sequence $\{x_k\}_{k=1}^\infty \subset X$ with the property that*

$$d(\operatorname{conv}\{x_k\}_{k=1}^n, \operatorname{conv}\{x_k\}_{k=n+1}^\infty) \geq \vartheta, \tag{7.6}$$

for all $n \in \mathbb{Z}_+$.

Proof. 1 ⇒ 2. Assume that X is not reflexive. Then X is a proper closed subspace of X^{**}, and we can thus pick an $x^{**} \in X^{**}$ at positive distance from X; for definiteness, fix $x^{**} \in X^{**}$ with $|x^{**}|_{X^{**}} < 1$ and also fix a $\vartheta > 0$ such that $d(x^{**}, X) > \vartheta$. Observe that this implies, since $0 \in X$, that $|x^{**}|_{X^{**}} > \vartheta$.

We construct a sequence as in condition 2 inductively; in fact, the sequence will in addition have the property $\langle x^{**}, x_i^*\rangle = \vartheta$ for all $i \in \mathbb{Z}_+$. For the initial step, since $|x^{**}|_{X^{**}} > \vartheta$, there is an $x_1^* \in X^*$ with norm at most 1 such that $|\langle x^{**}, x_1^* \rangle| = \vartheta$; multiplying x_1^* with an appropriate $\zeta \in S(0;1) \subset \mathbb{C}$ if necessary, we obtain the same equality without the absolute values. Since $|x^{**}|_{X^{**}} < 1$, we must then have $|x_1^*|_{X^*} > \vartheta$, and we can repeat the argument just given to get $x_1 \in \overline{B}(0;1) \subset X$ such that $\langle x_1^*, x_1 \rangle = \vartheta$. This completes the initial step.

Now assume that we have found $x_j \in X$ and $x_i \in X^*$, $i, j = 1, \dots, n-1$ so that the properties required in condition 2 of the lemma are satisfied for this range of indices. We should now find an $x_n^* \in \overline{B}(0;1) \subset X^*$ so that $\langle x^{**}, x_n^* \rangle = \vartheta$ and $\langle x_i, x_n^* \rangle = 0$ for $i = 1, \dots, n$. The existence of solution of this linear system of equation follows from Helly's condition (now applied to X^* and X^{**} in place of X and X^*, and using the observation that one of the scalars can be normalized to unity): Since $\left| x^{**} + \sum_{i=1}^{n-1} a_i x_i \right|_{X^{**}} \geq d(x^* *, X) > 0$, we have

$$\left| \vartheta + \sum_{i=1}^{n} a_i \cdot 0 \right| = \vartheta \leq \frac{\vartheta}{d(x^{**}, X)} \left| x^{**} + \sum_{i=1}^{n-1} a_i x_i \right|_{X^{**}}$$

for all choices of the scalars a_i. Since Helly's constant here is $M := \frac{\vartheta}{d(x^{**}, X)} < 1$, we can choose $\epsilon > 0$ such that $M + \epsilon \leq 1$ to deduce the existence of an appropriate x_n^*.

We still need to find an $x_n \in \overline{B}(0;1) \subset X$, which satisfies $\langle x_j^*, x_n \rangle = \vartheta$ for $j = 1, \dots, n$. The existence again follows from Helly's condition, using $\langle x^{**}, x_j^* \rangle = \vartheta$ and $|x^{**}|_{X^{**}} < 1$

$$\left| \sum_{j=1}^{n} a_j \vartheta \right| = \left| \left\langle x^{**}, \sum_{j=1}^{n} a_j x_j^* \right\rangle \right| \leq |x^{**}|_{X^{**}} \left| \sum_{j=1}^{n} a_j x_j^* \right|_{X^*},$$

so Helly's condition is satisfied again, with Helly's constant $|x^{**}|_{X^{**}} < 1$. This completes the induction step, and with it the proof of the implication.

$2 \Rightarrow 3$. Assume the existence of the sequences $\{x_j\}_{j=1}^{\infty}$ and $\{x_i^*\}_{i=1}^{\infty}$ as in condition 2. Then $\{x_j\}_{j=1}^{\infty}$ is an appropriate sequence also for condition 3, with the same ϑ. Indeed, let $\sum_{j=1}^{n} \lambda_j = \sum_{j=n+1}^{\infty} \mu_j = 1$, with $\lambda_j, \mu_j \geq 0$ and only finitely many of the μ_j non-zero. Then

$$\left| \sum_{j=1}^{n} \lambda_j x_j - \sum_{j=n+1}^{\infty} \mu_j x_j \right|_X \geq \left\langle x_{n+1}^*, \sum_{j=1}^{n} \lambda_j x_j - \sum_{j=n+1}^{\infty} \mu_j x_j \right\rangle = \sum_{j=1}^{n} \lambda_j \cdot 0 + \sum_{j=n+1}^{\infty} \mu_j \theta = \theta,$$

i.e., $d(\operatorname{conv}\{x_j\}_{j=1}^{n}, \operatorname{conv}\{x_j\}_{j=n+1}^{\infty}) \geq \vartheta$.

$3 \Rightarrow 1$. Assume there is a bounded sequence in X with the property (7.6). Then, given any $x \in X$, we claim that, for a large enough $n \in \mathbb{Z}_+$, $d(x, \operatorname{conv}\{x_k\}_{k=n+1}^{\infty}) \geq \frac{1}{2}\vartheta$. Indeed, if $d(x, \operatorname{conv}\{x_k\}_{k=n+1}^{\infty}) < \frac{1}{2}\vartheta$, then $|x - y|_X < \vartheta$ for some finite convex combination y of the x_k, $k > n+1$. Due to the finiteness, we can actually find an $m > n$ such that $y \in \operatorname{conv}\{x_k\}_{k=1}^{m}$, and this implies $d(y, \operatorname{conv}\{x_k\}_{k=m+1}^{\infty}) \geq \vartheta$. But then $d(x, \operatorname{conv}\{x_k\}_{k=m+1}^{\infty}) \geq d(y, \operatorname{conv}\{x_k\}_{k=m+1}^{\infty}) - |x - y|_X \geq \vartheta - \frac{1}{2}\vartheta = \frac{1}{2}\vartheta$, and this shows the claim.

Fix for the moment an $x \in X$ and $C := \overline{\operatorname{conv}}\{x_k\}_{k=n+1}^{\infty}$, where n is large enough, as above, so that $d(x, C) \geq \frac{1}{2}\vartheta$. Now $\{x\}$ is compact and C is closed, and both sets a convex. Then a version of the Hahn–Banach theorem (see [19]) implies that there exists a linear functional $x^* \in X^*$ such that $\Re \langle x^*, x_k \rangle < \alpha < \Re \langle x^*, x \rangle$ for all $k > n$ and some $\alpha \in \mathbb{R}$.

Now consider $x^{**} \in X^{**}$ defined by $\Re \langle x^{**}, y^* \rangle := \Lambda(\Re \langle y^*, x_k \rangle)_{k=1}^{\infty}$, where $\Lambda \in \mathcal{B}(\ell^{\infty}; \mathbb{R})$ is a Banach limit, i.e., a linear functional on bounded real sequences with the property $\liminf_{k \to \infty} \lambda_k \leq \Lambda\lambda \leq \limsup_{k \to \infty} \lambda_k$. (For the existence of such a functional, see [19]. Recall that it is legitimate to define a linear functional in terms of its real part, for every real-linear functional is the real part of a unique complex-linear functional.) From the definition it is immediate that the $x^{**} \in X^{**}$ satisfies $\Re \langle x^{**}, y^* \rangle \leq \limsup_{k \to \infty} \Re \langle y^*, x_k \rangle$. This property makes it impossible that $x^{**} \in X$, as we will see.

Given any $x \in X$, let x^* be the linear functional with the property $\Re \langle x^*, x_k \rangle < \alpha < \Re \langle x^*, x \rangle$ for all k greater than some n, as above. But then clearly $\Re \langle x^{**}, x^* \rangle \leq \limsup_k \Re \langle x^*, x_k \rangle \leq \alpha <$

$\Re\langle x^*, x\rangle$, and this means that $x^{**} \neq x$. Since this holds for arbitrary $x \in X$, we have constructed an element $x^{**} \in X^{**} \setminus X$, and this shows explicitly that X is not reflexive. $\square$

Corollary 7.10. *If X is a non-reflexive Banach space, then for every $\vartheta \in (0, \frac{1}{2})$, there exists a Paley–Walsh martingale $(f_k)_{k=0}^n \in L^\infty([0,1];X)^{n+1}$ satisfying pointwise*

$$|f_k|_X \leq 1 \qquad \text{and} \qquad |\delta f_k|_X \geq \vartheta$$

Proof. The martingale is constructed with the aid of the sequence $\{x_k\}_{k=0}^\infty$ provided by Lemma 7.9. Indeed, take the sequence satisfying the condition (7.6) with $2\vartheta \in (0,1)$ in place of ϑ, and set $f_n := \sum_{k=1}^{2^n} x_k \mathbf{1}_{[(k-1)2^{-n}, k2^{-n})}$. Then obviously f_n is $\mathfrak{D}_n$-measurable, and we obtain a martingale adapted to $(\mathfrak{D}_k)_{k=1}^\infty$ by defining $f_k := \mathbb{E}(\, f_n | \, \mathfrak{D}_k)$. The estimate $|f_k|_X \leq 1$ is obvious for $k = n$ (in fact with equality), since $x_k \in S(0;1)$, and for $k < n$, the point values are obtained by averaging, so the estimate follows from the triangle inequality.

The value of f_k on $[(j-1)2^{-k}, j2^{-k})$ is the average of the values $x_{2^{n-k}(j-1)+1}, \ldots, x_{2^{n-k}j}$, which f_n attains on this interval, i.e. the it is in $\operatorname{conv}\{x_i\}_{2^{(n-k)}(j-1) < i \leq 2^{n-k}j}$. It follows from the choice of the x_k that the values of f_k on two distinct basis intervals of $\mathfrak{D}_k$ differ at least by 2ϑ.

Now the value of $f_{k-1} = \mathbb{E}(\, f_k | \, \mathfrak{D}_{k-1})$ at any $I \in \operatorname{bs}\mathfrak{D}_k$ is obtained by averaging two distinct values of f_k, say $y_1 := f_k|_I$ and y_2 (attained on a neighbouring interval). Thus, at this arbitrarily chosen point, $|f_{k-1} - f_k|_X = \left|\frac{1}{2}(y_1 + y_2) - y_1\right|_X = \frac{1}{2}|y_1 - y_2|_X \geq \vartheta$. This completes the proof. $\square$

Proposition 7.11. *Every UMD-space is reflexive.*

Proof. Let X, contrary to the claim, be a non-reflexive UMD-space. Let $p \in (1, \infty)$ and $\vartheta \in (0, \frac{1}{2})$ be fixed. For $n \in \mathbb{Z}_+$, let $f := (f_k)_{k=1}^n$ be a Paley–Walsh martingale satisfying the conditions in Corollary 7.10. The idea is to show, using the existence of this special martingale, that the spaces $(\mathbb{C}^n, |\cdot|_\infty)$ are embedded in $L^p([0,1];X)$ (which is also UMD) in a manner which leads to contradiction by Lemmas 7.6 and 7.7.

We consider the operator $\Lambda_n : \mathbb{C}^n \to L^p([0,1];X)$ defined by $\Lambda_n a := (a \star f)_n$. Since X is UMD, and since f is pointwise bounded in norm by unity, we have

$$|\Lambda_n a|_{L^p([0,1];X)} = |(a \star f)_n|_{L^p([0,1];X)} \leq M_p(X)\,|a|_\infty\,|f_n|_{L^p([0,1];X)} = M_p(X)\,|a|_\infty\,.$$

On the other hand, by the pointwise boundedness of f from below, we have

$$\begin{aligned}\vartheta\,|a_k| \leq |a_k \delta f_k|_{L^p([0,1];X)} &\leq |(a \star f)_k|_{L^p([0,1];X)} + |(a \star f)_{k-1}|_{L^p([0,1];X)} \\ &\leq 2\,|(a \star f)_n|_{L^p([0,1];X)} = 2\,|\Lambda_n a|_{L^p([0,1];X)}\,.\end{aligned}$$

From Lemma 7.7 it follows that

$$M_p(L^p([0,1];X)) \geq \frac{1}{M_p(X)}\frac{\vartheta}{2} M_p(\mathbb{C}^n, |\cdot|_\infty),$$

but the right-hand side tends to infinity as $n \to \infty$ by Lemma 7.6. This is a contradiction; hence non-reflexive UMD-spaces do not exist. $\square$

Corollary 7.12. *A Banach space X is UMD if and only if its dual X^* is UMD.*

Proof. We already saw the necessity of the condition in Lemma 7.3. For the sufficiency, if X^* is UMD, then X^* is reflexive by Proposition 7.11 and X^{**} is UMD by Lemma 7.3. But if X^* is reflexive, so is X, and hence $X = X^{**}$ is UMD. $\square$

7.5 Notes and comments

This chapter follows de Pagter [5], except for the auxiliary results related to the characterization of reflexivity: Helly's condition (Lemma 7.8) is from Hille and Phillips [9], Theorem 2.7.8, and Lemma 7.9 comes from James [10]; this paper contains a remarkable total of 38 different characterizations of reflexivity.

In fact, UMD-spaces even have a property called *super-reflexivity*. On the other hand, there are reflexive, even super-reflexive, Banach spaces, which are not UMD [5].

Lemma 7.2 is *not* a direct consequence of the vector-valued Fubini's theorem, since it is not obvious in general that the point evaluations $\delta f_k(\gamma, \cdot)$, $\gamma \in \Gamma$, should constitute a martingale difference sequence in $L^p(\Omega; X)$, even if $\{\delta f_k\}_{k=1}^{\infty}$ was such a sequence in $L^p(\Omega; L^p(\Gamma; X))$.

Chapter 8

Hilbert and Riesz Transforms

8.1 Introduction

The Hilbert transform is one of the basic operators in harmonic analysis. It is defined (several equivalent definitions exist), for trigonometric polynomials on the torus $\mathbb{T}$ by the conjugation

$$H\left(\sum_{k=-n}^{n} x_k e^{\mathbf{i}2\pi k\cdot}\right) := \sum_{k=-n}^{n} -\mathbf{i}\,\mathrm{sgn}(k) x_k e^{\mathbf{i}2\pi k\cdot}, \qquad \mathrm{sgn}(k) := \begin{cases} \frac{k}{|k|} & k \neq 0 \\ 0 & k = 0 \end{cases},$$

and for $\phi = \int_{-\infty}^{\infty} \widehat{\phi}(\xi) e^{\mathbf{i}2\pi\xi t} d\xi \in \mathcal{S}(\mathbb{R}; X)$ by the analogous expression

$$H\phi := \int_{-\infty}^{\infty} -\mathbf{i}\,\mathrm{sgn}(\xi)\widehat{\phi}(\xi) e^{\mathbf{i}2\pi\xi t} d\xi.$$

The question is now whether these extend to bounded operators on $L^p(\mathbb{T}; X)$ and $L^p(\mathbb{R}; X)$, respectively. For $X = \mathbb{C}$, the L^2 case is immediate from Plancherel's formula, and a theorem of F. Riesz gives an affirmative answer for all $p \in (1, \infty)$ (see e.g. [7] for the transform on $L^p(\mathbb{R})$).

It is obvious from the formulae above that the Hilbert transform has a particularly simple multiplier structure when viewed in the Fourier domain; in fact, it corresponds to multiplication of the Fourier coefficients, respectively the Fourier transform, by $\mathbf{i} \cdot \mathbf{1}_{(-\infty,0)} - \mathbf{i} \cdot \mathbf{1}_{(0,\infty)}$. This also gives an indication of the significance of this operator in connection with the multiplier theorems.

A close variant of the Hilbert transform is the Riesz projection R, whose multiplier is $\mathbf{1}_{[0,\infty)}$ (interpreted as $\mathbf{1}_{\mathbb{N}}$ in the periodic case). This also extends conveniently to the d-dimensional setting, where we simply define R to be the operator whose multiplier is the characteristic function of the positive cone $[0, \infty)^d$.

It is easy to express the Riesz projection in terms of the Hilbert transform; indeed, since the relation between operators and their multipliers is linear, we find that $\frac{1}{2}(\mathrm{id} + \mathbf{i}H)$ has the multiplier $\mathbf{1}_{(0,\infty)}$. On $\mathbb{R}$, it is clear that open and closed intervals yield equal operators, so this is the Riesz projection. On $\mathbb{Z}$, we can perform a simple shift with the help of the pointwise multipliers of the form $e^{\mathbf{i}2\pi\nu\cdot}$, which are unitary operators on L^p. Conversely, in the non-periodic case we have $H = -\mathbf{i}R + \mathbf{i}\mathcal{R}R\mathcal{R}$, where $\mathcal{R}$ is the reflection, $\mathcal{R}f(t) := f(-t)$, which is also unitary on L^p and commutes with the Fourier transform $\mathcal{F}$. Appropriate modifications for the periodic case are easy to make.

The point of these observations is the fact that the boundedness of the Hilbert and Riesz transforms are equivalent. The Hilbert transform is sometimes more convenient in the one-dimensional setting, owing to the form of the conjugation, when written in terms of the trigonometric functions sin and cos instead of the imaginary exponentials above. The main result of this chapter concerning the boundedness of these operators will be shown for the Hilbert transform, but when extending this to the multidimensional context, we will change to the Riesz projection, and R will be our main tool also in the following chapters.

8.2 Random walk on the complex plane

Our goal is to deduce the boundedness of the Hilbert transform from the UMD-property. We first consider the transform for functions on the unit circle. The idea will be to consider an appropriate martingale representing random walk, starting from the origin and eventually (almost surely) crossing the unit circle. The bounds for the operator norm of the Hilbert transform will then be deduced from the estimates available for the auxiliary martingale due to the UMD-assumption.

We start with a lemma, in which we define the martingale to be used throughout this section and describe some of its properties.

Lemma 8.1. *Let ε_k be Rademacher functions, and $\mathfrak{F}_n$ be the σ-algebra generated by $\{\varepsilon_i\}_{i=1}^n$. Let $\eta \in (0, \frac{1}{2})$. Then the complex martingale f^η (adapted to $(\mathfrak{F}_{2k})_{k=1}^\infty$) and the stopping time τ defined by*

$$\delta f_k^\eta := \eta\,(\varepsilon_{2k-1} + \mathbf{i}\varepsilon_{2k}) \qquad \tau := \inf\{n \in \mathbb{Z}_+ : |f_n^\eta| \geq 1\}$$

(and $f_0^\eta := 0$) satisfy $\eta\,|\sqrt{\tau}|_{L^p(\Omega)} \leq C_p$, and $\tau < \infty$ (a.s.), as well as $1 \leq |f_\tau^\eta| \leq 1 + 2\eta$ (a.s.).

Thus the martingale f^η represents random walk in the complex plane, starting from the origin and (a.s.) crossing the unit circle $\mathbb{T}$ at some point. The last estimate guarantees that f^η is pretty close to the unit circle right after the first moment of crossing; this is useful in view of getting estimates of the behaviour of functions on $\mathbb{T}$. We will eventually pass to the limit $\eta \to 0$ to obtain the final estimates.

The symbols introduced in the assertion of the lemma will be used throughout this section.

Proof. Since $\{\tau \geq k\} = \{|f_{k-1}^\eta| < 1\}$ is $\mathfrak{F}_{2(k-1)}$-measurable, $w_k := \mathbf{1}_{\{\tau \geq k\}}$ defines a predictable sequence. It is easily seen that $(w \star f^\eta)_n = f_{n\wedge\tau}^\eta$. This is a bounded martingale: If $\tau = \infty$, then $|f_{n\wedge\tau}^\eta| = |f_n^\eta| < 1$ for all n, and if $\tau < \infty$, then $|f_{n\wedge\tau}^\eta| \leq |f_{n\wedge\tau-1}^\eta| + |\delta f_{n\wedge\tau}^\eta| \leq 1 + 2\eta \leq 2$.

The square function of $w \star f^\eta$ is

$$S(w \star f^\eta) = \sqrt{\sum_{k=1}^{\infty} |w_k \delta f_k^\eta|^2} = \sqrt{\sum_{k=1}^{\infty} \mathbf{1}_{\{\tau \geq k\}} 2\eta^2} = \sqrt{2}\eta \sqrt{\sum_{k=1}^{\tau} 1} = \sqrt{2}\eta\sqrt{\tau}.$$

By the square function estimate we then have

$$\sqrt{2}\eta\,|\sqrt{\tau}|_{L^p(\Omega)} = |S(w \star f^\eta)|_{L^p(\Omega)} \leq S_p\,|w \star f^\eta|_{\ell^\infty(\mathbb{Z}_+;L^p(\Omega))} \leq 2S_p$$

since $|(w \star f^\eta)_n| \leq 2$ pointwise. This was the first claim. It follows from this estimate that τ is a.s. finite; thus the estimate $|f_\tau^\eta| \leq |f_{\tau-1}^\eta| + |\delta f_\tau^\eta| \leq 1 + 2\eta$, valid on $\{\tau < \infty\}$, holds almost everywhere. ☐

We consider a trigonometric polynomial $u(e^{\mathbf{i}2\pi t}) := \sum_{j=1}^m (x_j \cos(2\pi kt) + y_j \sin(2\pi jt))$, $x_j, y_j \in X$, which will be kept fixed through this section. The conjugate polynomial of u is $v(e^{\mathbf{i}2\pi t}) := \sum_{j=1}^m (-y_j \cos(2\pi jt) + x_j \sin(2\pi jt))$; this is easily seen to agree with the definition of the conjugate using exponential functions. The aim is to show that $|v|_{L^p(\mathbb{T};X)} \leq C\,|u|_{L^p(\mathbb{T};X)}$ with the bound C depending only on p and X and not on u and v.

If we manage to do so, then the density of the trigonometric polynomials in $L^p(\mathbb{T};X)$, $p \in [1,\infty)$, allows us to deduce the boundedness of the conjugation for all L^p functions with zero average (since we omitted the constant term from u). The general case then follows as a simple application of the properties of Schauder decompositions: By Lemma 2.8, the norm $\|u + c\|_{L^p} := |u|_{L^p} \wedge |u + c|_{L^p}$ (for $\int_{\mathbb{T}} u dt = 0$, c constant) is equivalent to the original norm $|\cdot|_{L^p}$; in particular, $\|\cdot\|_{L^p} \leq K\,|\cdot|_{L^p}$ Thus, if H, the Hilbert transform or conjugation, is bounded on functions of zero average, then $|H(u + c)|_{L^p} = |Hu|_{L^p} \leq C\,|u|_{L^p} \leq CK\,|u + c|_{L^p}$, and H is bounded on L^p. (We used the fact that H maps constants to zero in the first step, but a slight modification of the argument shows that this property is not crucial.)

We extend u and v to harmonic C^∞ functions on $\mathbb{C}$ in the standard way:

$$u(re^{\mathbf{i}2\pi t}) := \sum_{j=1}^{m} r^j (x_j \cos(2\pi jt) + y_j \sin(2\pi jt)) \tag{8.1}$$

and similarly for v. We take the freedom to regard these functions of the complex variable $z = re^{\mathbf{i}2\pi t}$ also as functions of two real variables, $\Re z$ and $\Im z$, and partial derivatives of u and v will always refer to this interpretation. One easily verifies the harmonicity conditions $(D_1^2 + D_2^2)u = (D_1^2 + D_2^2)v = 0$, and the Cauchy–Riemann type coupling between the two: $D_1 u = D_2 v$, $D_2 u = -D_1 v$. (If x_j and y_j were real (rather than vectors of X), then $u + \mathbf{i}v$ would be a complex analytic function.) Also note that $u(0) = v(0) = 0$.

We already observed that f_τ^η is (a.s.) in the vicinity of the unit circle. In order to estimate u on $\mathbb{T}$, we will estimate $u(f_\tau^\eta)$ (and similarly v); for this purpose, we write this quantity as a sum of a difference sequence:

$$u(f_{\tau\wedge n}^\eta) = u(f_{\tau\wedge n}^\eta) - u(f_0^\eta) = \sum_{k=1}^{n\wedge\tau} (u(f_k^\eta) - u(f_{k-1}^\eta)). \tag{8.2}$$

This is not yet a martingale difference sequence, but Taylor's theorem will show that it is not too far away from one:

Lemma 8.2. *On $\{|f_{k-1}^\eta| < 1\} = \{\tau \geq k\}$, we have the estimate, for harmonic $u : \mathbb{C} \to X$ and f as in Lemma 8.1:*

$$\left|\left(u(f_k^\eta) - u(f_{k-1}^\eta)\right) - \left(\eta\varepsilon_{2k-1} D_1 u(f_{k-1}^\eta) + \eta\varepsilon_{2k} D_2 u(f_{k-1}^\eta) + 2\eta^2 D_{1,2} u(f_{k-1}^\eta)\right)\right|_X \leq C(u)\eta^3,$$

where $C(u)$ depends on the size of the derivatives of u in $\overline{B}(0;2)$.

Proof. Taylor's theorem gives

$$u(z + h) = u(z) + h \cdot Du(z) + (h \cdot D)^2 u(z) + R(z, h),$$

with

$$|R(z,h)|_X \leq \frac{|h|_\infty^3}{3!} \sup_{\lambda\in[0,1]} \left| \left(\frac{h}{|h|_\infty} \cdot D\right)^3 u(z + \lambda h) \right|_X.$$

The assertion follows now from simple observations: Since $f_k^\eta = f_{k-1}^\eta + h$ with $h := \eta\varepsilon_{2k-1} + \mathbf{i}\eta\varepsilon_{2k}$, we have $(h \cdot D)^2 u = \eta^2 (D_1^2 + D_2^2 + 2\varepsilon_{2k-1}\varepsilon_{2k} D_{1,2}) u = 2\eta^2 \varepsilon_{2k-1}\varepsilon_{2k} D_{1,2} u$ by harmonicity. For $|f_{k-1}^\eta| < 1$, we have $|f_{k-1}^\eta + \lambda h| \leq 1 + 2\eta \leq 2$ (recall $\eta \in (0, \frac{1}{2})$), and X-norms of the partial derivatives of u of order three, each continuous on the compact set $\overline{B}(0;2)$, attain a maximum on this set. For $C(u)$ we can then take the sum of these maxima (times a universal constant). $\square$

We are getting closer to estimating u on $\mathbb{T}$ in terms of martingales. The following lemma will settle this matter:

Lemma 8.3. *For harmonic $u : \mathbb{C} \to X$, the sequence*

$$\bar{u}_n := \sum_{k=1}^{n} \eta \mathbf{1}_{\{\tau\geq k\}} \left(D_1 u(f_{k-1}^\eta)\varepsilon_{2k-1} + D_2 u(f_{k-1}^\eta)\varepsilon_{2k}\right), \qquad n \in \mathbb{Z}_+,$$

defines a martingale, adapted to $(\mathfrak{F}_{2k})_{k=1}^\infty$, which satisfies $|u(f_{\tau\wedge n}^\eta) - \bar{u}_n|_{L^p(\Omega;X)} \leq C(p,u)\eta$, for $p \in (1,\infty)$, where the $C(p,u)$ depends on the size of derivatives of u in $B(0;2)$.

Proof. To see that $\bar{u}$ is a martingale, observe that $\mathbf{1}_{\{\tau\geq k\}}$ and f_{k-1}^η (thus $D_i u(f_{k-1}^\eta)$) are $\mathfrak{F}_{2(k-1)}$-measurable; hence

$$\mathbb{E}\left(\mathbf{1}_{\{\tau\geq k\}} D_i u(f_{k-1}^\eta)\varepsilon_{2k-\delta} \middle| \mathfrak{F}_{2(k-1)}\right) = \mathbf{1}_{\{\tau\geq k\}} D_i u(f_{k-1}^\eta) \mathbb{E}\left(\varepsilon_{2k-\delta} \middle| \mathfrak{F}_{2(k-1)}\right) = 0,$$

since $\varepsilon_{2k-\delta}$, $\delta = 0, 1$, has zero average and is independent of $\varepsilon_1, \ldots, \varepsilon_{2(k-1)}$, thus of $\mathfrak{F}_{2(k-1)}$, which is generated by the $2(k-1)$ Rademacher functions enumerated. This shows that $(\delta\bar{u}_n)_{n=1}^{\infty}$ is a martingale difference sequence; thus $\bar{u}$ is a martingale.

Using (8.2) and Lemma 8.2, we have

$$\begin{aligned}
&\left| u(f^{\eta}_{n\wedge\tau}) - \bar{u}_n - 2\eta^2 \sum_{k=1}^{n\wedge\tau} \varepsilon_{2k-1}\varepsilon_{2k} D_{1,2}u(f^{\eta}_{k-1}) \right|_X \\
&= \left| \sum_{k=1}^{n\wedge\tau} \left(u(f^{\eta}_k) - u(f^{\eta}_{k-1}) - D_1 u(f^{\eta}_{k-1})\varepsilon_{2k-1} - D_2 u(f^{\eta}_{k-1})\varepsilon_{2k} - 2\eta^2 \varepsilon_{2k-1}\varepsilon_{2k} D_{1,2}u(f^{\eta}_{k-1}) \right) \right|_X \\
&\leq \sum_{k=1}^{n\wedge\tau} C(u)\eta^3 \leq C(u)\eta^3\tau. \qquad (8.3)
\end{aligned}$$

From Lemma 8.1 we know that $\left(\int_\Omega \tau^{\frac{p}{2}} d\mathbb{P}\right)^{\frac{1}{p}} \leq C_p\eta^{-1}$, and substituting $2p$ in place of p yields $|\tau|_{L^p} \leq (C_p\eta^{-1})^2$. Thus the L^p norm of the right-hand side of (8.3) is bounded by $C(p,u)\eta$, and to estimate the L^p norm of $u(f^{\eta}_{n\wedge\tau})$, we need to estimate the summation on the left-hand side of (8.3) involving mixed second derivatives of u. An appropriate bound follows from the square function estimate, once we recall that $D_{1,2}u(f^{\eta}_{k-1}) = \sum_{j=1}^{m} \left(x_j\phi_j(f^{\eta}_{k-1}) + y_j\varphi_j(f^{\eta}_{k-1}) \right)$, where the ϕ_j and φ_j are the appropriate partial derivatives of $r^j\cos(2\pi jt)$ and $r^j\sin(2\pi jt)$, respectively. Now $(\mathbf{1}_{\{\tau\geq k\}}\varepsilon_{2k-1}\varepsilon_{2k}\phi(f^{\eta}_{k-1}))_{k=1}^{\infty}$ is a martingale difference sequence adapted to $(\mathfrak{F}_{2k})_{k=1}^{\infty}$ (which is seen essentially in the same way as for $\delta\bar{u}_n$ above), and we can compute

$$\begin{aligned}
\left| 2\eta^2 \sum_{k=1}^{n\wedge\tau} \varepsilon_{2k-1}\varepsilon_{2k}\phi(f^{\eta}_{k-1}) \right|_{L^p(\Omega)} &\leq s_p^{-1}(\mathbb{R}) \left| \sqrt{\sum_{k=1}^{n\wedge\tau} \left(2\eta^2\varepsilon_{2k-1}\varepsilon_{2k}D_{1,2}\phi(f^{\eta}_{k-1})\right)^2} \right|_{L^p(\Omega)} \\
&\leq 2\eta^2 s_p^{-1}(\mathbb{R}) \, |\phi|_{L^\infty(B(0;1))} \left| \sqrt{\sum_{k=1}^{n\wedge\tau} 1} \right|_{L^p(\Omega)} \leq 2\eta^2 s_p^{-1}(\mathbb{R}) C(u) \left|\sqrt{\tau}\right|_{L^p(\Omega)} \leq C(p,u)\eta,
\end{aligned}$$

where the last inequality was part of Lemma 8.1. Since $\sum_{k=1}^{n\wedge\tau} 2\eta^2\varepsilon_{2k-1}\varepsilon_{2k}u(f^{\eta}_{k-1})$ is a linear combination of a finite (fixed) number of terms of the form just estimated, multiplied by constant vectors x_j, y_j, an estimate of the similar form (with possibly larger $C(p,u)$) also follows for this sum, and the proof is complete. $\square$

Lemma 8.3 at our hands, it now seems reasonable to try to bound the martingale $\bar{v}$ in terms of $\bar{u}$. We should note that it is only now that we invoke the UMD-property. The first result is of technical nature.

Lemma 8.4. *Let X be a UMD-space, and $r_k : \Omega \to X$ be $\mathfrak{F}_{k-1}$ measurable for $k \in \mathbb{Z}_+$. Then $\sum_{k=1}^{n} \varepsilon_k r_k$, $n \in \mathbb{Z}_+$, defines a martingale on Ω satisfying*

$$M_p^{-p}(X) \left| \sum_{k=1}^{n} \varepsilon_k r_k \right|^p_{L^p(\Omega;X)} \leq \int_{\Omega^2} \left| \sum_{k=1}^{n} \varepsilon_k(\omega_1) r_k(\omega_2) \right|_X d\mathbb{P}^2(\omega_1,\omega_2) \leq M_p^p(X) \left| \sum_{k=1}^{n} \varepsilon_k r_k \right|^p_{L^p(\Omega;X)}.$$

Proof. The fact that $(\varepsilon_k r_k)_{k=1}^{\infty}$ is a martingale difference sequence is immediate. By Fubini's theorem, the double integral in the assertion can be evaluated as

$$\begin{aligned}
\int_\Omega d\mathbb{P}(\omega_2) \int_\Omega d\mathbb{P}(\omega_1) \left| \sum_{k=1}^{n} \varepsilon_k(\omega_1) r_k(\omega_2) \right|_X &= \int_\Omega d\mathbb{P}(\omega_2) \int_\Omega d\mathbb{P}(\omega_1) \left| \sum_{k=1}^{n} \varepsilon_k(\omega_2)\varepsilon_k(\omega_1) r_k(\omega_2) \right|_X \\
&= \int_\Omega d\mathbb{P}(\omega_1) \int_\Omega d\mathbb{P}(\omega_2) \left| \sum_{k=1}^{n} \varepsilon_k(\omega_1)\varepsilon_k(\omega_2) r_k(\omega_2) \right|_X,
\end{aligned}$$

where the first equality used the equidistributedness property of $(\varepsilon_k)_{k=1}^n$ and $(\epsilon_k\varepsilon_k)_{k=1}^n$, $\epsilon_k \in \{-1,1\}$; above we had $\epsilon_k = \varepsilon_k(\omega_2)$.

Since X is UMD, the inner integral in this last expression can be estimated from above by $M_p^p(X)\int_\Omega |\sum_{k=1}^n \varepsilon_k(\omega_2) r_k(\omega_2)|_X \, d\mathbb{P}(\omega_2)$ and from below by a similar expression with $M_p^{-p}(X)$ in place of $M_p^p(X)$. (Recall that the UMD-inequality is always two-sided.) □

It is now straightforward to estimate $|\bar{v}_n|_{L^p}$ by $|\bar{u}_n|_{L^p}$:

Lemma 8.5. *Let X be a UMD-space, $u,v : \mathbb{C} \to X$ conjugate polynomials and $\bar{u},\bar{v}$ corresponding martingales as in Lemma 8.3. Then $|\bar{v}_n|_{L^p(\Omega;X)} \le M_p^2(X)\,|\bar{u}_n|_{L^p(\Omega;X)}$.*

Proof. Recall that $\bar{u}_n = \sum_{k=1}^n \eta \mathbf{1}_{\{\tau\ge k\}}\left(\varepsilon_{2k-1}D_1u(f_{k-1}^\eta) + \varepsilon_{2k}D_2u(f_{k-1}^\eta)\right)$ and

$$\begin{aligned}\bar{v}_n &= \sum_{k=1}^n \eta\mathbf{1}_{\{\tau\ge k\}}\left(\varepsilon_{2k-1}D_1v(f_{k-1}^\eta) + \varepsilon_{2k}D_2v(f_{k-1}^\eta)\right)\\ &= \sum_{k=1}^n \eta\mathbf{1}_{\{\tau\ge k\}}\left(-\varepsilon_{2k-1}D_2u(f_{k-1}^\eta) + \varepsilon_{2k}D_1v(f_{k-1}^\eta)\right),\end{aligned}$$

where the last equality follows from the Cauchy–Riemann equations.

Observe that $\eta\mathbf{1}_{\{\tau\ge k\}}D_1u(f_{k-1}^\eta)$ and $\eta\mathbf{1}_{\{\tau\ge k\}}D_2u(f_{k-1}^\eta)$ are $\mathfrak{F}_{2(k-1)}$-measurable, so that $\bar{u}_n$, written in the form $\sum_{k=1}^{2n}\varepsilon_k r_k$ is a sum to which Lemma 8.4 applies; the same is also true of $\bar{v}_n$. Thus

$$\begin{aligned}|\bar{v}_n|_{L^p(\Omega;X)}^p &= \left|\sum_{k=1}^n \eta\mathbf{1}_{\{\tau\ge k\}}\left(-\varepsilon_{2k-1}D_2u(f_{k-1}^\eta) + \varepsilon_{2k}D_1v(f_{k-1}^\eta)\right)\right|_{L^p(\Omega;X)}^p\\ &\le M_p^p(X)\int_\Omega d\mathbb{P}(\bar\omega)\int_\Omega d\mathbb{P}(\omega)\left|\sum_{k=1}^n \eta\mathbf{1}_{\{\tau(\bar\omega)\ge k\}}\left(-\varepsilon_{2k-1}(\omega)D_2u(f_{k-1}^\eta(\bar\omega)) + \varepsilon_{2k}(\omega)D_1v(f_{k-1}^\eta(\bar\omega))\right)\right|_X^p\\ &= M_p^p(X)\int_\Omega d\mathbb{P}(\bar\omega)\int_\Omega d\mathbb{P}(\omega)\left|\sum_{k=1}^n \eta\mathbf{1}_{\{\tau(\bar\omega)\ge k\}}\left(\varepsilon_{2k}(\omega)D_2u(f_{k-1}^\eta(\bar\omega)) + \varepsilon_{2k-1}(\omega)D_1v(f_{k-1}^\eta(\bar\omega))\right)\right|_X^p.\end{aligned}$$

In the last equality we used the invariance of the joint distribution of the Rademacher functions with respect to changes of signs to remove the minus, and also the invariance with respect to permutations. The last expression in the inequality above can now be estimated by $M_p^{2p}(X)\,|\bar{u}_n|_{L^p(\Omega;X)}^p$ by applying the other side of Lemma 8.4.

This completes the proof. □

Now it is time to combine the estimates, and complete our random walk towards the unit circle.

Proposition 8.6. *The conjugation of trigonometric polynomials extends to a bounded linear transformation, the Hilbert transform, on $L^p(\mathbb{T};X)$, $p \in (1,\infty)$, whenever X is a UMD-space.*

Proof. As noted above, it suffices to show that $|v|_{L^p(\mathbb{T};X)} \le C\,|u|_{L^p(\mathbb{T};X)}$ for u as in (8.1) and v the harmonic conjugate, C only depending on p and X. Combining Lemmas 8.3 and 8.5 we know that

$$\begin{aligned}|v(f_{n\wedge\tau}^\eta)|_{L^p(\Omega;X)} &\le |\bar{v}_n|_{L^p(\Omega;X)} + C(p,v)\eta \le M_p^2(X)\,|\bar{u}_n|_{L^p(\Omega;X)} + C(p,v)\eta\\ &\le M_p^2(X)\,|u(f_{n\wedge\tau}^\eta)|_{L^p(\Omega;X)} + \left(M_p^2(X)C(p,u) + C(p,v)\right)\eta.\end{aligned}$$

As $n\to\infty$, $f_{n\wedge\tau}^\eta \to f_\tau^\eta$ (a.s.); since $|f_{n\wedge\tau}^\eta| \le 1+2\eta \le 2$, and u and v are bounded in $\overline{B}(0;2)$, it follows from the dominated convergence theorem that we can drop the n's from the two sides of the above inequality, to deduce $|v(f_\tau^\eta)|_{L^p(\Omega;X)} \le M(p,X)\,|u(f_\tau^\eta)|_{L^p(\Omega;X)} + C(p,u,X)\eta$, where we dropped the dependence of the constant on v, since v is anyway determined by u.

For $\zeta \in \mathbb{T}$ (the unit circle), it is easy to see that $u_\zeta(z) := u(\zeta z)$ and v_ζ defined similarly satisfy the same properties as u and v, so that we also have the inequalities

$$|v(f_\tau^\eta \zeta)|_{L^p(\Omega;X)} \le M(p,X)\,|u(f_\tau^\eta \zeta)|_{L^p(\Omega;X)} + C(p,u,X)\eta, \qquad \zeta \in \mathbb{T}.$$

Observe that we have the same constant $C(p,u,X)$ here, whatever ζ, since the factor $C(p,u,X) = M_p^2(X)C(p,u) + C(p,v)$ only depends on the size of the derivatives of u on $B(0;2)$ (Lemma 8.3), and the u_ζ are just rotations of u around the origin.

Raising the previous inequality to the power of p and using $\left(\frac{1}{2}(a+b)\right)^p \le \frac{1}{2}(a^p+b^p)$, we have (with new constants)

$$\int_\Omega |v(f_\tau^\eta(\omega)\zeta)|_X^p \, d\mathbb{P}(\omega) \le M(p,X) \int_\Omega |u(f_\tau^\eta(\omega)\zeta)|_X^p \, d\mathbb{P}(\omega) + C(p,u,X)\eta, \qquad \zeta \in \mathbb{T}. \tag{8.4}$$

We can clearly integrate this inequality over $\zeta \in \mathbb{T}$ (with normalized Lebesgue measure); the two double integrals appearing can then be written as

$$\int_\Omega d\mathbb{P}(\omega) \int_\mathbb{T} d\zeta\, |v(f_\tau^\eta(\omega)\zeta)|_X^p = \int_\Omega d\mathbb{P}(\omega) \int_\mathbb{T} d\zeta\, |v(|f_\tau^\eta(\omega)|\,\zeta)|_X^p\,,$$

where the equality follows from the fact that $f_\tau^\eta(\omega)\zeta$ and $|f_\tau^\eta(\omega)|\,\zeta$, where ζ ranges over $\mathbb{T}$, just give two different parametrizations of the circle centered at the origin, one point of which is $f_\tau^\eta(\omega)$.

Recall from Lemma 8.1 that $1 \le |f_\tau^\eta(\omega)| \le 1 + 2\eta$ for almost all $\omega \in \Omega$, where $\eta \in (0, \frac{1}{2})$ is the parameter adjusting the size of each step in our random walk, described by the martingale f. We now consider the limit as this parameter tends to zero: Then $|f_\tau^\eta(\omega)| \to 1$ for almost all ω, and the continuity of v implies that $v(|f_\tau^\eta(\omega)|\,\zeta) \to v(\zeta)$. Since v is bounded in $B(0;2)$, the dominated convergence theorem says that

$$\int_\Omega d\mathbb{P}(\omega) \int_\mathbb{T} d\zeta\, |v(|f_\tau^\eta(\omega)|\,\zeta)|_X^p \underset{\eta\to 0}{\longrightarrow} \int_\mathbb{T} |v(\zeta)|_X^p \, d\zeta = |v|_{L^p(\mathbb{T};X)}^p\,.$$

A similar result clearly holds for u in place of v. Thus the inequality (8.4) (integrated over $\mathbb{T}$) and the convergence just established yield

$$|v|_{L^p(\mathbb{T};X)}^p \le M(p,X)\,|u|_{L^p(\mathbb{T};X)} + \limsup_{\eta\to 0} C(p,u,X)\eta,$$

and obviously this last limit is zero, so we get the desired estimate with a factor only depending on p and (the UMD-constant of) X as asserted. $\square$

The extension from the unit-circle to the d-dimensional torus $\mathbb{T}^d$ follows from the general extension result of Lemma A.15.

Lemma 8.7. *If the Riesz projection R is bounded on $L^p(\mathbb{T};X)$, X a Banach space, then it is also bounded on $L^p(\mathbb{T}^d;X)$, and $|R|_{L^p(\mathbb{T}^d;X)} \le |R|_{L^p(\mathbb{T};X)}^d$.*

Proof. The Riesz transform on $L^p(\mathbb{T};X)$ is obviously an X-valued extension of the scalar-valued Riesz transform on $L^p(\mathbb{T})$. By Lemma A.15, R then also has an $L^p(\mathbb{T}^{d-1};X)$-valued extension R_d which is bounded in the operator norm by $|R|_{L^p(\mathbb{T};X)}$. This extension is, by definition, an operator on $L^p(\mathbb{T};L^p(\mathbb{T}^{d-1};X))$ with multiplier $\mathbf{1}_\mathbb{N}$. Alternatively, it can be viewed as an operator on $L^p(\mathbb{T}^d;X)$, whose multiplier is the indicator of the positive half-space $\mathbb{Z}^{d-1}\times\mathbb{N}$. The operators R_j whose multipliers are indicators of other similar half-spaces $\mathbb{Z}^{j-1}\times\mathbb{N}\times\mathbb{Z}^{d-j}$ are obtained from R_d by rotations (which are unitary on every L^p). The assertion follows, since the Riesz projection R on $L^p(\mathbb{T}^d;X)$ is the product $R = \prod_{j=1}^d R_j$. $\square$

8.3 From torus to $\mathbb{R}^d$

To establish the boundedness of the Hilbert transform on $L^p(\mathbb{R}^d;X)$, we will exploit the result already obtained for the torus, and the Poisson summation formula (Lemma A.47)

$$\sum_{k\in\mathbb{Z}^d}\varphi(t+\lambda k)=\lambda^{-d}\sum_{k\in\mathbb{Z}^d}\widehat{\varphi}(\lambda^{-1}k)e^{\mathbf{i}2\pi\lambda^{-1}k\cdot t}$$

valid for $\varphi\in\mathcal{S}(\mathbb{R}^d;X)$, $\lambda\in\mathbb{R}\setminus\{0\}$ and $t\in\mathbb{R}^d$.

The following result, from which we readily derive the desired boundedness of the Hilbert transform, is useful in other estimates, too. It gives the possibility to examine L^p properties of $\varphi\in\mathcal{S}(\mathbb{R}^d;X)$ by means of a sequence of functions on the torus, defined by

$$T_j^p\varphi:=2^{-\frac{dj}{p}}\sum_{k\in\mathbb{Z}^d}\widehat{\varphi}(k2^{-j})e^{\mathbf{i}2\pi k\cdot(\cdot)}. \tag{8.5}$$

Recalling the Fourier inversion formula $\varphi(t)=\int_{\mathbb{R}^d}\widehat{\varphi}(\xi)e^{\mathbf{i}2\pi\xi\cdot t}d\xi$, it seems reasonable to regard $T_j^p\varphi$ as a discrete approximation of φ. The following result gives this interpretation more quantitative content.

Lemma 8.8. *For $\varphi\in\mathcal{S}(\mathbb{R}^d;X)$ and $p\in[1,\infty)$, we have*

$$|\varphi|_{L^p(\mathbb{R}^d;X)}=\lim_{j\to\infty}\left|T_j^p\varphi\right|_{L^p(\mathbb{T}^d;X)}$$

Proof. It is here convenient to take our fundamental domain (see Section 1.2) to be $[-\frac{1}{2},\frac{1}{2})^d$, a symmetric region around the origin. Making first a change of variable and then applying the Poisson summation formula, we find that

$$\begin{aligned}\int_{[-\frac12,\frac12)^d}\left|\sum_{\kappa\in\mathbb{Z}^d}\widehat{\varphi}(2^{-j}\kappa)e^{\mathbf{i}2\pi\kappa\cdot t}\right|_X^p dt&=\int_{2^j[-\frac12,\frac12)^d}\left|\sum_{\kappa\in\mathbb{Z}^d}\widehat{\varphi}(2^{-j}\kappa)e^{\mathbf{i}2\pi2^{-j}\kappa\cdot t}\right|_X^p 2^{-jd}dt\\&=\int_{2^j[-\frac12,\frac12)^d}\left|2^{jd}\sum_{\kappa\in\mathbb{Z}^d}\varphi(t+2^j\kappa)\right|_X^p 2^{-jd}dt,\end{aligned}$$

and the powers of two give a total of $2^{jd(p-1)}$; hence, after taking the pth root and multiplying by $2^{-\frac{jd}{p}}$ the exponent is $(jd)(1-\frac{1}{p}-\frac{1}{p})=0$, explaining the choice of the factor in (8.5). Thus

$$\left|T_j^p\varphi\right|_{L^p(\mathbb{T}^d;X)}=\left|\sum_{\kappa\in\mathbb{Z}^d}\varphi(t+2^j\kappa)\right|_{L^p(2^j[-\frac12,\frac12);X)}.$$

Considering only the term in the summation with $\kappa=0$, it is clear that

$$\int_{2^j[-\frac12,\frac12)^d}|\varphi(t)|_X^p\,dt\underset{j\to\infty}{\longrightarrow}|\varphi|^p_{L^p(\mathbb{R}^d;X)}.$$

It remains to show that the rest of the summation vanishes in the limit. This seems reasonable, since $\varphi\in\mathcal{S}$ decreases rapidly away from the origin, e.g., we may estimate $|\phi(t)|_X\le C(1+|t|^2)^{-\frac12 N}\le C\,|t|_\infty^{-N}$, where N can be taken as large as desired, with an appropriate choice of C. Then, for $t\in2^j[-\frac12,\frac12)^d$,

$$\begin{aligned}\int_{2^j[-\frac12,\frac12)^d}\sum_{\kappa\in\mathbb{Z}^d\setminus\{0\}}|\varphi(t+2^j\kappa)|_X^p\,dt&\le2^{jn}\sum_{\kappa\in\mathbb{Z}^d\setminus\{0\}}\left(\frac{C}{(2^j(|\kappa|_\infty-\frac12))^N}\right)^p\\&=2^{j(n-Np)}\sum_{\kappa\in\mathbb{Z}^d\setminus\{0\}}\frac{C^{-p}}{(|\kappa|_\infty-\frac12)^{Np}}.\end{aligned}$$

For $Np > n$, the summation converges to a finite limit, whereas the factor $2^{j(n-Np)} \to 0$ as $j \to \infty$. This completes the proof. □

Now the boundedness of the Riesz projection in the non-periodic case is almost immediate:

Corollary 8.9. *If the Riesz projection is bounded on $L^p(\mathbb{T}^d; X)$, where X is a Banach space, then it is also bounded on $L^p(\mathbb{R}^d; X)$, and $|R|_{L^p(\mathbb{R}^d;X)} \leq |R|_{L^p(\mathbb{T}^d;X)}$.*

Proof. Consider $\varphi \in \mathcal{D}(\mathbb{R}^d; X)$. Then $T_j^p \varphi$ is a trigonometric polynomial, and

$$RT_j^p\varphi := 2^{-\frac{dj}{p}} \sum_{\kappa\in\mathbb{N}^d} \mathcal{F}\varphi(\kappa 2^{-j}) e^{\mathbf{i}2\pi\kappa\cdot(\cdot)} = 2^{-\frac{dj}{p}} \sum_{\kappa\in\mathbb{Z}^d} \mathcal{F}(R\varphi)(\kappa 2^{-j}) e^{\mathbf{i}2\pi\kappa\cdot(\cdot)} = T_j^p(R\varphi),$$

i.e., the operators T_j^p "commute" with the Riesz projection. (Strictly speaking, the operator R means two different things on the two sides of the above equality, but the point is nevertheless clear.) Thus

$$\begin{aligned} |R\varphi|_{L^p(\mathbb{R}^d;X)} &= \lim_{j\to\infty} \left|T_j^p R\varphi\right|_{L^p(\mathbb{T}^d;X)} = \lim_{j\to\infty} \left|RT_j^p\varphi\right|_{L^p(\mathbb{T}^d;X)} \\ &\leq |R|_{L^p(\mathbb{T}^d;X)} \lim_{j\to\infty} \left|T_j^p\varphi\right|_{L^p(\mathbb{T}^d;X)} = |R|_{L^p(\mathbb{T}^d;X)} |\varphi|_{L^p(\mathbb{R}^d;X)} . \end{aligned}$$

This shows the assertion. □

8.4 Notes and comments

Section 8.2 follows de Pagter [5]. The proof given for the boundedness of the Hilbert transform is originally due to Burkholder (1983). The transference result of Lemma 8.8 is from Zimmermann [29].

McConnell's [14] results on multiplier transformations lead to another proof of the boundedness of the Hilbert transform in UMD-spaces, which works directly for the non-periodic case. The proof is similar in spirit to the one given here in that it uses Brownian motion in the upper half-space $\mathbb{R}^d \times \mathbb{R}_+$ to deduce properties of functions on $\mathbb{R}^d$; however, this proof needs some fairly deep results from the theory of stochastic processes falling beyond the level of our treatment.

The facts that the boundedness of the Riesz projection on $\mathbb{T}$ implies its boundedness on $\mathbb{T}^d$ and on $\mathbb{R}^d$ are special cases of a more general theorem: If R is bounded on the unit-circle for some Banach space X, then it is also bounded on any ordered locally compact group with respect to an appropriate positive cone (Ben de Pagter, personal communication).

Proposition 8.6 has a converse of equal significance: The boundedness of the Hilbert transform or the Riesz projection also implies the UMD-condition, so the boundedness of these transformations actually characterizes UMD-spaces. This converse result is proved by Bourgain (see Burkholder [2]). We have omitted this proposition due to limited space; the implication we established is sufficient to obtain multiplier theorems in UMD-spaces, but it is good to observe that the UMD-condition is actually necessary for the boundedness of the Hilbert transform.

Because of the equivalence, the name "spaces of class $\mathcal{HT}$" is also used for UMD, e.g. by Hieber and Prüss [8].

There are other characterizations of these spaces, and in somewhat older literature, one finds the name ζ-convex, which also means the same thing. The condition of ζ-convexity simply requires the existence of a biconvex function $\zeta : X \times X \to \mathbb{R}$ satisfying

$$\zeta(0,0) > 0 \qquad \text{and} \qquad \zeta(x,y) \leq |x+y|_X \quad \text{if} \quad |x|_X = |y|_X = 1.$$

Equivalence proofs are found in Burkholder [2].

The condition of ζ-convexity appears simpler than the UMD-condition; however, this simplicity is somewhat misleading. Even if we know that a certain Banach space is ζ-convex, it is in practice almost impossible to actually find an appropriate function ζ explicitly. For this reason, this notion, which gives an oversimplified idea of the subject, should perhaps be avoided (Jan Prüss, personal communication).

Usually the Hilbert transform is known as the (principal value) integral operator (see Duoandikoetxea [7] for equivalence proofs)

$$Hf(t) := \frac{1}{\pi} \lim_{\epsilon \downarrow 0} \int_{|s|>\epsilon} \frac{f(t-s)}{s} ds.$$

One can show ([7], Theorem 5.17), for a general Banach space X, that a class of Calderón–Zygmund integral operators, which includes the Hilbert transform, are bounded on all $L^p(\mathbb{R}; X)$, $p \in (1, \infty)$, if they are bounded on one. This shows that the boundedness of the Hilbert transform is also independent of the exponent p, just like the UMD-condition. In fact, the proof in [7] for the assertion concerning the Calderón–Zygmund operators also goes via a weak-type L^1 estimate.

Chapter 9

Multipliers in UMD-Spaces

9.1 Introduction

We know by now that the UMD-property implies the boundedness of the Riesz projection R, whose multiplier is the characteristic function of the positive cone $\mathbb{N}^d$ or $\mathbb{R}^d_+$. From this, our intention is to deduce the boundedness of several other multiplier operators. The present chapter deals with scalar-valued Fourier multipliers acting on functions whose range is in a UMD-space. These results were developed prior to the recognition of the significance of R-boundedness, but the abstract framework perhaps streamlines the theory (see also Section 9.5). Furthermore, these results pave the way for the study of operator-valued multipliers, where the notion of R-boundedness becomes essential.

9.2 Simple multipliers

For a while, we simultaneously analyze some multiplier operators on both the periodic and non-periodic cases. The elementary results are similar in both cases, and the proofs also follow the same lines.

We immediately get the boundedness of all multipliers related to the translates of the positive cone; in fact, we even have R-boundedness as a direct consequence of Example 4.11 and the product rule of R-bounds, since $e^{\mathbf{i}2\pi\nu\cdot(\cdot)}Re^{-\mathbf{i}2\pi\nu\cdot(\cdot)}$ is the operator whose multiplier is the characteristic function of the positive cone translated from the origin to ν.

In the periodic case, similarly as before, T_λ, $\lambda \in \mathbb{C}^{\mathbb{Z}^d}$, denotes the linear operator acting on trigonometric polynomials, the action of which on monomials is given by $T_\lambda(e^{\mathbf{i}2\pi\kappa\cdot(\cdot)}x) := \lambda_\kappa e^{\mathbf{i}2\pi\kappa\cdot(\cdot)}$. It is easy to see, as in the proof of Corollary 2.14, that a necessary requirement for the boundedness of T_λ is that $\lambda \in \ell^\infty(\mathbb{Z}^d)$, but the converse need not be true in general. In the non-periodic case, T_g, $g \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, denotes the operator $\mathcal{F}^* m_g \mathcal{F}$, where m_g is simply multiplication by g. The action of such an operator is well defined at least on the class $\mathcal{F}^{-1}\mathcal{D}(\mathbb{R}^d;X)$, which is dense in $L^p(\mathbb{R}^d;X)$, $p \in [1,\infty)$ (Lemma A.39). Indeed, for $\psi \in \mathcal{F}^{-1}\mathcal{D}$, $\widehat{\psi} \in \mathcal{D}$ by definition; in particular, $\widehat{\psi}$ and thus $g\widehat{\psi}$ have compact support, so the local integrability of g implies $g\widehat{\psi} \in L^1(\mathbb{R}^d;X)$, and the transform $\mathcal{F}^*$ can be taken even in the ordinary L^1 sense.

The set of multipliers λ giving rise to bounded operators on $L^p(\mathbb{T}^d;X)$ is denoted by $\mathcal{M}^p(\mathbb{T}^d;X)$, and we define $|\lambda|_{\mathcal{M}^p(\mathbb{T}^d;X)} := |T_\lambda|_{\mathcal{B}(L^p(\mathbb{T}^d;X))}$. In the non-periodic case, $\mathcal{M}^p(\mathbb{R}^n;X)$ is defined similarly in the obvious way. Because of the fundamental role of the cones, we denote by ϱ the sequence with $\varrho_\kappa = \mathbf{1}_{\mathbb{N}^d}(\kappa)$. Furthermore, let ϱ^ν be the cone translated to $\nu \in \mathbb{Z}^d$, i.e., $\varrho^\nu_\kappa := \varrho_{\kappa-\nu}$. (When dealing with sequences indexed by vectors as here, we try to avoid ambiguity by using "the κth *term* of λ" for λ_κ, and "the ith *coordinate* of κ" for κ_i.) It should cause no confusion to use the same symbol ϱ^ν also for the function $\mathbf{1}_{\{t:t\geq\nu\}}$ on $\mathbb{R}^d$, since anyway the values of this ϱ^ν agree with the previous one in the lattice points $\kappa \in \mathbb{Z}^d$.

From the R-boundedness of the translated Riesz projectors, we can deduce the boundedness of any multiplier with finitely many non-zero terms, and we actually obtain a rather satisfying quantitative bound. Consider first the multiplier corresponding to a box $[\alpha;\beta) := \{\xi : \alpha \le \xi < \beta\}$.

Lemma 9.1. *The expression*

$$\sum_{\theta\in\{0,1\}^d} (-1)^{|\theta|_1} \varrho^{\alpha+\theta(\beta-\alpha)} \tag{9.1}$$

coincides with the indicator of the box $[\alpha;\beta)$.

Observe that the assertion is simultaneously stated in $\mathbb{Z}^d$ and in $\mathbb{R}^d$. In the discrete case, it is usually more convenient to consider "closed" boxes, but the present form of the lemma allows unified treatment most easily. Of course, in the discrete case, $[\alpha;\beta) = [\alpha;\beta-\iota]$, $\iota := (1,1,\ldots,1)$.

For simplicity, we give the proof with the discrete case terminology. The modifications for the case of $\mathbb{R}^d$ are obvious and only notational.

Proof. The assertion follows from a simple inclusion–exclusion argument. Indeed, for $\kappa \not\ge \alpha$, the κth term is clearly 0, since all the terms in the summation vanish. For $\kappa \ge \alpha$, assume for definiteness that exactly r coordinates of κ satisfy $\alpha_i \le \kappa_i < \beta_i$. Then among those θ for which $|\theta|_1 = k$, there are $\binom{d-r}{k}$ for which these r coordinates are 0, i.e., for which $\varrho_\kappa^{\alpha+\theta(\beta-\alpha)} = 1$. It follows that the κth coordinate is given by

$$\sum_{\theta\in\{0,1\}^d} (-1)^{|\theta|_1} \varrho_\kappa^{\nu+\theta} = \sum_{k=0}^{d-r} (-1)^k \binom{d-r}{k} = \begin{cases} 0 & d-r>0 \\ 1 & d-r=0 \end{cases},$$

i.e., the only non-zero terms of 9.1 are those, all coordinates of which (i.e., d of them) are bounded by the corresponding coordinates of α and β; these non-zero terms are all 1. But this is just what we claimed. ∎

Corollary 9.2. *For X a UMD-space, the set of all multiplier operators whose multipliers are indicators of boxes $[\alpha;\beta]$, $\alpha,\beta \in \mathbb{Z}^d$, is R-bounded on $L^p(U;X)$, $U = \mathbb{T}^d, \mathbb{R}^d$ and $p \in (1,\infty)$, with R-bound at most $8 \cdot 2^d |R|_{\mathcal{B}(L^p(U;X))}$.*

Proof. From the representation of an arbitrary box as (9.1), and

$$T_{\varrho^\nu} = e^{\mathbf{i}2\pi\nu\cdot(\cdot)} R e^{-\mathbf{i}2\pi\nu\cdot(\cdot)}, \tag{9.2}$$

it follows that every operator T corresponding to a multiplier of a box satisfies

$$T \in 2^d \operatorname{abco}\{m_\phi R m_\psi : |\phi|_{L^\infty(\mathbb{T}^d)}, |\psi|_{L^\infty(\mathbb{T}^d)} \le 1\}.$$

Now Lemma 4.12(2) together with Example 4.11 and the product rule for R-bounds give the desired bound for the set of operators considered. ∎

We now concentrate for a while on the periodic case, and turn to somewhat more general multipliers. Above we considered indicators of boxes, of which a particular example is a single point. Using (9.1) we can thus represent the naturals basis e^ν, with $e^\nu_\kappa := \delta_{\nu\kappa}$, in terms of the cones by

$$e^\nu = \sum_{\theta\in\{0,1\}^d} (-1)^{|\theta|_1} \varrho^{\nu+\theta}.$$

Since each sequence λ has the natural expression $\lambda = \sum_{\nu\in\mathbb{Z}^d} \lambda_\nu e^\nu$, we obtain, at least formally, the following alternative representation:

$$\lambda = \sum_{\nu\in\mathbb{Z}^d} \lambda_\nu \sum_{\theta\in\{0,1\}^d} (-1)^{|\theta|_1} \varrho^{\nu+\theta} = \sum_{\nu\in\mathbb{Z}^d} \left(\sum_{\theta\in\{0,1\}^d} (-1)^{|\theta|_1} \lambda_{\nu-\theta} \right) \varrho^\nu. \tag{9.3}$$

If λ has only finitely many non-zero terms, then the summations above are finite, the manipulations are valid, and the expression on the right is correct.

This representation motivates the following definition:

Definition 9.3. *For $\lambda \in \ell^\infty(\mathbb{Z}^d)$, the **variation** is defined by*

$$\operatorname{var}\lambda := \sum_{\nu\in\mathbb{Z}^d}\left|\sum_{\theta\in\{0,1\}^d}(-1)^{|\theta|_1}\lambda_{\nu-\theta}\right|.$$

With the help of the preceding considerations, we can easily prove the following result:

Lemma 9.4. *Let $\Lambda \subset \ell^\infty(\mathbb{Z}^d)$ be of uniformly bounded variation, and each $\lambda \in \Lambda$ consist of only finitely many non-zero terms. Then*

$$\mathcal{R}_p(\{T_\lambda\}_{\lambda\in\Lambda}) \leq 8\,|R|_{\mathcal{B}(L^p(\mathbb{T}^d;X))}\sup_{\lambda\in\Lambda}\operatorname{var}\lambda$$

Proof. Using (9.2) with (9.3), we have

$$T_\lambda = \sum_{\nu\in\mathbb{Z}^d}\left(\sum_{\theta\in\{0,1\}^d}(-1)^{|\theta|_1}\lambda_{\nu-\theta}\right)e^{\mathbf{i}2\pi\nu\cdot(\cdot)}Re^{-\mathbf{i}2\pi\nu\cdot(\cdot)}$$
$$\in \operatorname{var}\lambda\cdot\operatorname{abco}\{m_\phi R m_\psi : |\phi|_{L^\infty(\mathbb{T}^d)}, |\psi|_{L^\infty(\mathbb{T}^d)} \leq 1\}, \quad (9.4)$$

where $\operatorname{var}\lambda$ can further be estimated by the supremum of the relevant variations, and the rest follows by the same results as in the proof of Corollary 9.2. □

These results were rather straightforward consequences of the boundedness of the Riesz projection. They nevertheless serve as a starting point for a more thorough study of multiplier theorems valid in the UMD setting. In particular, we would like to allow for more general multipliers than only those with finitely many non-zero components.

9.3 Kernels and convolutions

To exploit the UMD-property, it is convenient to consider the fundamental domain $[0,1)^d$ of the torus $\mathbb{T}^d$ decomposed by means of products of the dyadic algebras $\mathfrak{D}_k$, $k\in\mathbb{N}$, on $[0,1)$. For $n = rd+j$, $r\in\mathbb{N}$, $j\in\{1,\dots,d\}$, we denote $\mathfrak{P}_n := \mathfrak{D}_{r+1}^j \times \mathfrak{D}_r^{d-j}$, and for $n=0$, $\mathfrak{P}_0 := \{\emptyset,[0,1)^d\}$.

The kernel related to the expression (5.3) of the conditional expectation as an integral operator will be of interest here. Due to the product nature of the algebra $\mathfrak{P}_n$, the kernel from (5.3) also factors into components depending on one coordinate only:

$$\mathbb{E}(\,f|\,\mathfrak{P}_n)\,(t) = \int_{\mathbb{T}^d} f(s)\left(\prod_{\ell=1}^d 2^{r+u_{j\ell}}\sum_{i=0}^{2^{r+u_{j\ell}}-1}\mathbf{1}_{I^2_{r+u_{j\ell};i}}(t_\ell,s_\ell)\right)ds =: \int_{\mathbb{T}^d} f(s)k_n(t,s)ds.$$

Here $u_{j\ell} := 1$ for $j\geq\ell$ and 0 otherwise, and $I_{r;i} := [i2^{-r},(i+1)2^{-r})$. Note that this same formula even applies to $\mathfrak{P}_0$, with $r=-1$, $j=d$.

We will actually need a slightly different version of this kernel. To motivate the modification to be given below, recall from Chapter 1 that the notion of multipliers emerged from the study of translation invariant bounded operators. The conditional expectation $\mathbb{E}(\,\cdot|\,\mathfrak{P}_n)$ does not commute with translations (except for $n=0$), but the modified operator below will.

Lemma 9.5. *The operator*

$$f \mapsto \int_{\mathbb{T}^d}\tau_h\mathbb{E}(\,\tau_{-h}f|\,\mathfrak{P}_n)\,dh$$

admits the convolution representation $f\mapsto f*g_n$, where

$$\widehat{g}_n(\kappa) = \prod_{\ell=1}^j \operatorname{sinc}^2(2^{-(r+1)}\pi\kappa_\ell)\cdot\prod_{\ell=j+1}^d\operatorname{sinc}^2(2^{-r}\pi\kappa_\ell).$$

Here $\operatorname{sinc}(t) := \frac{\sin(t)}{t}$ for $t \neq 0$ and $\operatorname{sinc}(0) := 1$. This is a C^∞ functions whose zeros are at $n\pi$, $n \in \mathbb{Z} \setminus \{0\}$.

Proof. Using the kernel representation for the conditional expectation, this is a straightforward computation:

$$\begin{aligned}
&\int_{\mathbb{T}^d} \tau_h \mathbb{E}(\,\tau_{-h} f | \mathfrak{P}_n)\,(t) dh = \int_{\mathbb{T}^d} dh \int_{\mathbb{T}^d} ds f(s+h) k_n(t-h, s) \\
&= \int_{\mathbb{T}^d} dh \int_{\mathbb{T}^d} ds' f(t-s') k_n(t-h, t-s'-h) = \int_{\mathbb{T}^d} ds' f(t-s') \int_{\mathbb{T}^d} dh k_n(h-t, h-t+s') \\
&\qquad = \int_{\mathbb{T}^d} ds' f(t-s') \int_{\mathbb{T}^d} dh' k_n(-h', s'-h').
\end{aligned}$$

In the second to last equality we used the even symmetry $k_n(t,s) = k_n(-t,-s)$, which follows from the symmetry of the dyadic algebras $\mathfrak{D}_k$. The last form in the above chain of equalities is clearly the convolution of f, evaluated at t, with

$$\begin{aligned}
g_n(s) &:= \int_{\mathbb{T}^d} k_n(-h, s+h) dh = \prod_{\ell=1}^{d} 2^{r+u_{j\ell}} \sum_{i=0}^{2^{r+u_{j\ell}}-1} \int_{\mathbb{T}^d} \mathbf{1}_{I_{r+u_{j\ell};i}}(-h_\ell) \mathbf{1}_{I_{r+1;i}}(s_\ell - h_\ell) dh \\
&= \prod_{\ell=1}^{d} 4^{r+u_{j\ell}} \int_{\mathbb{T}^d} \mathbf{1}_{-I_{r+u_{j\ell},0}}(h_\ell) \mathbf{1}_{I_{r+u_{j\ell};0}}(s_\ell - h_\ell) dh = \prod_{\ell=1}^{d} 4^{r+u_{j\ell}} \mathbf{1}_{-I_{r+u_{j\ell};0}} * \mathbf{1}_{I_{r+u_{j\ell};0}}(s_\ell),
\end{aligned}$$

where a change of variable was performed to get the second equality; we also moved the minus sign from the argument of the indicator function to the parameter set. It is easy to compute $\widehat{\mathbf{1}}_{\pm I_{r;0}}(k) = \frac{1}{\pi k} e^{\mp i\pi k 2^{-r}} \sin(\pi k 2^{-r})$, and since the Fourier coefficient of a convolution is the product of the individual coefficients, we have

$$\widehat{g}_n(\kappa) = \prod_{\ell=1}^{d} 4^{r+u_{j\ell}} \frac{\sin^2(2^{-r-u_{j\ell}} \pi \kappa_\ell)}{(\pi\kappa_\ell)^2} = \prod_{\ell=1}^{d} \operatorname{sinc}^2(2^{-r-u_{j\ell}} \pi \kappa_\ell),$$

as we claimed. □

We will establish multiplier theorems related to the following **dyadic decomposition** of $\mathbb{Z}^d$: $D_0 := \{0\}$ and

$$D_{rd+j} := (-2^{r+1}, 2^{r+1})^{j-1} \times \{(-2^{r+1}, -2^r] \cup [2^r, 2^{r+1})\} \times (-2^r, 2^r)^{d-j}, \qquad r \in \mathbb{N}, \quad j \in \{1, \ldots, d\}. \tag{9.5}$$

Later on, a decomposition of $\mathbb{R}^d$ (strictly, of $\mathbb{R}^d \setminus \{0\}$) will be considered, given by the same equation (9.5), but allowing $r \in \mathbb{Z}$. (When considering intervals of integers above, $\pm[2^r, 2^{r+1})$ contains no integers for $r < 0$.)

It is easy to see that

$$\bigcup_{k \le rd+j} D_k = (-2^{r+1}, 2^{r+1})^j \times (-2^r, 2^r)^{d-j}, \tag{9.6}$$

when the decomposition of $\mathbb{Z}^d$ is considered. For the decomposition of $\mathbb{R}^d$, we must subtract $\{0\}$ from the right-hand side.

We denote by S_k the Fourier multiplier operator whose multiplier is $\mathbf{1}_{D_k}$. Again, it should not cause confusion to use this same notation in both the periodic and the non-periodic case. The following result deals with both cases simultaneously, and it is the main reason for introducing the D_k and S_k.

Lemma 9.6. *$(S_k)_{k=0}^\infty$ is a Schauder decomposition of $L^p(\mathbb{T}^d; X)$, and $(S_n, S_{-n})_{n=0}^\infty$ of $L^p(\mathbb{R}^d; X)$, $p \in (1, \infty)$.*

Proof. We must verify the criteria given in Lemma 2.6. Since $D_k \cap D_\ell = \emptyset$ for $k \neq \ell$, and the square of an indicator is the indicator itself, it is clear that $S_k S_\ell = \delta_{k\ell} S_k$. In the discrete case, $\operatorname{ran}(S_k)_{k=0}^\infty$ consists of the trigonometric polynomials, which are dense in $L^p(\mathbb{T}^d;X)$. In the case of $\mathbb{R}^d$, $\operatorname{ran}(S_n, S_{-n})_{n=0}^N$ consists of $\varphi \in \mathcal{S}(\mathbb{R}^d;X)$ whose Fourier transform vanishes outside $\cup_{k \le N} D_k$ and inside $\cup_{k<-N} D_k$. As $N \to \infty$, we find that $\operatorname{ran}(S_n, S_{-n})_{n=0}^\infty$ consists of those $\varphi \in \mathcal{F}^{-1}\mathcal{D}(\mathbb{R}^d;X)$, which satisfy $0 \notin \operatorname{supp} \mathcal{F}\varphi$, and such functions are dense in $L^p(\mathbb{R}^d;X)$, $p \in (1,\infty)$, by Lemma A.39. All that remains to show is the uniform boundedness of the partial sums $\sum_{k=0}^K S_k$, respectively $\sum_{n=-N}^N S_n$, but this follows from Lemma 9.2, since the $\sum_{k \le K} S_k$ is the multiplier operator corresponding to a box like the one in (9.6), and $\sum_{-N \le n \le N} S_n$ is the difference of the indicators of two such boxes. $\square$

The decompositions are actually unconditional, but establishing this fact requires some more effort. Once this is done, the abstract machinery of Chapter 2 can be applied. The following result gives us convenient auxiliary functions to work with these decompositions, based on the kernels studied above.

Corollary 9.7. *For X a UMD-space and $p \in (1,\infty)$, the functions h_n, $n \in \mathbb{N}$, characterized by $h_0 = 1$ and*

$$\widehat{h}_n(\kappa) = \prod_{\ell=1}^{j-1} \operatorname{sinc}^2(2^{-(r+1)}\pi\kappa_\ell) \cdot (\sin\cdot\operatorname{sinc})^2(2^{-(r+1)}\pi\kappa_j) \cdot \prod_{\ell=j+1}^{d} \operatorname{sinc}^2(2^{-r}\pi\kappa_\ell), \qquad n = rd + j,$$

with $r \in \mathbb{N}$, $j = 1, \dots, d$, satisfy the inequality

$$\left| \sum_{k=0}^{n} \epsilon_k (h_k * f) \right|_{L^p(\mathbb{T}^d;X)} \le M_p(X)\, |f|_{L^p(\mathbb{T}^d;X)},$$

for all $\epsilon_k = \pm 1$ and all $f \in L^p(\mathbb{T}^d;X)$.

Note that, for $\kappa \in D_{n-d} = D_{(r-1)d+j}$, we have $|\kappa_\ell| < 2^{r-1}$ for $\ell > j$, and thus $2^{-r}\pi\kappa_\ell \in (-\frac{1}{2}\pi, \frac{1}{2}\pi)$, so that sinc of this quantity is strictly away from zero. The same argument holds for $\ell < j$, with r replaced by $r+1$. For $\ell = j$ we have $|\kappa_j| \in [2^{r-1}, 2^r)$, so that $\left|2^{-(r+1)}\pi\kappa_j\right| \in [\frac{1}{4}\pi, \frac{1}{2}\pi)$, and sin is strictly away from zero on this interval, thus also $\sin\cdot\operatorname{sinc}$. This property, together with the inequality in the assertion of the lemma, is the main reason for introducing the functions h_n.

Proof. Taking the difference of two operators of the kind considered in Lemma 9.5 we find that $f \mapsto \int_{\mathbb{T}^d} \tau_h \left(\mathbb{E}(\tau_{-h} f | \mathfrak{P}_n) - \mathbb{E}(\tau_{-h} f | \mathfrak{P}_{n-1}) \right)$, where $n = rd + j$, is a convolution operator with kernel $h_n = g_n - g_{n-1}$. The Fourier coefficients are then given by $\widehat{h}_n(\kappa) = \widehat{g}_n(\kappa) - \widehat{g}_{n-1}(\kappa)$, i.e.,

$$\prod_{\ell=1}^{j-1} \operatorname{sinc}^2(2^{-(r+1)}\pi\kappa_\ell) \cdot \left(\frac{\sin^2(2^{-(r+1)}\pi\kappa_j)}{(2^{-(r+1)}\pi\kappa_j)^2} - \frac{\sin^2(2\cdot 2^{-(r+1)}\pi\kappa_j)}{4\cdot(2^{-(r+1)}\pi\kappa_j)^2} \right) \cdot \prod_{\ell=j+1}^{d} \operatorname{sinc}^2(2^{-r}\pi\kappa_\ell).$$

Applying the trigonometric identities $\sin^2 t - \frac{1}{4}\sin^2(2t) = \sin^2 t - \sin^2 t \cdot \cos^2 t = \sin^4 t$, the difference term can be written as $\sin^4(2^{-(r+1)}\pi\kappa_j)(2^{-(r+1)}\pi\kappa_j)^{-2} = (\sin\cdot\operatorname{sinc})^2(2^{-(r+1)}\pi\kappa_j)$, so that $\widehat{h}_n$ here coincides with the one in the assertion.

The asserted inequality then follows from the UMD-property:

$$\begin{aligned}
\left|\sum_{k=0}^{n} \epsilon_k (h_k * f)\right|_{L^p(\mathbb{T}^d;X)} &= \left|\epsilon_0 (g_0 * f) + \sum_{k=1}^{n} \epsilon_k (g_k - g_{k-1}) * f\right|_{L^p(\mathbb{T}^d;X)} \\
&= \left|\int_{\mathbb{T}^d} \tau_h \left(\epsilon_0 \mathbb{E}(\tau_{-h} f | \mathfrak{P}_0) + \sum_{k=1}^{n} \epsilon_k \left\{\mathbb{E}(\tau_{-h} f | \mathfrak{P}_k) - \mathbb{E}(\tau_{-h} f | \mathfrak{P}_{k-1})\right\}\right) dh\right|_{L^p(\mathbb{T}^d;X)} \\
&\le \int_{\mathbb{T}^d} \left|\epsilon_0 \mathbb{E}(\tau_{-h} f | \mathfrak{P}_0) + \sum_{k=1}^{n} \epsilon_k \left\{\mathbb{E}(\tau_{-h} f | \mathfrak{P}_k) - \mathbb{E}(\tau_{-h} f | \mathfrak{P}_{k-1})\right\}\right|_{L^p(\mathbb{T}^d;X)} dh \\
&\le M_p(X) \int_{\mathbb{T}^d} |\tau_{-h} f|_{L^p(\mathbb{T}^d;X)}\, dh = M_p(X)\, |f|_{L^p(\mathbb{T}^d;X)} .
\end{aligned}$$

The second equality simply re-expressed the convolutions with the conditional expectation operators, and the last inequality was a consequence of the UMD-property (Remark 7.1). The proof is then complete. □

We have now translated the UMD-inequality concerning conditional expectation to one involving convolutions. Bearing in mind that convolution means multiplication in the domain of Fourier coefficients, and recalling the observation following the statement of Corollary 9.7 that the Fourier coefficients $\widehat{h}_n(\kappa)$ are away from zero for $\kappa \in D_{n-d}$ for all $n > d$, we find that it is possible to define convolution operators whose kernels have Fourier coefficients $\widehat{h}_n^{-1}(\kappa)$, which can be used to invert conditional expectations in a certain limited sense. If we can do so, we should be able to break functions into appropriate pieces by taking conditional expectations, then use the UMD-estimates available, and collect the pieces with the aid of the inverse operations. The ultimate goal is to deduce the unconditionality of the Schauder decomposition $(S_k)_{k=1}^{\infty}$, which then yields multiplier theorems.

We now turn to the details. Define the inverse kernels $\bar{h}_n$ by

$$\bar{h}_n := \sum_{\kappa \in D_{n-d}} \widehat{h}_n^{-1}(\kappa) e^{\mathbf{i}2\pi\kappa\cdot(\cdot)}, \qquad n > d.$$

Observe in particular that the κth Fourier coefficient of $\bar{h}_n * h_n * f$ is given by $\widehat{\bar{h}}_n \cdot \widehat{h}_n \cdot \widehat{f}(\kappa) = \widehat{f}(\kappa)\mathbf{1}_{D_{n-d}}(\kappa)$ for $n > d$.

What we need is boundedness, in fact, R-boundedness, of the convolution operators whose kernels are $\bar{h}_n$, or equivalently, of the multiplier operators T_n, related to the multipliers $\lambda_\kappa^n = \widehat{\bar{h}}_n(\kappa)$. For this purpose we invoke Lemma 9.4; however, we first need to estimate the variation of the multiplier sequences. A tool for doing so is provided in the following:

Lemma 9.8. *Let $\lambda \in \ell^\infty(\mathbb{Z}^d)$ be given, on a certain box $[\alpha;\beta] := \{\nu : \alpha \le \nu \le \beta\}$, by a function φ with continuous partial derivatives of order θ, $|\theta|_\infty \le 1$, so that $\lambda_\kappa = \varphi(\kappa)$ for $\kappa \in [\alpha;\beta]$, and $\lambda_\kappa = 0$ for $\kappa \notin [\alpha;\beta]$. Then*

$$\operatorname{var} \lambda \le \sum_{\theta \in \{0,1\}^d} \max_{t \in [\alpha;\beta]} \left|D^\theta \varphi(t)\right| (\beta - \alpha)^\theta$$

Proof. Each single term appearing in the definition of variation can be written, by the fundamental theorem of calculus, in the form

$$\sum_{\theta \in \{0,1\}^d} (-1)^{|\theta|_1} \varphi(\nu - \theta) = \int_{[\nu-\iota;\nu]} D^\iota \varphi(t) dt, \qquad \iota := (1,1,\ldots,1).$$

On the boundary, where $\nu - \theta$ falls outside $[\alpha;\beta]$ for some θ, it is convenient first to drop the zero terms from the summation and only then apply the fundamental theorem. If, for definiteness,

$I \subset \{1, \ldots, n\}$, $\#I = r$, contains the coordinate directions to which we can move from ν a step of -1 staying inside $[\alpha, \beta]$, then (see Section A.5, in particular around Lemma A.21, for notation)

$$\sum_{\theta \in \{0,1\}^d} (-1)^{|\theta|_1} \varphi(\nu - \theta) = \sum_{\theta \in \{0,1\}^r} (-1)^{|\theta|_1} \varphi_I(\nu_I - \theta; \nu) = \int_{[\nu_I - \iota_I; \nu_I]} D_I \varphi_I(t; \nu) dt \le \max_{t \in [\alpha; \beta]} |D_I \varphi(t)| .$$

Here $D_I = D^\theta$, where $\theta_i := 1$ if $i \in I$ and 0 otherwise.

Now we need to count the total contribution from all the terms in the expression of variation, observing different contributions from the boundary of the box $[\alpha; \beta]$. The set of those lattice points in the box from which we can move a step of -1 exactly to each direction of $i \in I$, $\#I = r$, consists of the points, whose ith coordinate is between $\alpha_i + 1$ and β_i for $i \in I$, and equals α_i for $i \in I^c$. Thus the total number of such points is $\prod_{i \in I} (\beta_i - \alpha_i) = (\beta - \alpha)^\theta$, where $\theta_i := 1$ for $i \in I$ and 0 otherwise.

Therefore, for each $\theta \in \{0,1\}^d$, the expression of variation contains $(\beta - \alpha)^\theta$ terms, which can be estimated by $\max_t \left| D^\theta \varphi(t) \right|$, and no other terms. The assertion follows by summing the estimates for different values of θ. $\square$

Corollary 9.9. *The multiplier sequences corresponding to the convolution operators $\bar{h}_n * \cdot$ are of uniformly bounded variation. Thus $\mathcal{R}_p(\{\bar{h}_n * \cdot\}_{n=0}^\infty) \le C \, |R|_{L^p(\mathbb{T}^d; X)}$, where C is a universal constant.*

Proof. The second assertion is immediate from Lemma 9.4, with C the supremum of the variations of the multiplier sequences times 4, once we prove that the supremum is finite.

The multiplier sequence of $\bar{h}_n * \cdot$, which consists of the Fourier coefficients of $\bar{h}_n$, vanishes outside $D_{n-d} = D_{(r-1)d+j}$, and is given in the dyadic block by $\hat{\bar{h}}_n(\kappa) = \varphi_n(\kappa)$, where

$$\varphi_n(t) := \prod_{\ell=1}^{j-1} \operatorname{sinc}^{-2}(2^{-(r+1)} \pi t_\ell) \cdot (\sin \cdot \operatorname{sinc})^{-2} (2^{-(r+1)} \pi t_j) \cdot \prod_{\ell=j+1}^{d} \operatorname{sinc}^{-2}(2^{-r} \pi t_\ell).$$

It is easy to see that the variation operator satisfies the triangle inequality. Since each dyadic block $D_{(r-1)d+j}$ defined in (9.5) is a union of two boxes $(-2^r + 1, 2^r)^{j-1} \times \left(\pm[2^{r-1}, 2^r)\right) \times (-2^{r-1}, 2^{r-1})^{d-j}$, and since φ_n has even symmetry, we only need to consider the variation of the multiplier sequence in one of the boxes, say the one with plus sign above. We then invoke Lemma 9.8 to estimate this variation. We first estimate $\sum_{\theta \in \{0,1\}^d}$ by $2^d \max_{\theta \in \{0,1\}^d} = 2^d \max_{\theta_1 \in \{0,1\}} \cdots \max_{\theta_d \in \{0,1\}}$. Since φ also factors into components, each depending on a single coordinate only, we need to estimate separately the quantities

$$\max_{t_\ell \in [2^{r-u_{\ell j}}, 2^{r-u_{\ell j}}]} \left(\operatorname{sinc}^{-2}(2^{-(r+u_{j\ell})} \pi t_\ell) \vee \left\{ (\operatorname{sinc}^{-2})'(2^{-(r+u_{j\ell})} \pi t_\ell) \cdot 2^{-(r+u_{j\ell})} \pi \cdot 2 \cdot 2^{r-u_{\ell j}} \right\} \right),$$

for $\ell \ne j$, and

$$\max_{t_j \in [2^{r-1}, 2^r]} \left(((\sin \cdot \operatorname{sinc})^{-2})'(2^{-(r+1)} \pi t_j) \vee \left\{ ((\sin \cdot \operatorname{sinc})^{-2})'(2^{-(r+1)} \pi t_j) \cdot 2^{-(r+1)} \pi \cdot 2^{r-1} \right\} \right).$$

After a change of variable these reduce to

$$\max_{t \in [-\frac{\pi}{2}, \frac{\pi}{2}]} \left(\operatorname{sinc}^{-2}(t) \vee (\operatorname{sinc}^{-2})'(t) \cdot \pi \right) \quad \text{and} \quad \max_{t \in [\frac{\pi}{4}, \frac{\pi}{2}]} \left((\sin \cdot \operatorname{sinc})^{-2}(t) \vee ((\sin \cdot \operatorname{sinc})^{-2})'(t) \cdot \frac{\pi}{4} \right),$$

which are finite quantities, since the functions sinc and $\sin \cdot \operatorname{sinc}$ are non-vanishing C^∞ mappings on the compact intervals considered. Since we no longer have dependence on n in these last estimates, the assertion is established. $\square$

Now we are ready to prove the result which gives us access to various multiplier theorems.

Lemma 9.10. *Let X be a UMD-space. Then $(S_k)_{k=0}^\infty$ is an unconditional Schauder decomposition of $L^p(\mathbb{T}^d; X)$, $p \in (1, \infty)$.*

Proof. We already know (by Lemma 9.6) that $(S_k)_{k=0}^{\infty}$ is a Schauder decomposition. According to Lemma 3.7, we must then show that both $(S_k)_{k=0}^{\infty}$ and $(S_k^*)_{k=0}^{\infty}$ are random unconditionals (Definition 3.3) on $L^p(\mathbb{T}^d;X)$ and its dual, respectively.

Since X is UMD, thus reflexive, we know that $(L^p(\mathbb{T}^d;X))^* = L^{\overline{p}}(\mathbb{T}^d;X^*)$. Furthermore, from the orthogonality of the trigonometric monomials it follows that

$$\begin{aligned}\langle g, S_k f\rangle_{L^p(\mathbb{T}^d;X)} &= \int_{\mathbb{T}^d} \langle g(t), S_k f(t)\rangle_X \, dt = \int_{\mathbb{T}^d} \langle S_k g(t), S_k f(t)\rangle_X \, dt \\ &= \langle S_k g, S_k f\rangle_{L^p(\mathbb{T}^d;X)} = \langle S_k g, f\rangle_{L^p(\mathbb{T}^d;X)}\end{aligned}$$

for all trigonometric polynomial $g \in L^{\overline{p}}(\mathbb{T}^d;X^*)$, $f \in L^p(\mathbb{T}^d;X)$, and, by density and the continuity of S_k, for all g, f in these spaces. Thus $S_k^* = S_k$, and since X^* is UMD whenever X is, and $\overline{p} \in (1,\infty)$ whenever p is, all we have to prove is the random unconditionality of $(S_k)_{k=0}^{\infty}$ on $L^p(\mathbb{T}^d;X)$, for X UMD and $p \in (1,\infty)$.

Recall from the definition that the action of S_n on trigonometric monomials is given by $S_n\left(e^{\mathrm{i}2\pi\kappa\cdot(\cdot)}x\right) := e^{\mathrm{i}2\pi\kappa\cdot(\cdot)}\mathbf{1}_{D_n}(\kappa)x$. Then, by the definition of the inverse kernels $\bar{h}_n$, we have $\bar{h}_n * h_n * f = S_{n-d}f$ for $n > d$, and $S_0 f = \int_{\mathbb{T}^d} f dm = 1 * f = 1 * 1 * f$. Thus

$$\begin{aligned}&\left(\int_\Omega \left|\sum_{k=0}^{n} \varepsilon_k(\omega) S_k f\right|^p_{L^p(\mathbb{T}^d;X)} d\mathbb{P}(\omega)\right)^{\frac{1}{p}} \\ &\quad= \left(\int_\Omega \left|1 * 1 * f + \sum_{k=1}^{n} \varepsilon_k(\omega)\bar{h}_{k+d} * h_{k+d} * f\right|^p_{L^p(\mathbb{T}^d;X)} d\mathbb{P}(\omega)\right)^{\frac{1}{p}} \\ &\quad\le C\,|R|_{\mathcal{B}(L^p(\mathbb{T}^d;X))}\left(\int_\Omega \left|1 * f + \sum_{k=d+1}^{d+n} \varepsilon_k(\omega) h_k * f\right|^p_{L^p(\mathbb{T}^d;X)} d\mathbb{P}(\omega)\right)^{\frac{1}{p}} \\ &\quad\le C\,|R|_{\mathcal{B}(L^p(\mathbb{T}^d;X))} \cdot 3M_p(X)\,|f|_{L^p(\mathbb{T}^d;X)}\,.\end{aligned}$$

The first inequality follows from Corollary 9.9 (recalling that we can add a single bounded operator, in this case $1 * \cdot$, to an R-bounded collection, preserving the R-boundedness with possibly larger R-bound) and the second from Lemma 9.7 (after setting $\epsilon_k := \varepsilon_k(\omega)$ in the conclusion of that lemma, integrating over Ω, and applying the triangle inequality to $\sum_{k=d+1}^{n+d} = \sum_{k=0}^{n+d} - \sum_{k=0}^{d}$, and also to estimate $1 * f$ as a separate quantity). The inequality obtained is recognized as that in the definition of a random unconditional, once we substitute $\sum_{k=0}^{n} S_k f$ in place of f. □

The multiplier theorems then follow with almost no effort at all.

Corollary 9.11 (Littlewood–Paley-type multiplier theorem, Zimmermann 1989). *Let X be a UMD-space. Then $\lambda \in \ell^\infty(\mathbb{Z}^d)$ is in $\mathcal{M}^p(\mathbb{T}^d;X)$, $p \in (1,\infty)$, whenever λ attains a constant value on each of the dyadic blocks D_k, $k \in \mathbb{N}$.*

Proof. The assertion is immediate from the unconditionality of the dyadic decomposition shown in Lemma 9.10 and the abstract multiplier result of Corollary 2.14. □

While the previous result was a direct consequence of the very general abstract theory, the following one, which allows for a wider variety of multipliers, combines this with the estimate for R-bounds of the finitely non-zero multipliers studied above. For $\lambda \in \ell^\infty(\mathbb{Z}^d)$ and $Q \subset \mathbb{Z}^d$, we denote by $\lambda|_Q$ the sequence whose κth term is λ_κ whenever $\kappa \in Q$, and 0 otherwise. With this notion, we have the following result:

Theorem 9.12 (Marcinkiewicz-type multiplier theorem, Zimmermann 1989). *Let X be a UMD-space and $p \in (1,\infty)$. Then there is a constant $C = C(X,p,d)$ such that*

$$|\lambda|_{\mathcal{M}^p(\mathbb{T}^d;X)} \le C \sup_{k\in\mathbb{N}} \operatorname{var} \lambda|_{D_k}$$

for all $\lambda \in \ell^\infty(\mathbb{Z}^d)$ of uniformly bounded variation on the dyadic blocks D_k.

Proof. For any trigonometric polynomial $f \in L^p(\mathbb{T}^d;X)$, we have $T_\lambda f = \sum_{k=0}^\infty T_{\lambda|_{D_k}} S_k f$ (where the sum is actually finite). Thus

$$\begin{aligned}
|T_\lambda f|_{L^p(\mathbb{T}^d;X)} = \left|\sum_{k=0}^\infty T_{\lambda|_{D_k}} S_k f\right|_{L^p(\mathbb{T}^d;X)} &\le C_p \left(\int_\Omega \left|\sum_{k=0}^\infty \varepsilon_k(\omega) T_{\lambda|_{D_k}} S_k f\right|^p_{L^p(\mathbb{T}^d;X)} d\mathbb{P}(\omega)\right)^{\frac{1}{p}} \\
&\le C_p \cdot 8\,|R|_{\mathcal{B}(L^p(\mathbb{T}^d;X))} \sup_{k\in\mathbb{N}} \operatorname{var} \lambda|_{D_k} \left(\int_\Omega \left|\sum_{k=0}^\infty \varepsilon_k(\omega) S_k f\right|^p_{L^p(\mathbb{T}^d;X)} d\mathbb{P}(\omega)\right)^{\frac{1}{p}} \\
&\le C_p \cdot 8\,|R|_{\mathcal{B}(L^p(\mathbb{T}^d;X))} \sup_{k\in\mathbb{N}} \operatorname{var} \lambda|_{D_k} \cdot C_p\,|f|_{L^p(\mathbb{T}^d;X)},
\end{aligned}$$

where the first and last inequalities exploited the unconditionality of $(S_k)_{k=0}^\infty$ via Lemma 3.2, and the second inequality follows from Lemma 9.4. □

9.4 Non-periodic case

We now turn to the study of multiplier transformations acting on functions on $\mathbb{R}^d$. In this section, our goal is to establish analogues on $L^p(\mathbb{R}^D;X)$ of the multiplier theorems of the previous section, and to prove the unconditionality of the Schauder decomposition $(S_n)_{n\in\mathbb{Z}}$. This last result will be used in Chapter 10 in the study of operator-valued Fourier multipliers.

We start by showing that rapidly decreasing functions give rise to multiplier operators.

Theorem 9.13 (Mikhlin-type multiplier theorem, Zimmermann 1989). *Let X be UMD and $p \in (1,\infty)$. Then $\mathcal{S}(\mathbb{R}^d) \subset \mathcal{M}^p(\mathbb{R}^d;X)$, and there is a constant $C = C(X,p,d)$ such that*

$$|\psi|_{\mathcal{M}^p(\mathbb{R}^d;X)} \le C \sup_{|\theta|_\infty \le 1} \sup_{t\in\mathbb{R}^d} |t|^{|\theta|_1} \left|D^\theta \psi(t)\right|$$

Observe that the right-hand side of the asserted inequality is certainly finite for $\psi \in \mathcal{S}(\mathbb{R}^d)$.

Proof. For $\psi \in \mathcal{S}(\mathbb{R}^d)$ and $\phi \in \mathcal{S}(\mathbb{R}^d;X)$, products and Fourier transforms of these functions are also rapidly decreasing, the inversion formula is valid, and Lemma 8.8 can be applied to give

$$\begin{aligned}
\left|\mathcal{F}^{-1}(\psi\widehat{\phi})\right|_{L^p(\mathbb{R}^d;X)} &= \lim_{n\to\infty} \left|T_n^p(\mathcal{F}^{-1}(\psi\widehat{\phi}))\right|_{L^p(\mathbb{T}^d;X)} \\
&= \lim_{n\to\infty} \left|2^{-\frac{dn}{p}} \sum_{\kappa\in\mathbb{Z}^d} \psi(\kappa 2^{-n})\widehat{\phi}(\kappa 2^{-n}) e^{\mathbf{i}2\pi\kappa\cdot(\cdot)}\right|_{L^p(\mathbb{T}^d;X)}.
\end{aligned}$$

The expression inside the norm on the right-hand side is now recognized as the discrete multiplier transform by the sequence $(\psi(\kappa 2^{-n}))_{\kappa\in\mathbb{Z}^d}$ of the function

$$2^{-\frac{dn}{p}} \sum_{\kappa\in\mathbb{Z}^d} \widehat{\phi}(\kappa 2^{-n}) e^{\mathbf{i}2\pi\kappa\cdot(\cdot)} = T_n^p \phi.$$

Thus we find that

$$\left|\mathcal{F}^{-1}(\psi\widehat{\phi})\right|_{L^p(\mathbb{R}^d;X)} \le \left|(\psi(\kappa 2^{-n}))_{\kappa\in\mathbb{Z}^d}\right|_{\mathcal{M}^p(\mathbb{T}^d;X)} \lim_{n\to\infty} |T_n^p\phi|_{L^p(\mathbb{T}^d;X)},$$

and $|T_n^p\phi|_{L^p(\mathbb{T}^d;X)} \to |\phi|_{L^p(\mathbb{R}^d;X)}$ as $n\to\infty$ by Lemma 8.8. It is thus required to find an appropriate bound for the discrete multiplier norm appearing above.

Using Theorem 9.12 and Lemma 9.8, we have

$$\left|(\psi(\kappa 2^{-n}))_{\kappa\in\mathbb{Z}^d}\right|_{\mathcal{M}^p(\mathbb{T}^d;X)} \leq C \sup_{k\in\mathbb{N}} \operatorname{var}(\psi(\kappa 2^{-n}))_{\kappa\in D_k}$$
$$\leq C' \sup_{rd+j\in\mathbb{N}} \max_{|\theta|_\infty\leq 1} \left|D^\theta\psi(\cdot 2^{-n})\right|_{L^\infty(D_{rd+j})} 2^{r|\theta|_1}.$$

Lemma 9.8 actually deals with simple boxes, but it can also be applied to D_k, which is a union of two boxes, after first estimating the variation by the triangle inequality. It is immediate from the definition of the dyadic decomposition (9.5) that each side of the two parts of D_{rd+j} is proportional to 2^r (with the constant of proportionality between 1 and 4, these are taken into account when replacing C by C'); this gives rise to the factor $2^{k\theta}$ in the last expression, since in Lemma 9.8 we have the factor $(\beta-\alpha)^\theta$, with α,β the corners of the box in question.

Using the chain rule, we are lead to estimate the maximum of $\left|D^\theta(t2^{-n})\right| 2^{(r-n)|\theta|_1}$ for $t\in D_{rd+j}$. For such t, we have $|t|\geq|t_j|\geq 2^r$, so that we can estimate

$$\left|D^\theta(t2^{-n})\right| 2^{(r-n)|\theta|_1} \leq \left|D^\theta(t2^{-n})\right| \left|t2^{-n}\right|^{|\theta|_1},$$

but this last quantity, after making the change of variable $t' := t2^{-n}$ and estimating by the supremum over $t'\in\mathbb{R}^d$, is of the form appearing in the assertion. This proves the proposition. □

We need some simple observations concerning the dyadic blocks. The continuous linear one-to-one mapping L_{rd+j}, $r\in\mathbb{Z}$, $j=1,\ldots,d$, defined on $t=(t',t_j,t'')\in\mathbb{R}^{j-1}\times\mathbb{R}\times\mathbb{R}^{d-j}$ by

$$t\to(2^{-(r+1)}t', 2^{-r}t'', 2^{-(r+1)}t_j),$$

takes

$$D_{rd+j} := (-2^{r+1},2^{r+1})^{j-1}\times\left(\cup\pm[2^r,2^{r+1})\right)\times(-2^r,2^r)^{d-j} \quad\text{onto}\quad D_0 := (-1,1)^{d-1}\times\cup\pm[\tfrac{1}{2},1).$$

Furthermore, it takes

$$\bigcup_{k\leq(r+p)d+j} D_k = (-2^{r+1+p},2^{r+1+p})^j\times(-2^{r+p},2^{r+p})^{d-j} \quad\text{onto}\quad \bigcup_{k\leq pd} D_k = (-2^p,2^p)^d. \tag{9.7}$$

With the help of the mappings L_{rd+j}, we can construct the following auxiliary functions:

Lemma 9.14. *There exist $\varphi_n\in\mathcal{D}(\mathbb{R}^d)$, $n\in\mathbb{Z}$, with the following properties:*

1. $\varphi_n|_{D_n}=1$,
2. $\operatorname{supp}\varphi_n\subset\bigcup_{n-2d<k\leq n+d} D_k$, *and*
3. $\sup_{|\theta|_\infty\leq 1}\sup_{t\in\mathbb{R}^d}|t|^{|\theta|_1}\left|D^\theta\varphi_n(t)\right|\leq C$ *for some $C<\infty$, for all $n\in\mathbb{Z}$.*

The φ_n provide smoothed approximations of the multipliers $\mathbf{1}_{D_n}$.

Proof. From (9.7) it is clear that $\cup_{k\leq-2d}D_k \subset\subset D_0\subset\subset \sup_{k\leq d}$. Thus there exists a $\varphi_0\in\mathcal{D}$ such that $\varphi_0=1$ on D_0 and 0 outside $\cup_{-2d<k\leq d}$. We then define $\varphi_n:=\varphi_0\circ L_n$. By the properties of the mappings L_n above, φ_n equals 1 and vanishes in the appropriate sets. To deduce the uniform bound in item 3, observe that $D^\theta\varphi_{rd+j}=D^\theta(\varphi_0\circ L_{rd+j})=2^s 2^{-(r+1)|\theta|_1}D^\theta\varphi_0$, where s is between 0 and d, and thus

$$|t|^{|\theta|_1}\left|D^\theta\varphi_{rd+j}(t)\right|\leq 2^d\left|2^{-(r+1)}t\right|^{|\theta|_1}\left|D^\theta\varphi_0(L_{rd+j}t)\right|\leq 2^d\left|L_{rd+j}t\right|^{|\theta|_1}\left|D^\theta\varphi_0(L_{rd+j}t)\right|.$$

After the change of variable $t':=L_{rd+j}t$, the boundedness of this last quantity for $t'\in\mathbb{R}^d$ is obvious, since $\varphi_0\in\mathcal{D}\subset\mathcal{S}$. □

Lemma 9.15. *$(S_n)_{n\in\mathbb{Z}}$ is an unconditional Schauder decomposition of $L^p(\mathbb{R}^d;X)$, $p\in(1,\infty)$.*

Proof. As in the proof of Lemma 9.10, we verify that both $(S_n)_{n\in\mathbb{Z}}$ and $(S_n^*)_{n\in\mathbb{Z}}$ are random unconditionals, and again, it suffices to do this for one of them, since S_n^* is the operator S_n acting on the space $L^{\overline{p}}(\mathbb{R}^d;X^*)$. The proof exploits the smoothed approximations T_{φ_n} of the S_n provided by Lemma 9.14. Observe in particular that $S_nT_{\varphi_n}=S_n$. Furthermore, since the functions $\phi\in\mathcal{F}^{-1}\mathcal{D}(\mathbb{R}^d;X)$ with $0\notin\operatorname{supp}\mathcal{F}\phi$ are dense in $L^p(\mathbb{R}^d;X)$ (by Lemma A.39), it suffices to study the action of the operators S_n on such ψ. Then all the sums in the following are actually finite:

$$\left(\int_\Omega\left|\sum_{n\in\mathbb{Z}}\varepsilon_n(\omega)S_n\psi\right|^p_{L^p(\mathbb{R}^d;X)}d\mathbb{P}(\omega)\right)^{\frac{1}{p}}=\left(\int_\Omega\left|\sum_{n\in\mathbb{Z}}\varepsilon_n(\omega)S_nT_{\varphi_n}\psi\right|^p_{L^p(\mathbb{R}^d;X)}d\mathbb{P}(\omega)\right)^{\frac{1}{p}}$$
$$\leq C\left(\int_\Omega\left|\sum_{n\in\mathbb{Z}}\varepsilon_n(\omega)T_{\varphi_n}\psi\right|^p_{L^p(\mathbb{R}^d;X)}d\mathbb{P}(\omega)\right)^{\frac{1}{p}}\leq C\sup_{\omega\in\Omega}\left|\sum_{n\in\mathbb{Z}}\varepsilon_n(\omega)\varphi_n\right|_{\mathcal{M}^p(\mathbb{R}^d;X)}|\psi|_{L^p(\mathbb{R}^d;X)}.$$

The first inequality, where $C=16\cdot 2^d\,|R|_{L^p(\mathbb{R}^d;X)}$, followed from the R-boundedness of the multipliers of boxes (Lemma 9.2, recall that D_n consists of two boxes), and the second inequality simply uses the definition of the multiplier norm. To estimate this norm, we invoke Theorem 9.13 and Lemma 9.14 to give

$$\left|\sum_{n\in\mathbb{Z}}\varepsilon_n(\omega)\varphi_n\right|_{\mathcal{M}^p(\mathbb{R}^d;X)}\leq C\sup_{|\theta|_\infty\leq 1}\sup_{t\in\mathbb{R}^d}|t|^{|\theta|_1}\left|D^\theta\sum_{n\in\mathbb{Z}}\varepsilon_n(\omega)\varphi_n(t)\right|$$
$$\leq C\cdot 3d\cdot\sup_{|\theta|_\infty\leq 1}\sup_{t\in\mathbb{R}^d}\sup_{n\in\mathbb{Z}}|t|^{|\theta|_1}\left|D^\theta\varphi_n(t)\right|\leq C',$$

where the first inequality was a direct application of Theorem 9.13, and the second used the fact, following from Lemma 9.14(2), that any point lies in the support of at most $3d$ of the φ_n. Then the third inequality is simply Lemma 9.14(3). This completes the proof. ∎

A Littlewood–Paley-type multiplier result is now as immediate as in the discrete case, Corollary 9.11.

Corollary 9.16. *Let X be a UMD-space. Then $g\in L^\infty(\mathbb{R}^d;X)$ is in $\mathcal{M}^p(\mathbb{R}^d;X)$, $p\in(1,\infty)$, whenever g attains a constant value on each of the dyadic blocks D_n, $n\in\mathbb{Z}$.*

Proof. This follows from Lemma 9.15 and Corollary 2.14. ∎

We could also obtain a non-periodic version of Theorem 9.12 involving similar variation estimates. Instead, we decide to proceed to the operator-valued theorems in the next chapter.

9.5 Notes and comments

This chapter is based on the results of Zimmermann [29], a paper which appeared before the significance of R-boundedness was widely recognized. The UMD theory is older, and the results are stated and proved in this same setting in [29]. In the non-periodic case, we have not treated the full strength of the results in [29] to avoid some technical points.

Instead of R-boundedness, Zimmermann defines another special notion of boundedness, for which one can establish similar properties as the ones for R-bounds used in this chapter. This boundedness involves the norm

$$\|T\|_R:=|R|_{\mathcal{B}(L^p(\mathbb{T}^d;X))}\inf\left\{c>0:T\in c\cdot\overline{\operatorname{conv}}\{m_\phi Rm_\psi\}_{|\phi|_{L^\infty},|\psi|_{L^\infty}\leq 1}\right\},\tag{9.8}$$

where R is the Riesz projection. One should observe the striking similarity between this expression and equation (9.4).

Today, Zimmermann's results can neatly be set in the general framework of the theory of R-boundedness and unconditional Schauder decompositions, as we have done here. (A somewhat similar translation work is also found in Witvliet [28], chapter 3.) The methods of proof in [29] are essentially the same as here, the main difference being the fact that the lemmas are proved in the concrete setting instead of using the abstract versions.

It is an interesting typographical point that the "R" also appears explicitly in Zimmermann's norm (9.8), referring to the operator R. Independently of this, we should note that R-boundedness was originally referred to as the Riesz property, and it was only later that the R was reinterpreted as "randomized" (see Weis [26]).

Chapter 10

Operator-Valued Multipliers

10.1 Introduction

The characterization of operator-valued Fourier multipliers is a place where the notion of R-boundedness shows its power. In Chapter 9, it provided a convenient framework for results originally obtained without this notion, but in the present setting, it becomes necessary. This chapter is a short introduction to the state of art in the field of multipliers.

For a function $G \in L^1_{\mathrm{loc}}(\mathbb{R}^d; \mathcal{B}(X;Y))$, the multiplier operator T_G is defined by $T_G\psi := \mathcal{F}^{-1}(M_G\widehat{\psi})$, where M_G denotes pointwise action of the operator G, i.e., $(M_G f)(t) := G(t)f(t)$. For $\psi \in \mathcal{F}^{-1}\mathcal{D}(\mathbb{R}^d; X)$ (a dense subset of $\mathcal{S}(\mathbb{R}^d; X)$, thus of $L^p(\mathbb{R}^d; X)$ for $p \in [1,\infty)$), we have (by definition) $\widehat{\psi} \in \mathcal{D}(\mathbb{R}^d; X)$, and thus $G\widehat{\psi}$ has compact support, and hence belongs to $L^1(\mathbb{R}^d; Y)$. Thus we can evaluate the transform $\mathcal{F}^*$ in the ordinary L^1 sense; alternatively, we have $L^1(\mathbb{R}^d; Y) \subset \mathcal{S}^*(\mathbb{R}^d; X)$ and we can then take the inverse Fourier transform $\mathcal{F}^{-1}$ in the distribution sense. If the mapping T_G so defined satisfies $|T_G\psi|_{L^q(\mathbb{R}^d;Y)} \leq C\,|\psi|_{L^p(\mathbb{R}^d;X)}$ for some $C < \infty$ independent of $\psi \in \mathcal{D}(\mathbb{R}^d; X)$, then T_G extends to a bounded linear mapping from $L^p(\mathbb{R}^d; X)$ to $L^q(\mathbb{R}^d; Y)$. Whenever this is the case, G is called a Fourier multiplier, and we define

$$|G|_{\mathcal{M}^{p,q}(\mathbb{R}^d;X,Y)} := |T_G|_{\mathcal{B}(L^p(\mathbb{R}^d;X);L^q(\mathbb{R}^d;Y))}\,.$$

We mostly study the case $q = p$, in which case we omit the double superscript, and simply write $\mathcal{M}^p$, as earlier. Also, the main result of the following proposition is the case $q = p$, but it is instructive to observe the byproduct in the case $q > p$, which is obtained at the same strike.

10.2 Recent theorems

We begin with a proposition showing the necessity of R-boundedness.

Proposition 10.1. *Let X, Y be Banach spaces, $p, q \in (1,\infty)$. Let $G \in L^1_{\mathrm{loc}}(\mathbb{R}^d; \mathcal{B}(X;Y))$ be a Fourier multiplier: $G \in \mathcal{M}^{p,q}(\mathbb{R}^d; X, Y)$. Then, if $p = q$, $\{G(t)\}_{t\in\mathfrak{L}G} \subset \mathcal{B}(X;Y)$ is R-bounded with*

$$\mathcal{R}_p(\{G(t)\}_{t\in\mathfrak{L}G}) \leq C\,|G|_{\mathcal{M}^p(\mathbb{R}^d;X,Y)}\,, \tag{10.1}$$

where $\mathfrak{L}G$ denoted the set of Lebesgue points of G and C is a universal constant. In particular, $G \in L^\infty(\mathbb{R}^d; \mathcal{B}(X;Y))$, with the norm bounded by the right-hand side of (10.1).

If $q > p$, then $G = 0$.

Observe that this result does not require the UMD-property of either of the spaces. Henceforth, in the study of L^p multipliers, we can restrict to multiplier functions $G \in L^\infty(\mathbb{R}^d; \mathcal{B}(X;Y))$ without loss of generality.

Proof. The idea is to estimate $G(t)$ at the Lebesgue points by an integral expression, and apply the convergence theorems. We pick a $\psi \in \mathcal{D}(\mathbb{R}^d)$, $\operatorname{ran}\psi \subset [0,1]$, which is symmetric, so that $\widetilde{\psi} = \psi$, and $\int_{\mathbb{R}^d} \psi^2 dm = 1$. Then, from Lemma A.38, applied to ψ^2, it follows that

$$G(t) = \lim_{k\to\infty} \int_{\mathbb{R}^d} k^d \psi^2(k(s-t))G(s)ds, \qquad t \in \mathfrak{L}G,$$

where convergence takes place in the norm of $\mathcal{B}(X;Y)$. It is convenient to write the kernel of the integral operator above in a different form, using $k^d\psi(k\cdot) = \mathcal{F}(\check{\psi}(k^{-1}\cdot))$. Since $\psi \in \mathcal{D} \subset \mathcal{S}$ is symmetric, we further have $\check{\psi} = \mathcal{F}(\widetilde{\psi}) = \widehat{\psi}$, so that $k^d\psi^2(k\cdot) = \mathcal{F}(\widehat{\psi}(k^{-1}\cdot))\psi(k\cdot)$. We use this expression to estimate the R-bound of $\{G(t)\}_{t\in\mathfrak{L}G}$ from the definition:

$$\begin{aligned}\int_\Omega \left| \sum_{j=1}^n \varepsilon_j(\omega) G(t_j)x_j \right|_Y^p d\mathbb{P}(\omega) \\ = \int_\Omega \lim_{k\to\infty} \left| \sum_{j=1}^n \varepsilon_j(\omega) \int_{\mathbb{R}^d} G(s)\mathcal{F}(\widehat{\psi}(k^{-1}\cdot))(s-t_j)x_j\psi(k(s-t_j))ds \right|_Y^p d\mathbb{P}(\omega).\end{aligned}$$

Fatou's lemma can be applied to take out the limit, converting it to $\liminf$ and yielding an upper bound for the quantity above. Inside the integral, observe that $\mathcal{F}(\widehat{\psi}(k^{-1}\cdot))(s-t_j) = \mathcal{F}(e^{\mathbf{i}2\pi t_j\cdot(\cdot)}\widehat{\psi}(k^{-1}\cdot))(s)$, and $G(s)$ applied to this quantity times x_j is the same as

$$\mathcal{F}T_G\left(e^{\mathbf{i}2\pi t_j\cdot(\cdot)}\widehat{\psi}(k^{-1}\cdot)x_j\right)$$

evaluated at s, by the definition of $T_G := \mathcal{F}^{-1}M_G\mathcal{F}$. Furthermore, using $\int \mathcal{F}f \cdot g dm = \int f \cdot \mathcal{F}g dm$ with basic properties of the Fourier transform, we have

$$\begin{aligned}\int_{\mathbb{R}^d} \mathcal{F}T_G\left(e^{\mathbf{i}2\pi t_j\cdot(\cdot)}\widehat{\psi}(k^{-1}\cdot)x_j\right)(s)\psi(k(s-t_j))ds \\ = \int_{\mathbb{R}^d} T_G\left(e^{\mathbf{i}2\pi t_j\cdot(\cdot)}\widehat{\psi}(k^{-1}\cdot)x_j\right)(s) \cdot e^{-\mathbf{i}2\pi t_j\cdot s}k^{-d}\widehat{\psi}(k^{-1}s)ds.\end{aligned}$$

We can then use Hölder's inequality to estimate

$$\begin{aligned}\left| \int_{\mathbb{R}^d} \sum_{j=1}^n \varepsilon_j(\omega)e^{-\mathbf{i}2\pi t_j\cdot s}T_G\left(e^{\mathbf{i}2\pi t_j\cdot(\cdot)}\widehat{\psi}(k^{-1}\cdot)x_j\right)(s)\cdot k^{-d}\widehat{\psi}(k^{-1}s)ds \right|_Y^p \\ \leq \int_{\mathbb{R}^d} \left| \sum_{j=1}^n \varepsilon_j(\omega) m_{e_{-t_j}} T_G m_{e_{t_j}} \left(\widehat{\psi}(k^{-1}\cdot)x_j\right) \right|_Y^p ds \cdot \left(\int_{\mathbb{R}^d} \left| k^{-d}\widehat{\psi}\right|^{\overline{p}} (k^{-1}s)ds \right)^{\frac{p}{\overline{p}}},\end{aligned}$$

where $e_t := e^{\mathbf{i}2\pi t\cdot(\cdot)}$. A change of variable shows that the latter factor equals $k^{-d}\left|\widehat{\psi}\right|^p_{L^{\overline{p}}(\mathbb{R}^d)}$; the $L^{\overline{p}}$ norm here is some finite quantity, since $\widehat{\psi} \in \mathcal{S} \subset L^{\overline{p}}$. For the first term we exploit the assumed boundedness of $T_G : L^q(\mathbb{R}^d;X) \to L^p(\mathbb{R}^d;Y)$ via the resulting R-boundedness of the family of operators $\{m_{e_{-t}}T_G m_{e_t}\}_{t\in\mathbb{R}^d}$ (by Example 4.11 and the product rule of R-bounds; also observe Remark 4.2 on different exponents p, q on the two sides of the R-boundedness inequality). Thus

$$\begin{aligned}\left(\int_\Omega \left| \sum_{j=1}^n \varepsilon_j(\omega) m_{e_{-t_j}} T_G m_{e_{t_j}} \left(\widehat{\psi}(k^{-1}\cdot)x_j\right) \right|^p_{L^p(\mathbb{R}^d;Y)} d\mathbb{P}(\omega) \right)^{\frac{1}{p}} \\ \leq 4\left|T_G\right|_{\mathcal{B}(L^q(\mathbb{R}^d;X);L^p(\mathbb{R}^d;Y))} \left(\int_\Omega \left| \sum_{j=1}^n \varepsilon_j(\omega)\widehat{\psi}(k^{-1}\cdot)x_j \right|^q_{L^q(\mathbb{R}^d;X)} d\mathbb{P}(\omega) \right)^{\frac{1}{q}}.\end{aligned}$$

The integrand in the last expression factors into parts depending only on $\omega \in \Omega$ and on $t \in \mathbb{R}^d$, respectively, so that we have

$$\left(\int_\Omega \left| \sum_{j=1}^n \varepsilon_j(\omega)\widehat{\psi}(k^{-1}\cdot)x_j \right|^q_{L^q(\mathbb{R}^d;X)} d\mathbb{P}(\omega) \right)^{\frac{1}{q}} = \left(\int_\Omega |\varepsilon_j(\omega)x_j|^q_X \, d\mathbb{P}(\omega) \right)^{\frac{1}{q}} \left| \widehat{\psi}(k^{-1}\cdot) \right|_{L^q(\mathbb{R}^d)},$$

and a change of variable shows that $\left|\widehat{\psi}(k^{-1}\cdot)\right|^q_{L^q(\mathbb{R}^d)} = k^d \left|\widehat{\psi}\right|_{L^q(\mathbb{R}^d)}$. Combining all the estimates, we find that

$$\left(\int_\Omega \left| \sum_{j=1}^n \varepsilon_j(\omega)G(t_j)x_j \right|^p_Y d\mathbb{P}(\omega) \right)^{\frac{1}{p}}$$
$$\leq 4 \left|T_G\right|_{\mathcal{B}(L^q;L^p)} \cdot \lim_{k\to\infty} k^{\frac{d}{q}-\frac{d}{p}} \cdot \left|\widehat{\psi}\right|_{L^{\overline{p}}} \left|\widehat{\psi}\right|_{L^q} \left(\int_\Omega |\varepsilon_j(\omega)x_j|^q_X \, d\mathbb{P}(\omega) \right)^{\frac{1}{q}}.$$

If $q = p$, the factor containing k equals unity, whatever k, and the asserted R-boundedness follows. If $q > p$, then $\frac{1}{q} - \frac{1}{p} < 0$, and the limit expression is 0; then the R-bound is zero, and $G(t) = 0$ at every Lebesgue point, so the operator vanishes (a.e.). ☐

The following proposition gives a criterion of convergence for multipliers.

Proposition 10.2. *Let $G_n \in L^\infty(\mathbb{R}^d;\mathcal{B}(X;Y))$, $n \in \mathbb{Z}_+$, be uniformly bounded multipliers, i.e., $|G_n|_{\mathcal{M}^p(\mathbb{R}^d;X,Y)} \leq K$ for all $n \in \mathbb{Z}_+$. Let the Banach space Y be reflexive. If $G_n(t) \to G(t)$ strongly for almost all $t \in \mathbb{R}^d$ as $n \to \infty$, then G is also a multiplier, and*

$$|G|_{\mathcal{M}^p(\mathbb{R}^d;X,Y)} \leq \limsup_{n\to\infty} |G_n|_{\mathcal{M}^p(\mathbb{R}^d;X,Y)}.$$

Since UMD-spaces are reflexive, the conclusion holds in particular for X, Y UMD.

Proof. By (10.1) in Proposition 10.1, the functions G_n are also uniformly bounded in the norm of $L^\infty(\mathbb{R}^d;\mathcal{B}(X;Y))$. For $\phi \in \mathcal{S}(\mathbb{R}^d;X)$ and $\psi \in \mathcal{S}(\mathbb{R}^d)$ we then have, denoting by $\langle\cdot,\cdot\rangle_{\mathcal{S}}$ the pairing of $\mathcal{S}^*(\mathbb{R}^d;Y)$ and $\mathcal{S}(\mathbb{R}^d)$,

$$\left|\langle (T_{G_n} - T_G)\phi, \psi\rangle_{\mathcal{S}}\right| = \left|\left\langle (M_{G_n} - M_G)\widehat{\phi}, \check{\psi}\right\rangle_{\mathcal{S}}\right| = \left|\int_{\mathbb{R}^d} (G_n - G)\widehat{\phi}\check{\psi}dm\right| \leq \int_{\mathbb{R}^d} \left|G_n\widehat{\phi} - G\widehat{\phi}\right|_Y \left|\check{\psi}\right| dm.$$

For fixed ϕ and ψ, the integrand is dominated by $CK\left|\widehat{\phi}\right|_{L^\infty(\mathbb{R}^d;X)} \left|\check{\psi}\right| \in L^1(\mathbb{R}^d)$, since also $\widehat{\phi}$ and $\check{\psi}$ are rapidly decreasing. Furthermore, the integrand tends to zero (a.e.), since $G_n(t)\widehat{\phi}(t) \to G(t)\widehat{\phi}(t)$ for almost every t. The dominated convergence theorem then shows that $\langle T_{G_n}\phi, \psi\rangle_{\mathcal{S}} \to \langle T_G\phi, \psi\rangle_{\mathcal{S}}$ as $n \to \infty$, and this being true for arbitrary $\psi \in \mathcal{S}(\mathbb{R}^d;X)$, we find that $T_{G_n}\phi \to T_G\phi$ as tempered distributions, i.e., in the topology of $\mathcal{S}^*(\mathbb{R}^d;Y)$.

On the other hand, since $\phi \in \mathcal{S}(\mathbb{R}^d;X) \subset L^p(\mathbb{R}^d;X)$ and $|T_{G_n}|_{\mathcal{B}(L^p(\mathbb{R}^d;X);L^p(\mathbb{R}^d;Y))} \leq K$, we find that $\{T_{G_n}\phi\}_{n=1}^\infty$ is a bounded sequence in $L^p(\mathbb{R}^d;Y)$. Since Y is reflexive, $L^p(\mathbb{R}^d;Y) = (L^{\overline{p}}(\mathbb{R}^d;Y^*))^*$, and the Banach–Alaoglu theorem [19] provides us with a subsequence $\{T_{G_{n_k}}\}_{k=1}^\infty$, which converges weakly* to some $g \in L^p(\mathbb{R}^d;Y)$.

Consider then a test function ψy^*, with $\psi \in \mathcal{S}(\mathbb{R}^d)$ and $y^* \in Y^*$. Obviously $\psi y^* \in L^{\overline{p}}(\mathbb{R}^d;Y^*)$. We have, on the one hand,

$$\left\langle T_{G_{n_k}}\phi, \psi y^*\right\rangle_{L^{\overline{p}}(\mathbb{R}^d;Y^*)} \underset{k\to\infty}{\longrightarrow} \langle g, \psi y^*\rangle_{L^{\overline{p}}(\mathbb{R}^d;Y^*)} = \int_{\mathbb{R}^d} \langle g(t), \psi(t)y^*\rangle_{Y^*} \, dt$$
$$= \left\langle \int_{\mathbb{R}^d} g(t)\psi(t)dt, y^* \right\rangle_{Y^*} = \langle \langle g, \psi\rangle_{\mathcal{S}}, y^*\rangle_{Y^*},$$

and on the other

$$\langle T_{G_n}\phi, \psi y^*\rangle_{L^{\overline{p}}(\mathbb{R}^d;Y^*)} = \int_{\mathbb{R}^d} \langle T_{G_n}\phi(t), \psi(t)y^*\rangle_{Y^*}\, dt = \left\langle \int_{\mathbb{R}^d} T_{G_n}\phi(t)\psi(t)dt, y^* \right\rangle_{Y^*}$$
$$= \langle\langle T_{G_n}\phi, \psi\rangle_{\mathcal{S}}, y^*\rangle_{Y^*} \underset{n\to\infty}{\longrightarrow} \langle\langle T_G\phi, \psi\rangle_{\mathcal{S}}, y^*\rangle_{Y^*}.$$

Since the limit is unique, and since these computations are valid for all $y^* \in Y^*$ and all $\psi \in \mathcal{S}(\mathbb{R}^d)$, we must necessarily have $T_G\phi = g$, i.e., $T_{G_{n_k}}\phi \overset{*}{\rightharpoonup} T_G\phi$ in $(L^{\overline{p}}(\mathbb{R}^d;Y^*))^*$ as $n \to \infty$. But then, for an appropriate $f \in L^{\overline{p}}(\mathbb{R}^d;Y^*)$ of unit norm we have

$$|T_G\phi|_{L^p(\mathbb{R}^d;Y)} = \left|\langle T_G\phi, f\rangle_{L^{\overline{p}}(\mathbb{R}^d;Y^*)}\right| = \lim_{k\to\infty}\left|\left\langle T_{G_{n_k}}\phi, f\right\rangle_{L^{\overline{p}}(\mathbb{R}^d;Y^*)}\right|$$
$$\le \limsup_{n\to\infty} |T_{G_n}\phi|_{L^p(\mathbb{R}^d;Y)} \le \limsup_{n\to\infty} |G_n|_{\mathcal{M}^p(\mathbb{R}^d;X,Y)} \cdot |\phi|_{L^p(\mathbb{R}^d;X)}.$$

This shows the claim. ☐

Finally, we give a Mikhlin-type theorem with sufficient conditions for an operator-valued Fourier-multiplier. We here restrict ourselves to the one-dimensional domain $\mathbb{R}$; the reason for this is the fact that the dyadic decomposition $\{D_n\}_{n\in\mathbb{N}}$ introduced in Chapter 9 is only fine enough for a Mikhlin-type theorem in this setting; see Section 10.3.

Theorem 10.3 (Mikhlin-type multiplier theorem, Weis 2000). *Let X and Y be UMD-spaces and $p \in (1,\infty)$. Let $G \in C^1(\mathbb{R}\setminus\{0\};\mathcal{B}(X;Y))$ be an operator-valued function for which*

$$\{G(t)\}_{t\in\mathbb{R}\setminus\{0\}} \qquad \text{and} \qquad \{tG'(t)\}_{t\in\mathbb{R}\setminus\{0\}}$$

are R-bounded. Then G is a Fourier multiplier, and $|G|_{\mathcal{M}^p(\mathbb{R};X,Y)}$ is bounded by a constant $C = C(p,X,Y)$ times the sum of the R-bounds of the two above mentioned sets.

Proof. We may assume that the support of G lies in $[0,\infty)$, since a general operator can be obtained from two such functions with reflection applied to the other. We then consider discretized approximations of G of the following kind:

$$G_k := \sum_{j\in\mathbb{Z}} \left(G(2^j)\mathbf{1}_{[2^j,2^{j+1})} + \sum_{\ell=1}^{2^k} 2^{j-k}G'(2^j + (\ell-1)2^{j-k})\mathbf{1}_{[2^j+\ell 2^{j-k},2^{j+1})} \right).$$

Observe that the jth term of the summation is non-zero only on the interval $[2^j, 2^{j+1})$, so there is no convergence problem in the definition. Furthermore, since G' is continuous on every dyadic interval $[2^j, 2^{j+1})$, it follows that the evaluation in $t \in [2^j, 2^{j+1})$ of summation in ℓ above, easily recognized as a Riemann sum, tends to

$$\int_{2^j}^t G'(s)ds = G(t) - G(2^j),$$

thus $G_k(t) \to G(t)$ for every $t \in \mathbb{R}_+$. By Proposition 10.2, it now suffices to show that the discretized approximations give rise to uniformly bounded multiplier operators.

Denote by S_j the operator whose multiplier is $\mathbf{1}_{[2^j,2^{j+1})}$ and by $S_{j,k,\ell}$ the one corresponding to $\mathbf{1}_{[2^j+\ell 2^{j-k},2^{j+1})}$. Then $\{S_j\}_{j=1}^\infty$ is an unconditional Schauder decomposition of the $L^p(\mathbb{R}_+;X)$, as one easily sees from Lemma 9.15, and

$$\{S_{j,k,\ell}\}_{j\in\mathbb{Z},k\in\mathbb{Z}_+,1\le\ell\le 2^k} \subset \{m_\phi R m_\psi : |\phi|_{L^\infty}, |\psi|_{L^\infty} \le 1\}$$

is R-bounded. (R again denotes the Riesz projection.)

For the operator

$$T_{G_k} = \sum_{j\in\mathbb{Z}} G(2^j)S_j + \sum_{j\in\mathbb{Z}} \left(2^{-k} \sum_{\ell=1}^{2^k} 2^j G'(2^j + (\ell-1)2^{j-k}) S_{j,k,\ell} \right) S_j$$

the boundedness of the first part then follows from Theorem 4.13, since S_j is an unconditional Schauder decomposition of both $L^p(\mathbb{R}_+;X)$ and $L^p(\mathbb{R}_+;Y)$, the set $\{G(2^j)\}_{j\in\mathbb{Z}}$ is certainly R-bounded by the assumption, and since it consists of constant operators, it is also $\{S_j\}_{j\in\mathbb{Z}}$-invariant between the two L^p-spaces. For the second part, we observe that

$$2^{-k} \sum_{\ell=1}^{2^k} 2^j G'(2^j + (\ell-1)2^{j-k}) \in \operatorname{conv}([0,1]\{tG'(t)\}_{t>0}),$$

an R-bounded set by Lemmas 4.10 and 4.12(2). Since $\{S_{j,k,\ell}\}_{j,k,\ell}$ is also R-bounded, we deduce the R-boundedness of the operators multiplying S_j from the product rule, and we can again apply Theorem 4.13. The $\{S_j\}_{j\in\mathbb{Z}}$-invariance is again clear, since the multipliers $S_{j,k,\ell}$ and S_j obviously commute.

By the boundedness of the approximations and their convergence to the original operator, the assertion is established. □

10.3 Notes and comments

Propositions 10.1 and 10.2 are taken from Hieber and Prüss [8]. A multiplier convergence result similar to Proposition 10.2 is stated for scalar-valued multipliers (without proof) in Zimmermann [29].

Theorem 10.3 was first proved by Weis [26]. Another proof is given in Hieber and Prüss [8]. The proof given above has some elements from both of the two.

As mentioned before the statement of Theorem 10.3, the dyadic decompositions introduced in Chapter 9 are not fine enough for a Mikhlin-type theorem with domain $\mathbb{R}^d$, $d>1$. The reason is roughly the following: The natural generalization of the Mikhlin-type condition on the (R-)boundedness of $tG'(t)$ would in higher dimension involve expressions of the form $t^\theta D^\theta G(t)$, $\theta \in \{0,1\}^d$. In the one dimensional case, the derivative G' is allowed to blow up at one point, the origin, but on $\mathbb{R}^d$, $d>1$, the Mikhlin-type condition allows such behaviour on all the coordinate axes. The dyadic decomposition $\{D_n^{(1)}\}_{n\in\mathbb{Z}}$ of $\mathbb{R}$ leaves out the origin in a convenient manner, but the blocks $D_n^{(d)}$ of the d-dimensional decompositions do intersect with coordinate axes. For a d-dimensional Mikhlin-type theorem, a refined decomposition would be needed, say, the product decomposition

$$\{\Delta_\nu\}_{\nu\in\mathbb{Z}^d} \qquad \Delta_\nu := D_{\nu_1}^{(1)} \times \cdots \times D_{\nu_d}^{(1)}.$$

Unfortunately, the UMD-condition is insufficient to deduce the unconditionality of this refined decomposition. To obtain improved theorems related to the decomposition $\{\Delta_\nu\}_{\nu\in\mathbb{Z}^d}$, additional assumptions on the underlying Banach space are required; Witvliet [28] studies UMD-spaces with *property* (α) and Zimmermann [29] those with *local unconditional structure.*

We should emphasize that this chapter is only a very brief introduction to the modern theory of multipliers. Out of the many important results that have been obtained recently, we have only presented a very minimal collection to give an idea of the techniques involved.

Both [8] and [26] contain other multiplier results related to Theorem 10.3, and also applications to maximal L^p-regularity. A discrete analogue of Theorem 10.3 is contained in Arendt and Bu [1]:

Theorem 10.4 (Marcinkiewicz-type multiplier theorem, Arendt–Bu 2000). *Let X and Y be UMD-spaces, $p\in(1,\infty)$, and $(G_k)_{k\in\mathbb{Z}} \in \mathcal{B}(X;Y)^{\mathbb{Z}}$. If*

$$\{G_k\}_{k\in\mathbb{Z}} \qquad \text{and} \qquad \{k\,(G_{k+1}-G_k)\}_{k\in\mathbb{Z}}$$

are R-bounded, then $(G_k)_{k\in\mathbb{Z}}$ is an L^p Fourier multiplier.

Witvliet's thesis [28] contains an extensive treatment of modern multiplier theorems and applications. Some of the results have also appeared in Clément et al. [3].

A versatile overview of the modern theory of R-boundedness and Fourier-multipliers, from the point of view of applications to maximal L^p-regularity, is given in Weis [25].

Chapter 11

Summary

In this work, we have studied the modern extensions of the classical theory of Fourier multiplier transformations, setting emphasis on the techniques and ideas behind the recent developments. We have explained the emergence of the notion of multipliers in the classical context, its vector-valued generalizations, and the use of R-boundedness to characterize operator-valued multiplier transformations. Furthermore, we have introduced the UMD-spaces and proved the main theorems related to them. Finally, we have given examples of the modern multiplier results, which combine the UMD-theory and the notion of R-boundedness.

The notion of multipliers naturally emerged from the structure of bounded linear translation invariant operators between L^p-spaces when viewed in the Fourier domain. A complete characterization was available for the multipliers on L^1, L^2 and L^∞.

In order to view the problem from an abstract point of view, the L^p-spaces were replaced by a general Banach space X, and the decomposition of a function $f \in L^p$ into its harmonic components by a general Schauder decomposition $x = \sum_{k=1}^{\infty} x_k$, $x_k \in X_k$. A particularly simple characterization of multiplier operators was obtained when the underlying decomposition was unconditional.

The multiplier problem in any concrete setting is then related to finding appropriate unconditional decompositions. On the usual L^p-spaces, the division of a function into its harmonic components is a Schauder decomposition, but unconditional only on L^2. However, grouping the Fourier coefficients into appropriate blocks leads to the dyadic decomposition, which is unconditional.

In characterizing decompositions, the behaviour under randomization was found to be of interest, and in this connection, the Khintchine–Kahane inequality provided a useful tool. We also saw that the notion of R-boundedness of operator families is very naturally related to the boundedness of generalized multiplier transformations, which are obtained from sequences of operators acting componentwise on a Schauder decomposition. These considerations lead to interesting abstract multiplier theorems with a relatively small effort.

In order to efficiently apply the abstract machinery to obtain strong Fourier multiplier theorems for vector-valued functions, some conditions on the geometry of the Banach space are needed. It is natural to impose a condition related to the boundedness of a representative "test multiplier", and for this purpose, the Hilbert and Riesz transforms were found to be appropriate. However, it was more convenient to take as a starting point the requirement of the unconditionality of martingale differences (UMD), which implies the other two (and is, in fact, equivalent to them). Working with the UMD-condition, we showed the independence of the condition on the exponent p; furthermore, we found that UMD-spaces are reflexive, and X^* and $L^p(\Gamma; X)$ are UMD whenever X is.

The boundedness of the Hilbert transform was applied to deduce the boundedness of a large class of multiplier operators acting on functions taking values in a UMD-space, and we found that the dyadic decomposition is unconditional on $L^p(\mathbb{R}; X)$ for X UMD. Finally, the interplay of R-boundedness with the unconditionality of the dyadic decomposition lead to operator-valued Fourier multiplier theorems on UMD-spaces, the most recent results in the field.

The results in this work have appeared elsewhere, but the collective presentation is new. In preparing this work, we have in particular exploited the lecture notes of de Pagter [5], the monograph of Hieber and Prüss [8], the thesis of Witvliet [28] (and the related article by Clément et al. [3]), and the paper of Zimmermann [29].

In addition to the arrangement, the following attributes are, up to our knowledge, new in this work: We have introduced here the notion of relative R-boundedness (Definition 4.9), which does appear implicitly in [3] and in [28], though. As Theorem 4.14 shows, this notion is appropriate in characterizing certain abstract multipliers. Also, the proof of the vector-valued Jensen's inequality (Lemma 5.7) is possibly new, and it differs from the usual argument in the scalar-valued situation. Chapter 9 is largely translated to the language of R-boundedness and Schauder decompositions from the somewhat more cumbersome original notion.

Appendix A

Vector-Valued Analysis

A.1 Introduction

The appendix provides an account of the principles of mathematical analysis of vector-valued functions, which are used throughout this work. In the present context, "vector" always refers to a point of a Banach space. The topics to be touched are vector-valued integration and the theory of the Lebesgue–Bôchner spaces $L^p(\Omega; X)$ and their duals, vector-valued extensions of linear operators, differentiability properties of vector-valued functions, Fourier analysis, and the theory distributions.

A.2 Abstract integration

We will here develop an appropriate theory for integrating functions defined on a measure space Ω (with a σ-algebra $\mathcal{F}$ and a positive measure μ), and taking values in a Banach space X. For the moment, we will restrict the considerations to separable spaces; this ensures that we can use simple functions as an auxiliary device in much the same way as in real analysis. At the end of the section we discuss how to remove this assumption.

There are various ways of defining measurability and integration in the vector-valued setting, but we only present one; alternative procedures are discussed at some length in Section A.9.

A **Borel measurable** function is defined, as in the real case, as $f : \Omega \to X$, for which the preimage of every Borel set of X is a measurable set, i.e., $f^{-1} : \mathfrak{B}(X) \to \mathcal{F}$. Recall (or see [20]) that this is equivalent to merely requiring that the preimage of every open set be measurable; the vector-valued setting does not affect this topological fact. As usual, the class $\mathfrak{B}(X)$ of Borel sets of X means the σ-algebra generated by the topology $\mathcal{T}$ induced by the norm of X.

We now state the first lemma, which gives us the possibility to approximate arbitrary measurable functions by simple functions in a convenient way. The separability of X will play the key role in the proof.

Lemma A.1. *Let $\phi : X \to \mathbb{R}$ be continuous and convex and attain a minimum at $x_0 \in X$. Then each measurable function $f : \Omega \to X$ is a pointwise limit of simple measurable functions f_n satisfying $\phi(f_n(\omega)) \le \phi(f(\omega)) + \frac{1}{n}$ and $|f_n(\omega) - f(\omega)|_X \le |f(\omega)|_X + |x_0|_X$ for all $n \in \mathbb{Z}_+$, $\omega \in \Omega$. If ϕ attains the minimum at a unique point, then the $\frac{1}{n}$ can be omitted.*

For the construction of the integral, we merely need the special case of the lemma with ϕ given by $\phi(x) = |x|_X$. This clearly satisfies the hypotheses, and the unique point of minimum is $x_0 = 0$; thus the estimate for the difference of f and the approximating functions in the assertion reduces to $|f_n(\omega) - f(\omega)|_X \le |f(\omega)|_X$.

The full strength of the lemma is exploited in proving a vector-valued version of Jensen's inequality.

Proof. Let $\{x_k\}_{k=1}^{\infty}$ be a dense sequence in X, and consider the sequence $\{x_k\}_{k=0}^{\infty}$, where we have added the point x_0 where ϕ has the minimum. (If the point of minimum is not unique, we just pick one.) We define the auxiliary functions $s_n : X \to \{x_k\}_{k=0}^{n} \subset X$, $n \in \mathbb{Z}_+$, as follows: $s_n(x) := x_k$, where $|x - x_k|_X$ is the minimum, subject to the constraints $k \in \{0, \dots, n\}$ and $\phi(x_k) \leq \phi(x) + \frac{1}{n}$. For definiteness, let k be the smallest value with this property. Note that at least x_0 satisfies the constraint conditions, so that the mappings s_n are well-defined for each n. Writing the definition of s_n the other way round, we find that

$$s_n^{-1}(x_k) = \{x \in X : \phi(x) \geq \phi(x_k) - \frac{1}{n}, |x - x_k|_X < |x - x_\ell|_X \text{ for } 0 \leq \ell < k, \\ |x - x_k|_X \leq |x - x_\ell|_X \text{ for } k < \ell \leq 2n\}.$$

This is obviously a Borel set: it is an intersection of a finite number of open and closed sets, which appear almost explicitly in the expression. Note in particular that $\{x : \phi(x) \geq \phi(x_k)\} = \phi^{-1}[\phi(x_k) - \frac{1}{n}, \infty)$ is a closed set, since ϕ is continuous. (The same conclusion is clearly true with the $\frac{1}{n}$ omitted, which is the case if the point of minimum of ϕ is unique.) Furthermore, the image under s_n^{-1} of any Borel set of X (in fact, of any subset of X) is the union of the images of the finite number of x_k it contains, which is again a Borel set. Thus s_n is a Borel function, for each n.

Now $\{s_n\}_{n=1}^{\infty}$ has the following properties, for each $x \in X$:

1. $\phi(s_n(x)) \leq \phi(x) + \frac{1}{n}$,
2. $|s_n(x) - x|_X \leq |x_0 - x|_X \leq |x_0|_X + |x|_X$,
3. $|s_n(x) - x|_X = \min\{|x - x_k|_X : k \leq n, \phi(x_k) \leq \phi(x) + \frac{1}{n}\} \downarrow 0$ as $n \uparrow \infty$.

(Again, we omit the $\frac{1}{n}$ if x_0 is the unique point of minimum of ϕ.) Items 1 and 2 are immediate from the definition. If we did not have the condition $\phi(x_k) \leq \phi(x) + \frac{1}{n}$, item 3 would follow directly from the density of $\{x_k\}_{k=0}^{\infty}$. Even with this condition, the conclusion is quite readily established:

If x_0 is the unique point where ϕ attains its minimum, then $x = x_0$ is among each set $\{x_k\}_{k=0}^{n}$, so there is nothing to prove. If x_0 is the unique minimum but $x \neq x_0$, then we clearly find a point $y \in X$ such that $\phi(y) < \phi(x)$ (say, $y = x_0$), and even if ϕ has several points of minimum, we nevertheless find, for any $x \in X$, a $y \in X$ such that $\phi(y) < \phi(x) + \frac{1}{n}$. We proceed with this latter form; it is understood that the $\frac{1}{n}$ is omitted in case ϕ has a unique minimum point.

Now for $\lambda \in (0, 1)$, $\phi(\lambda y + (1 - \lambda)x) \leq \lambda\phi(y) + (1 - \lambda)\phi(x) < \phi(x) + \frac{1}{n}$. As $\lambda \downarrow 0$, $y_\lambda := \lambda y + (1 - \lambda)x \to x$. Thus, arbitrarily close to x, we can find a y_λ such that $\phi(y_\lambda) < \phi(x) + \frac{1}{n}$. Then due to the density of $\{x_k\}_{k=0}^{\infty}$, we can find x_k in the vicinity of any such y_λ, and by the continuity of ϕ, they also satisfy $\phi(x_k) < \phi(x) + \frac{1}{n}$, when chosen close enough to y_λ. Thus clearly $|x - x_k|_X$ can be made as small as desired, retaining the condition $\phi(x_k) \leq \phi(x) + \frac{1}{n}$. Now all the items listed above are verified.

The auxiliary functions s_n with the above mentioned properties at hand, the lemma is almost proved: If $f : \Omega \to X$ is measurable, let $f_n := s_n \circ f$, whence the f_n are obviously simple; also, they satisfy the norm bounds of the assertion by 1 and 2 above, and f is the pointwise limit of f_n by 3. Finally, as a composition of a measurable and Borel functions, each f_n is measurable. ☐

In addition to Lemma A.1, it is useful to know that any pointwise limit (when it exists) of a sequence of measurable functions, whether simple or not, is again measurable. Unfortunately, the familiar real analysis results of similar kind concerning suprema, infima and lower and upper limits do not have meaning in the vector-valued setting.

Lemma A.2. *The pointwise limit f of measurable functions f_n, $n \in \mathbb{Z}_+$, is measurable.*

Once this is proved, it follows immediately that also the pointwise limit a.e. of measurable functions is measurable, possibly after being redefined on a set of measure zero (indeed, on the set where convergence does not take place, for instance).

Proof. We have to show that the preimage of every open set $G \subset X$ is measurable. Let $G_k := \{x \in G : d(x, G^c) > \frac{1}{k}\}$, where $d(\cdot,\cdot)$ denotes the distance of an element from a (closed) set; G^c is certainly closed as the complement of the open G. Then the G_k are open, and $G_k \uparrow G$. Furthermore, for $f(\omega) \in G$ we have $f(\omega) \in G_k$ for some k and then $f_\ell(\omega) \in G_k$ for all $\ell > n$ (say). Conversely, is $f_\ell(\omega) \in G_k$ for all $\ell > n$, for some k, then $f_\ell(\omega)$ must converge to a point $f(\omega)$ in $\overline{G_k} \subset G_{k+1} \subset G$ (the convergence to *some* point $f(\omega)$ was in the hypothesis of the lemma). It follows that

$$\{f \in G\} = \bigcup_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{\ell=n}^{\infty} \{f_\ell \in G_k\},$$

and each of the $\{f_\ell \in G_k\}$ is a preimage of an open set G_k, thus measurable, and the measurability of $\{f \in G\}$ follows, since the class $\mathcal{F}$ of measurable sets is a σ-algebra. $\square$

It is now time to define what we mean by integration of measurable vector-valued functions. We call a measurable f **integrable** whenever the real-valued function $|f(\cdot)|_X$ is integrable. Observe that $|f(\cdot)|_X : \Omega \to [0,\infty)$ is certainly measurable as a composition of the measurable $f : \Omega \to X$ and the continuous $|\cdot|_X : X \to [0,\infty)$. We use the customary notation $L^1(\Omega; X)$ for the (Lebesgue–Bôchner) space of integrable functions, unless different σ-algebras are involved, in which case we switch to the more accurate notation $L^1(\mathcal{F}; X)$. Only one measure (possibly restricted to different σ-algebras) will be used on the space Ω, so there is no need to indicate the measure in the notation of integrability. The norm is defined in analogy with the scalar case by $|f|_{L^1(\Omega;X)} := \int_\Omega |f(\omega)|_X \, d\mu(\omega)$. The spaces $L^p(\Omega; X)$ for $p > 1$, including $p = \infty$, and the corresponding norms are defined similarly in the obvious way.

With the definitions as above, many of the properties of vector-valued integrals (to be defined pretty soon) boil down to the corresponding properties of the real-valued Lebesgue integral. The facts that $L^p(\Omega; X)$ is a Banach space for each $p \in [1,\infty]$ can be shown essentially as in the real case (see e.g. [20]; a proof in a Banach space is given in [9]), the crucial point being the fact that an absolutely convergent series is convergent in a Banach space (see Section 2.2; the completeness is needed here). Also, if $f_n \to f$ in $L^p(\Omega; X)$, then a subsequence $f_{n_j} \to f$ a.e. Note that these properties entirely relied on the ordinary Lebesgue integral, and did not even require any definition of vector-valued integration! Familiar properties of probability spaces Ω with a probability measure $\mathbb{P}$, such as the monotonicity of $|f|_{L^p(\Omega;X)}$ in p for fixed f, or more generally, Jensen's inequality $\phi(|f|_{L^1(\Omega;X)}) \leq \int_\Omega \phi(|f(\omega)|_X) d\mu(\omega)$, for $\phi : \mathbb{R} \to \mathbb{R}$ convex, also follow in the same way.

Now finally, the **integral** of a simple integrable function is defined in the obvious way:

$$\int_\Omega \left(\sum_{k=1}^{n} x_k \mathbf{1}_{E_k} \right) d\mu := \sum_{k=1}^{n} x_k \mu(E_k).$$

For E_k disjoint (and clearly they can always be chosen that way), we have

$$\left| \sum_{k=1}^{n} x_k \mathbf{1}_{E_k} \right|_X = \sum_{k=1}^{n} |x_k|_X \, \mathbf{1}_{E_k},$$

and thus

$$\left| \int_\Omega \sum_{k=1}^{n} x_k \mathbf{1}_{E_k} d\mu \right|_X = \left| \sum_{k=1}^{n} x_k \mu(E_k) \right|_X \leq \sum_{k=1}^{n} |x_k|_X \, \mu(E_k) = \int_\Omega \left| \sum_{k=1}^{n} x_k \mathbf{1}_{E_k} \right|_X d\mu,$$

i.e., $\left| \int_\Omega f d\mu \right|_X \leq \int_\Omega |f|_X \, d\mu = |f|_{L^1(\Omega;X)}$ for simple integrable f.

Clearly a simple measurable $f = \sum_{k=1}^{n} x_k \mathbf{1}_{E_k}$ (with E_k disjoint and $x_k \neq 0$) is integrable if and only if $\mu(E_k) < \infty$ for $k = 1, \ldots, n$, or equivalently, $\mu(f \neq 0) < \infty$. We will denote the space of simple $L^1(\Omega; X)$ functions by $S(\Omega; X)$. (This is obviously a vector space.) The basic norm inequality above soon allows us to extend the definition of the bounded linear operator

$\int_\Omega (\cdot) d\mu : S(\Omega; X) \to X$ to all of $L^1(\Omega; X)$ by continuity. Indeed, all we need to show is that S is dense in $L^1(\Omega; X)$. This is the next task.

Lemma A.3. *1. The space $S(\Omega; X)$ is dense in $L^p(\Omega; X)$, $1 \le p < \infty$.*

2. If $\mu(\Omega) < \infty$, and $\phi : X \to \mathbb{R}$ is convex, continuous and attains a minimum at $x_0 \in X$, and $f \in L^p(\Omega; X)$, then $\{f_n\}_{n=1}^\infty \subset S(\Omega; X)$ converging to f in L^p can be chosen so that $\phi(f_n(\omega)) \le \phi(f(\omega)) + \frac{1}{n}$, $|f_n(\omega) - f(\omega)|_X \le |f(\omega)|_X + |x_0|_X$ for all $n \in \mathbb{Z}_+$, $\omega \in \Omega$.

If $x_0 = 0$, then the restriction $\mu(\Omega) < \infty$ can be dropped.

It is obvious that $S(\Omega; X) \subset L^p(\Omega; X)$, also for $p = \infty$, since a simple function attains only finitely many values (thus certainly of uniformly bounded norm) and since the set on which a simple integrable function differs from zero is of finite measure.

Proof. By Lemma A.1, $f \in L^p(\Omega; X)$ is the pointwise limit of a sequence $\{f_k\}_{k=1}^\infty$ of simple functions, which satisfy the inequalities (recall that $\phi = |\cdot|_X$, $x_0 = 0$ satisfy the conditions of that lemma) $|f_k(\omega)|_X \le |f(\omega)|_X$ and $|f_k(\omega) - f(\omega)|_X \le |f(\omega)|_X$ for every $\omega \in \Omega$. It follows from the first of these inequalities that $f_k \in L^p(\Omega; X)$, thus $f_k \in S(\Omega; X)$. (A priori, we only knew that the f_k are simple and measurable, not necessarily integrable.) Using the second inequality, Lebesgue theorem of dominated convergence (for real integrals) shows that $\int_\Omega |f_k - f|_X^p \, d\mu \to 0$ as $k \to \infty$, but this is what we wanted to prove.

For the second part of the lemma, we also take the sequence $\{f_n\}_{n=1}^\infty$ given by Lemma A.1, but now related to the given function ϕ instead of $|\cdot|_X$, and we must show the L^p convergence. But this follows again by dominated convergence, since the assumption $\mu(\Omega) < \infty$ guarantees that $|f(\cdot)|_X + |x_0|_X \in L^p(\Omega) + L^\infty(\Omega)$ is integrable. If $x_0 = 0$, the integrability holds regardless of whether $\mu(\Omega) < \infty$ or not. □

This result at hand, the operator $\int_\Omega \cdot d\mu$ extends to all of $L^1(\Omega; X)$, as indicated above, and thus the celebrated vector-valued integration over Ω is now defined. Integrals over measurable subsets $F \in \mathcal{F}$ are defined in the usual way by $\int_F f d\mu := \int_\Omega f \mathbf{1}_F d\mu$. This integral has many of the properties we would like and expect; for instance, if Y is another separable Banach space, $A \in \mathcal{B}(X; Y)$ and $f \in L^p(\Omega; X)$, we see that $Af : \Omega \to Y$ is measurable (as a composition of continuous and measurable functions) and $|Af|_{L^p(\Omega;Y)} \le |A|_{\mathcal{B}(X;Y)} |f|_{L^p(\Omega;X)}$. Furthermore, we immediately deduce the equality

$$A \int_\Omega f(\omega) d\mu(\omega) = \int_\Omega Af(\omega) d\mu(\omega), \tag{A.1}$$

which is readily verified for simple f and extends to the general case by the density of the simple functions and the continuity of $\int_\Omega \cdot d\mu$. As an important special case, take $A = x^* \in X^* := \mathcal{B}(X; \mathbb{C})$ to deduce $\langle x^*, \int_\Omega f d\mu \rangle = \int_\Omega \langle x^*, f \rangle \, d\mu$. Observe, in addition to this nice result, that the deduction via the simple functions and approximation was essentially the same as the standard procedure in the scalar case. (The equality (A.1) holds more generally for closed operators A, see [6], Theorem II.2.6, or [9], Theorem 3.7.12, but we will not need this here.)

In real analysis, it is an almost trivial fact that the constant zero function is essentially the only one whose integral vanishes over every measurable set. This result extends to the vector-valued case, with the help of the following lemma.

Lemma A.4. *If X is separable, then there is a sequence $\{\xi_k^*\}_{k=1}^\infty \subset S(0; 1) \subset X^*$ such that $|x|_X = \sup_{k \in \mathbb{Z}_+} |\langle \xi_k^*, x \rangle|$.*

Proof. Let $\{\xi_k\}_{k=1}^\infty$ be a dense sequence in X. By the Hahn–Banach theorem we can find $\{\xi_k^*\}_{k=1}^\infty \subset X^*$ such that $\langle \xi_k^*, \xi_k \rangle = |\xi_k|_X$ and $|\xi_k^*|_{X^*} = 1$ for each k. Then, given $x \in X$ and $\epsilon > 0$, we can find a ξ_k such that $|x - \xi_k|_X < \epsilon$. Thus $|\langle \xi_k^*, x \rangle| \ge |\langle \xi_k^*, \xi_k \rangle| - |\langle \xi_k^*, \xi_k - x \rangle| \ge |\xi_k^*|_- \epsilon$. □

Corollary A.5. *If $f \in L^1(\mathcal{F}; X)$ satisfies $\int_F f d\mu = 0$ for every $F \in \mathcal{F}$, then $f = 0$ (a.e.).*

Proof. If f satisfies the assumption, then $\int_F \langle x^*, f\rangle \, d\mu = \left\langle x^*, \int_F f d\mu \right\rangle = 0$ for all $F \in \mathcal{F}$ and all $x^* \in X^*$. We hence know from real analysis that $\langle x^*, f(\omega)\rangle = 0$ except possibly for $\omega \in Z_{x^*}$ (say), where $\mu(Z_{x^*}) = 0$.

Now the set $Z := \cup_{k=1}^{\infty} Z_{\xi_k^*}$, where $\{\xi_k^*\}_{k=1}^{\infty}$ is a sequence as in Lemma A.4, has probability zero, and for $\omega \in Z^c$, i.e., almost surely, $\langle \xi_k^*, f(\omega)\rangle = 0$ for all ξ_k^*, $k \in \mathbb{Z}_+$. But this implies $f(\omega) = 0$. Thus (a.e.) $f = 0$. □

We will also exploit a vector version of Fubini's theorem.

Lemma A.6 (Fubini's theorem). *Let $(\Omega_i, \mathcal{F}_i, \mu_i)$, $i = 1, 2$, be two σ-finite probability spaces and $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, \mu_1 \times \mu_2)$ the product space defined in the usual way. If $f \in L^1(\Omega_1 \times \Omega_2; X)$, then*

$$\int_{\Omega_1 \times \Omega_2} f(\omega_1, \omega_2) d(\mu_1 \times \mu_2)(\omega_1, \omega_2) = \int_{\Omega_1} d\mu_1(\omega_1) \left(\int_{\Omega_2} f(\omega_1, \omega_2) d\mu_2(\omega_2) \right).$$

Proof. If $f = \sum_{k=1}^{n} x_k \mathbf{1}_{E_k}$ is simple and integrable, then the claim reduces to the equations $\int_{\Omega_1 \times \Omega_2} \mathbf{1}_{E_k} d(\mu_1 \times \mu_2) = \int_{\Omega_1} d\mu_1 \left(\int_{\Omega_2} \mathbf{1}_{E_k} d\mu_2 \right)$, which is just real-valued Fubini's theorem [20] for the indicator functions. In the general case, use Lemma A.1 to get simple f_k such that $f_k \to f$ pointwise as $k \to \infty$, $|f_k - f|_X \le |f|_X$ (pointwise) for all k and all points of the product space $\Omega_1 \times \Omega_2$. Then, by the dominated convergence theorem, it follows that $\int_{\Omega_1 \times \Omega_2} |f - f_k|_X \, d(\mu_1 \times \mu_2) \to 0$. By the real-valued Fubini's theorem, $|f(\cdot,\cdot)|_X \in L^1(\Omega_1 \times \Omega_2)$ implies $|f(\omega_1, \cdot)|_X \in L^1(\Omega_2)$ for a.e. $\omega_1 \in \Omega_1$; thus we can also use dominated convergence to deduce $\int_{\Omega_2} |f(\omega_1, \cdot) - f_k(\omega_1, \cdot)|_X \, d\mu_2 \to 0$ for a.e. $\omega_1 \in \Omega_1$. Then we have one more dominated convergence

$$\int_{\Omega_1} d\mu_1(\omega_1) \left(\int_{\Omega_2} |f(\omega_1, \omega_2) - f_k(\omega_1, \omega_2)|_X \, d\mu_2(\omega_2) \right) \to 0,$$

and the result follows for general $f \in L^1(\Omega; X)$. □

We conclude this section by having a look beyond separable spaces.

Remark A.7. *The integration theory can be extended to non-separable Banach spaces X by restricting the treatment to measurable functions f with* ***essentially separable range****, i.e., we require that there is an $A \subset \Omega$ such that $\mu(A^c) = 0$ and $f(A) \subset X$ is separable.*

Indeed, we can redefine f on A^c to take values in $f(A)$ without doing any harm in view of the usual equivalence class philosophy. If $\{\xi_k\}_{k=1}^{\infty}$ is dense in $f(A)$ (which is $\operatorname{ran} f$ after the redefinition), then the set of finite linear combinations of the ξ_k, $k \in \mathbb{Z}_+$, with rational coefficients is dense in $\operatorname{span} \operatorname{ran} f$, and it is a countable set; thus $\operatorname{span} \operatorname{ran} f$ is separable. Furthermore, if a set is dense in $\operatorname{span} \operatorname{ran} f$, it is also dense in $\overline{\operatorname{span}} \operatorname{ran} f$, and we conclude that $\overline{\operatorname{span}} \operatorname{ran} f$ is a separable Banach space in which f can be integrated. Furthermore, if we have countably many measurable, essentially separably-valued functions f_k, a similar argument as above, involving linear combinations with rational coefficients, shows that $\overline{\operatorname{span}}(\cup_{k=1}^{\infty} \operatorname{ran} f_k)$ is also separable, whence integrals of the f_k, $k \in \mathbb{Z}_+$, can be defined in a common Banach space. Thus all of the integration theory works, if we simply add the requirement of essentially separable range to the definition of measurability.

It is obvious by now that all this hardly extends anything from the separable setting. The whole procedure may almost appear like cheating, but we nevertheless have integration defined in a general setting. It is possible to give other, perhaps more "natural" definitions and prove the essential separability of the range of a measurable function as a theorem (see Section A.9); however, this does not give us greater generality, and we can equally well take the essential separability as a definition.

A.3 Dual of $L^p(\Omega; X)$

Since duality arguments provide powerful tools for treating various problems of analysis, it is useful to have some knowledge of the duals of the Lebesgue–Bôchner spaces of integrable vector-valued functions. We begin with an embedding result. In order not to confuse different duality pairings, we use the notation $\langle x^*, x\rangle_X$ for the pairing between $x \in X$ and $x^* \in X^*$.

Lemma A.8. *If X is a Banach space, $p \in (1,\infty)$ and $g \in L^{\overline{p}}(\Omega; X^*)$, then*

$$L^p(\Omega; X) \ni f \mapsto \int_\Omega \langle g(\omega), f(\omega)\rangle_X \, d\mu(\omega)$$

defines a continuous linear functional g^* on $L^p(\Omega; X)$, of norm $|g|_{L^{\overline{p}}(\Omega;X^*)}$.

Thus $L^{\overline{p}}(\Omega; X^*)$ can be identified with a (closed) linear subspace of $L^p(\Omega; X)^*$, the mapping $L^{\overline{p}}(\Omega; X^*) \ni g \mapsto g^* \in L^p(\Omega; X)^*$ being a linear isometry.

Proof. Since f and g are limits (a.e.) of simple functions, say f_k and g_k, respectively, it is clear that $\omega \mapsto \langle g(\omega), f(\omega)\rangle_X$ is the limit (a.e.) of the simple functions $\omega \mapsto \langle g_k(\omega), f_k(\omega)\rangle_X$, and thus measurable. Furthermore

$$\left|\int_\Omega \langle g(\omega), f(\omega)\rangle_X \, d\mu(\omega)\right| \le \int_\Omega |g(\omega)|_{X^*} \, |f(\omega)|_X \, d\mu(\omega) \le |g|_{L^{\overline{p}}(\Omega;X^*)} \, |f|_{L^p(\Omega;X)} \, ,$$

whence $|g^*|_{L^p(\Omega;X)^*} \le |g|_{L^{\overline{p}}(\Omega;X^*)}$. We must still show that actually the equality holds.

Take first $g \in S(\Omega; X^*)$, $g = \sum_{k=1}^n x_k^* \mathbf{1}_{E_k}$, E_k disjoint. The norm is then given by $|g|_{L^{\overline{p}}(\Omega;X^*)} = \left(\sum_{k=1}^n |x_k^*|_{X^*}^{\overline{p}} \, \mu(E_k)\right)^{\frac{1}{\overline{p}}}$, and by the duality of the scalar-valued function spaces $L^p(\mathfrak{E})$ and $L^{\overline{p}}(\mathfrak{E})$, where $\mathfrak{E}$ is the finite algebra generated by $\{E_k\}_{k=1}^n$, there exist $a_k \ge 0$, $k = 1, \dots, n$, such that

$$|g|_{L^{\overline{p}}(\Omega;X^*)} = \sum_{k=1}^n a_k \, |x_k^*|_{X^*} \, \mu(E_k) \qquad \text{and} \qquad \left(\sum_{k=1}^n a_k^p \mu(E_k)\right)^{\frac{1}{p}} \le 1. \tag{A.2}$$

Given any $\epsilon > 0$, we can find $x_k \in X$ of norm at most 1 such that $\langle x_k^*, x_k\rangle \ge |x_k^*|_{X^*} - \epsilon$. Then $f := \sum_{k=1}^n a_k x_k \mathbf{1}_{E_k} \in S(\Omega; X)$ satisfies

$$|f|_{L^p(\Omega;X)} = \left(\sum_{k=1}^n a_k^p \, |x_k|_X^p \, \mu(E_k)\right)^{\frac{1}{p}} \le 1,$$

where the last inequality follows from $|x_k|_X \le 1$ and equation (A.2). Furthermore

$$\langle g^*, f\rangle_{L^p(\Omega;X)} = \sum_{k=1}^n a_k \, \langle x_k^*, x_k\rangle_X \, \mu(E_k) > \sum_{k=1}^n a_k \left(|x_k^*|_{X^*} - \epsilon\right) \mu(E_k) = |g|_{L^{\overline{p}}(\Omega;X^*)} - \epsilon \sum_{k=1}^n a_k \mu(E_k).$$

It follows that $|g^*|_{L^p(\Omega;X)^*} \ge |g|_{L^{\overline{p}}(\Omega;X)} - \epsilon \sum_{k=1}^n a_k \mu(E_k)$, and this being true for all $\epsilon > 0$, we have the desired equality. Thus $g \mapsto g^*$ is a linear isometry from $S(\Omega; X^*) \subset L^{\overline{p}}(\Omega; X^*)$ into $L^p(\Omega; X)^*$, and so extends to a linear isometry on all of $L^{\overline{p}}(\Omega; X)$. ☐

Lemma A.9. *Let $(\Omega, \mathfrak{F}, \mu)$ be a finite measure space, $p \in (1,\infty)$. If X^* has the Radon–Nikodým property with respect to $(\Omega, \mathfrak{F}, \mu)$, then $L^p(\Omega; X)^* = L^{\overline{p}}(\Omega; X^*)$.*

Proof. We must show that every $\Lambda \in L^p(\Omega; X)^*$ can be represented as an integral operator like the one in Lemma A.8. So let Λ be such a functional and define the vector-valued measure $\Psi : \mathfrak{F} \to X^*$ by

$$\langle \Psi(E), x\rangle_X := \langle \Lambda, x\mathbf{1}_E\rangle_{L^p(\Omega;X)} \qquad \text{for} \quad E \in \mathfrak{F}, \quad x \in X.$$

Then $|\langle\Psi(E),x\rangle_X| \le |\Lambda|_{L^p(\Omega;X)^*} |x|_X \, \mu(E)^{\frac{1}{p}}$, whence $|\Psi(E)|_{X^*} \le |\Lambda|_{L^p(\Omega;X)^*} \, \mu(E)^{\frac{1}{p}}$. It follows that Ψ is continuous on the null-set (which is equivalent to σ-additivity) since μ is; thus Ψ is a proper vector measure, and it is clearly absolutely continuous with respect to μ. Furthermore, for $E = \cup_{k=1}^n E_k$ (a disjoint union) and $x_k \in X$, $|x_k|_X \le 1$, we compute

$$\begin{aligned} \left|\sum_{k=1}^n \langle\Psi(E_k), x_k\rangle_X\right| &= \left|\left\langle \Lambda, \sum_{k=1}^n x_k \mathbf{1}_{E_k}\right\rangle_{L^p(\Omega;X)}\right| \\ &\le |\Lambda|_{L^p(\Omega;X)^*} \left(\sum_{k=1}^n |x_k|_X^p \, \mu(E_k)\right)^{\frac{1}{p}} \le |\Lambda|_{L^p(\Omega;X)^*} \, \mu(E)^{\frac{1}{p}}. \end{aligned}$$

Taking supremum over all disjoint unions $E = \sum_{k=1}^n E_k$ and all $x_k \in \overline{B}(0;1) \subset X$, we find that $|\Psi|_X (E) \le |\Lambda|_{L^p(\Omega;X)^*} \, \mu(E)^{\frac{1}{p}}$. With $E = \Omega$, we find that Ψ is of bounded variation. Since X^* has the Radon–Nikodým property with respect to $(\Omega, \mathfrak{G}, \mu)$, the X^*-valued measure Ψ has a Radon–Nikodým derivative $g \in L^1(\Omega; X^*)$ such that $\Psi(E) = \int_E g d\mu$. Thus

$$\langle\Lambda, x\mathbf{1}_E\rangle_{L^p(\Omega;X)} = \langle\Psi(E), x\rangle_X = \left\langle \int_E g(\omega) d\mu(\omega), x\right\rangle_X = \int_\Omega \langle g(\omega), x\mathbf{1}_E(\omega)\rangle_X \, d\mu(\omega),$$

and by linearity we may replace $x\mathbf{1}_E$ by any simple function f.

We must still show that $g \in L^{\overline{p}}(\Omega; X^*)$. To this end, observe that $\int_\Omega \langle g\mathbf{1}_{E_n}, \cdot\rangle_X \, d\mu$, $E_n := \{|g|_{X^*} \le n\}$, is a bounded linear functional on $L^p(\Omega; X)$ (since $g\mathbf{1}_{E_n}$ is bounded). Furthermore, this functional coincides with Λ for all simple functions supported on E_n, and by continuity for all $f \in L^p(E_n; X)$. Thus $g \in L^\infty(E_n; X^*) \subset L^{\overline{p}}(E_n; X^*)$ represents the linear functional $\Lambda|_{L^p(E_n;X)}$ in the sense of Lemma A.8, so by that lemma we have $|g|_{L^{\overline{p}}(E_n;X^*)} = |\Lambda|_{L^p(E_n;X)^*} \le |\Lambda|_{L^p(\Omega;X)^*}$. Since $E_n \uparrow \Omega$ as $n \uparrow \infty$, it follows from the monotone convergence theorem that also $|g|_{L^{\overline{p}}(\Omega;X^*)} \le |\Lambda|_{L^p(\Omega;X)^*}$. Then, with the help of Hölder's inequality we can apply the dominated convergence theorem to the equation

$$\langle\Lambda, f\mathbf{1}_{E_n}\rangle_{L^p(\Omega;X)} = \int_\Omega \langle g\mathbf{1}_{E_n}(\omega), f(\omega)\rangle_X \, d\mu(\omega)$$

to show that we actually have $\langle\Lambda, f\rangle_{L^p(\Omega;X)} = \int_\Omega \langle g, f\rangle_X \, d\mu$, so $g \in L^{\overline{p}}(\Omega; X^*)$ represents the arbitrarily chosen operator $\Lambda \in L^p(\Omega; X)^*$. Thus $L^p(\Omega; X)^* = L^{\overline{p}}(\Omega; X^*)$ with obvious identifications. ☐

The finiteness of the measure space in Lemma A.9 is not a serious restriction. For if $(\Omega, \mathfrak{F}, \mu)$ is a σ-finite measure space, we can find a positive function $w \in L^1(\Omega)$ so that $(\Omega, \mathfrak{F}, w\mu)$ is a finite measure space, and $f \mapsto w^{\frac{1}{p}} f$ is a unitary bijection from $L^p(w\mu; X)$ to $L^p(\mu; X)$. (See [20]). Furthermore, if $\Lambda \in L^p(\mu; X)^*$, then $\Lambda m_{w^{1/p}} \in L^p(w\mu; X)^*$, so if X^* has the Radon–Nikodým property with respect to $(\Omega, \mathfrak{F}, \mu)$, then $\Lambda m_{w^{1/p}}$ is represented by some $g \in L^{\overline{p}}(w\mu; X^*)$, i.e., for each $fw^{-\frac{1}{p}} \in L^p(w\mu; X)$ $(f \in L^p(\mu; X))$, we have

$$\langle\Lambda, f\rangle_{L^p(\mu;X)} = \left\langle \Lambda m_{w^{1/p}}, w^{-\frac{1}{p}} f\right\rangle_{L^p(w\mu;X)} = \int_\Omega \left\langle g, w^{-\frac{1}{p}} f\right\rangle_X w d\mu = \int_\Omega \left\langle w^{\frac{1}{\overline{p}}} g, f\right\rangle_X d\mu,$$

and $w^{\frac{1}{\overline{p}}} g \in L^{\overline{p}}(\mu; X)$, so $\Lambda \in L^p(\mu; X)^*$ is representable by this function, and $L^p(\mu; X)^* = L^{\overline{p}}(\mu; X^*)$.

A more difficult task is determining when the Radon–Nikodým property is satisfied by X^*. In one particular case it is easy:

Remark A.10. *Every Banach space X has the Radon–Nikodým property with respect to every $(\Omega, \mathfrak{F}, \mu)$, where $\mathfrak{F}$ is a finite algebra. Indeed, if $\operatorname{bs}\mathfrak{F} = \{F_k\}_{k=1}^n$, then*

$$\frac{d\Psi}{d\mu} = \sum_{k=1}^n \frac{\Psi(E_k)}{\mu(E_k)} \mathbf{1}_{E_k}.$$

Consequently, $L^p(\mathfrak{F};X)^* = L^{\overline{p}}(\mathfrak{F};X^*)$ whenever $\mathfrak{F}$ is a finite algebra.

The condition of a finite algebra is too restrictive for most applications though. The following result provides a more useful condition. Unfortunately, the limitations of space force us to omit the proof.

Lemma A.11. *Every reflexive Banach space has the Radon–Nikodým property with respect to every finite measure space. Consequently, $L^p(\Gamma;X)^* = L^{\overline{p}}(\Gamma;X^*)$ whenever $p \in (1,\infty)$, Γ is σ-finite and X is reflexive.*

Observe that "finite" now refers to $\mu(\Omega) < \infty$, as usual.

Proof. The proof is found in [6], Corollary III.2.13. □

A.4 Vector-valued extensions of linear operators

Many operators classically viewed as acting on scalar-valued functions can be extended to the vector-valued setting in the following canonical way:

Definition A.12. *Let $T \in \mathcal{B}(L^p(\Omega);L^q(\Omega))$. We say that $\widetilde{T} \in \mathcal{B}(L^p(\Omega;X);L^q(\Omega;X))$ is an **X-valued extension** of T if*

$$\widetilde{T}\left(\sum_{k=1}^{n} x_k f_k\right) = \sum_{k=1}^{n} x_k T f_k \tag{A.3}$$

for all $n \in \mathbb{Z}_+$, $x_k \in X$, $f_k \in L^p(\Omega)$, $k = 1,\dots,n$.

Remark A.13.

1. *Owing to the density of simple integrable functions in L^p, $1 \le p < \infty$, condition (A.3) can be replaced by the formally weaker requirement*

   $$\widetilde{T}\left(\sum_{k=1}^{n} x_k \mathbf{1}_{E_k}\right) = \sum_{k=1}^{n} x_k T \mathbf{1}_{E_k}$$

   for all $n \in \mathbb{Z}_+$, $x_k \in X$, E_k of finite measure, $k = 1,\dots,n$.

2. *The vector-valued extension is always unique (when it exists); indeed, its value is given for every simple function, so the general case follows from the requirement that the extension be continuous.*

The observation that $S(\Omega;X) \subset X \bigotimes L^p(\Omega) := \{\sum_{k=1}^n x_k f_k : x_k \in X; f_k \in L^p(\Omega); k = 1,\dots,n; n \in \mathbb{N}\}$, and the consequent density of $X \bigotimes L^p(\Omega)$ in $L^p(\Omega;X)$ are occasionally useful.

An example of a vector-valued extension is the vector-valued conditional expectation $\mathbb{E}(\,\cdot|\,\mathfrak{G})$ constructed in Section 5.2. The fact that $\mathbb{E}(\,\cdot|\,\mathfrak{G})$ acting on $L^1(\mathfrak{F};X)$ is an extension of $\mathbb{E}(\,\cdot|\,\mathfrak{G})$ acting on $L^1(\mathfrak{F})$ is sometimes emphasized by writing $\mathbb{E}_X(\cdot|\mathfrak{G})$ instead of $\mathbb{E}(\,\cdot|\,\mathfrak{G})$ when the vector-valued operator is in question. We find this unnecessary, for it should always be clear from the context which operator is used. In the extension lemmas below, we nevertheless follow this convention for clarity.

There are several results concerning vector-valued extensions. The simplest is probably the following one, which guarantees the existence of an X-valued extension, with the same norm, of every bounded positive linear operator on $L^p(\Omega)$.

Lemma A.14. *Let Ω be a σ-finite measure space, X a Banach space and $p,q \in [1,\infty)$. Then every positive $T \in \mathcal{B}(L^q(\Omega);L^p(\Omega))$ has an X-valued extension T_X of the same norm.*

The existence of the vector-valued conditional expectation could have been derived from this general result, but due to the fundamental role of $\mathbb{E}(\,\cdot|\,\mathfrak{G})$, we found it instructive to give a separate proof in the main text.

Proof. Define $\widetilde{T}$ on simple functions by

$$\widetilde{T}\left(\sum_{k=1}^{n} x_k \mathbf{1}_{E_k}\right) := \sum_{k=1}^{n} x_k T\mathbf{1}_{E_k},$$

whenever $x_k \in X$ and the E_k are measurable. Then

$$\left|\widetilde{T}\left(\sum_{k=1}^{n} x_k \mathbf{1}_{E_k}\right)(\omega)\right|_X \leq \sum_{k=1}^{n} |x_k|_X \, T\mathbf{1}_{E_k}(\omega) = T\left(\sum_{k=1}^{n} |x_k|_X \, \mathbf{1}_{E_k}\right)(\omega) = T\left|\sum_{k=1}^{n} x_k \mathbf{1}_{E_k}\right|_X (\omega)$$

by the triangle inequality and the fact that $T\mathbf{1}_{E_k} \geq 0$ since $\mathbf{1}_{E_k} \geq 0$. We assumed that the E_k are disjoint to obtain the last equality.

We have shown, for simple f, that $\left|\widetilde{T}f\right|_X(\omega) \leq T\,|f|_X(\omega)$. Thus

$$\left(\int_\Omega \left|\widetilde{T}f\right|_X^p (\omega) d\mu(\omega)\right)^{\frac{1}{p}} \leq \left(\int_\Omega \left(T\,|f|_X(\omega)\right)^p d\mu(\omega)\right)^{\frac{1}{p}} \leq |T|_{\mathcal{B}(L^q(\Omega);L^p(\Omega))} \left(\int_\Omega |f(\omega)|_X \, d\mu(\omega)\right)^{\frac{1}{q}}.$$

Hence $\left|\widetilde{T}f\right|_{L^p(\Omega;X)} \leq |T|_{\mathcal{B}(L^q(\Omega);L^p(\Omega))} \, |f|_{L^q(\Omega;X)}$ for all simple f. The operator $\widetilde{T}$ then extends to a bounded operator on all of $L^q(\Omega;X)$ with the same norm, and the assertion is established. □

The following results show that once we have the extension of an operator in one Banach space, we also have it in certain others.

Lemma A.15. *Let X be a Banach-space, Ω, Γ σ-finite measure spaces and $1 < p < \infty$. If $T \in \mathcal{B}(L^p(\Omega))$ has an X-valued extension T_X, then T also has an $L^p(\Gamma;X)$-valued extension $T_{L^p(\Gamma;X)}$, which satisfies $\left|T_{L^p(\Gamma;X)}\right|_{\mathcal{B}(L^p(\Omega;L^p(\Gamma;X)))} \leq |T_X|_{\mathcal{B}(L^p(\Omega;X))}$.*

Proof. The operator $\widetilde{T} := T_{L^p(\Gamma;X)}$ should satisfy

$$\widetilde{T}\left(\sum_{k=1}^{n} g_k f_k\right) = \sum_{k=1}^{n} g_k T f_k$$

for $n \in \mathbb{N}$, $g_k \in L^p(\Gamma;X)$, $f_k \in L^p(\Omega)$.

By the density of $X \bigotimes L^p(\Gamma)$ in $L^p(\Gamma;X)$ and the required continuity and linearity of $\widetilde{T}$, it suffices to construct a $\widetilde{T}$ satisfying the previous equality for all $g_k \in X \bigotimes L^p(\Gamma)$, i.e., we need to have

$$\widetilde{T}\left(\sum_{k=1}^{n} f_k h_k x_k\right) = \sum_{k=1}^{n} (Tf_k) h_k x_k \tag{A.4}$$

for $f_k \in L^p(\Omega)$. $h_k \in L^p(\Gamma)$ and $x_k \in X$.

Since the right-hand side of (A.4) is well-defined, we can take (A.4) as the definition of an operator $\widetilde{T}$ on $L^p(\Omega) \bigotimes L^p(\Gamma) \bigotimes X$. For the operator defined like this we compute

$$\begin{aligned}
\left|\sum_{k=1}^{n} (Tf_k) h_k x_k\right|_{L^p(\Omega;L^p(\Gamma;X))}^p &= \int_\Gamma \left|\sum_{k=1}^{n} (Tf_k)(\cdot) h_k(\gamma) x_k\right|_{L^p(\Omega;X)}^p d\mu(\gamma) \\
= \int_\Gamma \left|T_X\left(\sum_{k=1}^{n} f_k(\cdot) h_k(\gamma) x_k\right)\right|_{L^p(\Omega;X)}^p d\mu(\gamma) &\leq |T_X|_{\mathcal{B}(L^p(\Omega;X))}^p \int_\Gamma \left|\sum_{k=1}^{n} f_k(\cdot) h_k(\gamma) x_k\right|_{L^p(\Omega;X)}^p d\mu(\gamma) \\
&= |T_X|_{\mathcal{B}(L^p(\Omega;X))}^p \left|\sum_{k=1}^{n} f_k h_k x_k\right|_{L^p(\Omega;L^p(\Gamma;X))}^p.
\end{aligned}$$

This computation applied to the difference of two expressions $\sum_{k=1}^{n} f_k h_k x_k$ representing the same function shows that $\widetilde{T}$ is well defined, and thus it defines a linear operator whose norm is bounded by $|T_X|_{L^p(\Omega;X)}$ by the same calculation.

The operator $\widetilde{T}$ then extends to a bounded linear operator on all of $L^p(\Omega; L^p(\Gamma; X))$, since $L^p(\Omega) \bigotimes L^p(\Gamma) \bigotimes X$ is dense in $L^p(\Omega; L^p(\Gamma; X))$; indeed, any function f in this last space can be approximated in the L^p norm by a simple function $\sum_{k=1}^n g_k \mathbf{1}_{E_k}$, $g_k \in L^p(\Gamma; X)$, and each g_k can again be approximated by a simple function $\sum_{j=1}^{m_k} x_j^k \mathbf{1}_{F_j^k}$, $x_j^k \in X$. This completes the proof. □

Lemma A.16. *Let $p, q \in (1, \infty)$. If $T \in \mathcal{B}(L^q(\Omega); L^p(\Omega))$ has an X-valued extension $T_X \in \mathcal{B}(L^q(\Omega; X); L^p(\Omega; X))$, then the dual operator $(T_X)^* \in \mathcal{B}(L^p(\Omega; X)^*; L^q(\Omega; X)^*)$, restricted to $L^{\overline{p}}(\Omega; X^*) \subset L^p(\Omega; X)^*$, is the X^*-valued extension of $T^* \in \mathcal{B}(L^{\overline{p}}(\Omega); L^{\overline{q}}(\Omega))$.*

In particular, since dual operators have equal norms, we find that T^* has an X^*-valued extension whenever T has an X-valued extension, of no larger norm.

Proof. Consider functions $f := \sum_{k=1}^n x_k^* f_k \in L^{\overline{p}}(\Omega) \bigotimes X^* \subset L^{\overline{p}}(\Omega; X^*) \subset L^p(\Omega; X)^*$ and $g := \sum_{j=1}^m x_j g_j \in L^q(\Omega) \bigotimes X \subset L^q(\Omega; X)$. Then we compute

$$\begin{aligned}
\langle (T_X)^* f, g \rangle_{L^q(\Omega;X)} &= \left\langle f, T_X\left(\sum_{j=1}^m x_j g_j\right)\right\rangle_{L^p(\Omega;X)} = \left\langle f, \sum_{j=1}^m x_j T g_j \right\rangle_{L^p(\Omega;X)} \\
&= \sum_{k,j} \langle x_k^*, x_j \rangle_X \langle f_k, T g_j \rangle_{L^p(\Omega)} = \sum_{k,j} \langle x_k^*, x_j \rangle_X \langle T^* f_k, g_j \rangle_{L^q(\Omega)} \\
&= \left\langle \sum_{k=1}^n x_k^* T^* f_k, \sum_{j=1}^n x_j g_j \right\rangle_{L^q(\Omega;X)}.
\end{aligned}$$

Since this is true for all simple $g \in L^q(\Omega) \bigotimes X$, a dense subset of $L^q(\Omega; X)$, we conclude that

$$(T_X)^* \left(\sum_{k=1}^n x_k^* f_k\right) = \sum_{k=1}^n x_k^* T^* f_k,$$

i.e., (the restriction of) $(T_X)^*$ is the desired X^*-valued extension as asserted. □

A.5 Functions on Euclidean spaces

For X-valued functions f defined on $\mathbb{R}^d$, which are measurable with respect to the Lebesgue measure $m = m_d$, we can discover some further structure resulting from the properties of the Lebesgue measure and the geometry of the Euclidean spaces. We begin with the change-of-variable formula. The version we give is quite a bit more powerful than what we will need; however, since the vector-valued result follows almost instantly from the scalar case, we can equally well present the general form.

Lemma A.17 (Change of variable). *Let $E \subset G \subset \mathbb{R}^d$, E Lebesgue-measurable and G open. Let $g : G \to \mathbb{R}^d$ be continuous on G and differentiable at each point of E. Let $g|_E$ be one-to-one and $m(g(G \setminus E)) = 0$.*

Then, for $f \in L^1(\mathbb{R}^d; X)$,

$$\int_{g(E)} f dm = \int_E f \circ g \cdot |\det Dg| \, dm. \tag{A.5}$$

Proof. For $f = x\mathbf{1}_F$, $m(F) < \infty$, we can extract the x from both sides of the asserted equality (A.5), and the claim reduces to the scalar-valued theorem [20]. By linearity of the integral, we know that (A.5) is valid for all simple integrable functions. For an arbitrary $f \in L^1(\mathbb{R}^d; X)$, we can take simple integrable f_k, which tend to f in $L^1(\mathbb{R}^d; X)$ (by Lemma A.3). Then clearly

$\int_{g(E)} f dm = \lim_{k\to\infty} f_k dm$. Applying the change-of-variable formula to the positive function $|f - f_k|_X$ we also deduce

$$\int_E |f - f_k|_X \circ g \cdot |\det Dg|\, dm = \int_{g(E)} |f - f_k|_X \, dm \le |f - f_k|_{L^1(\mathbb{R}^d;X)} \,,$$

and this tends to zero as $k \to \infty$. Thus $\int_E f \circ g \cdot |\det Dg|\, dm = \lim_{k\to\infty} \int_E f_k \circ g \cdot |\det Dg|\, dm$ and the claim follows for general $f \in L^1(\mathbb{R}^d; X)$. □

Another question of interest is the existence of Lebesgue points. This matter can be settled with the aid of the Hardy–Littlewood maximal function

$$Mf(t) := \sup_{0<r<\infty} \frac{1}{m(B(t;r))} \int_{B(t;r)} |f|_X \, dm$$

as in real analysis with almost no modifications. Since it is clear from above that $Mf = M\,|f(\cdot)|_X$, and of course, for measurable $f : \Omega \to X$, we have $f \in L^1(\Omega; X)$ if and only if $|f(\cdot)|_X \in L^1(\Omega)$, and the norms agree, we can immediately generalize the familiar weak type inequality for the maximal function (see [20] or [23]) to obtain

$$m(Mf > \lambda) \le \frac{c_d}{\lambda} |f|_{L^1(\Omega;X)} \,.$$

To deduce the existence of Lebesgue points, we also need to be able to approximate integrable functions by continuous ones. In fact, it is easy to obtain a much stronger result, which will be useful later. We need a preliminary result concerning smooth scalar-valued functions on $\mathbb{R}^d$.

Lemma A.18. *If $K \subset G \subset \mathbb{R}^d$ with K compact and G open, then there exists a $\psi \in C^\infty(\mathbb{R}^d)$ such that $\psi|_K = 1$, $\psi|_{G^c} = 0$ and $\operatorname{ran}\psi = [0,1]$.*

Proof. The function $x \mapsto e^{-\frac{1}{x}}$ on $\mathbb{R}_+$ tends to 0, together with all its derivatives as $x \to 0$. Since this function is clearly C^∞ on $\mathbb{R}_+$, it follows that $\phi(x) := e^{-\frac{1}{x}} \mathbf{1}_{(0,\infty)}(x)$ is C^∞ on $\mathbb{R}$, and vanishes on $(-\infty, 0]$. Then, for $a < b$, $\varphi_{a,b}(x) := \phi(x-a)\phi(b-x)$ is positive for $x \in (a,b)$ and vanishes elsewhere, and clearly $\varphi_{a,b} \in C^\infty$. Furthermore, for a box $Q = (a_1, b_1) \times \cdots \times (a_d, b_d) \subset \mathbb{R}^d$, the C^∞ function $\varphi_Q(x) := \varphi_{a_1,b_1}(x^1) \cdots \varphi_{a_d,b_d}(x^d)$ is positive on Q and vanishes elsewhere. Also, the function

$$\Psi_\epsilon(x) := \frac{\int_0^x \varphi_{0,\epsilon} dt}{\int_0^\epsilon \varphi_{0,\epsilon} dt} \mathbf{1}_{\mathbb{R}_+}(x)$$

is C^∞, vanishes on $(-\infty, 0]$, and increases on $(0, \epsilon)$ to the constant value 1 on $[\epsilon, \infty)$.

Since G is open, for every $t \in G$, we can find a box Q_x such that $x \in Q_x \subset G$. The boxes Q_x, $x \in G$, clearly cover the compact set K; whence a finite number of them, say $Q_1, \dots, Q_k$, cover K. Then the function $\Phi := \sum_{j=1}^k \varphi_{Q_j}$ is positive on K and vanishes outside $\cup_{j=1}^k Q_i$, in particular outside G.

The continuous function Φ attain a minimum on K, say $\epsilon > 0$ (since the function is positive on K). A function of the desired type is now given by $\Psi_\epsilon \circ \Phi$. □

Lemma A.19. *Functions of the form $\sum_{k=1}^n x_k \psi_k$, $x_k \in X$, $\psi_k \in \mathcal{D}(\mathbb{R}^d)$ (infinitely differentiable with compact support) are dense in $L^p(\mathbb{R}; X)$, $1 \le p < \infty$.*

It is obvious that the functions $\sum_{k=1}^n x_k \psi_k$ are continuous with compact support. In fact, they are in the space $\mathcal{D}(\mathbb{R}^d; X)$ to be defined below.

Proof. By Lemma A.3, simple integrable functions, i.e., ones of the form $\sum_{k=1}^n x_k \mathbf{1}_{E_k}$, $m(E_k) < \infty$, are dense in $L^p(\mathbb{R}^d; X)$, $1 \le p < \infty$. Thus it clearly suffices to show that $x\mathbf{1}_E$ can be approximated in the $L^p(\mathbb{R}^d; X)$ norm arbitrarily well by functions of the form $x\psi$, $\psi \in \mathcal{D}(\mathbb{R}^d)$. This follows readily: From the properties of the Lebesgue measure we know [20] that $m(E)$ can be approximated from below and above by $m(K)$, $K \subset E$ compact, and $m(G)$, $G \supset E$ open. But for $K \subset G \subset \mathbb{R}^d$

as above, one can find (by Lemma A.18) a $\psi \in \mathcal{D}(\mathbb{R}^d)$ such that $\operatorname{ran}\psi = [0,1]$ with $\psi|_K = 1$, $\psi|_{G^c} = 0$. Then $|x\mathbf{1}_E - x\psi|_{L^p(\mathbb{R}^d;X)} \leq |x|_X\, m(G^c \setminus K) \leq |x|_X\,(m(G^c \setminus E) + m(E \setminus K))$, and the measures of the two difference sets in the last expression can be made as small as one likes. ☐

The existence of Lebesgue points now follows:

Lemma A.20. *If $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d;X)$ then almost every $t \in \mathbb{R}^d$ satisfies*

$$\lim_{r\to 0} \frac{1}{m(B(t;r))} \int_{B(t;r)} |f - f(t)|_X\, dm = 0. \tag{A.6}$$

(The existence of the limit is part of the assertion.)

A $t \in \mathbb{R}^d$ satisfying (A.6) is called a **Lebesgue point** of f; the set of all such points is denoted by $\mathfrak{L}(f)$. It is easy to see that every point of continuity of f is a Lebesgue point of f.

Once Lemma A.20 is shown, it follows from (A.6) that a similar convergence is true at every Lebesgue point even if the balls $B(x;r)$ are replaced by any sequence of measurable sets E_k **shrinking nicely** to t, i.e. satisfying $E_k \subset B(x;r_k)$, $m(E_k) > cm(B(x;r_k))$ for some fixed $c \in (0,1)$ and a sequence $\{r_k\}_{k=1}^\infty \subset \mathbb{R}_+$ with $r_k \to 0$. This is an immediate consequence of the observation

$$\frac{1}{m(E_k)} \int_{E_k} |f - f(t)|_X\, dm \leq \frac{m(B(x;r_k))}{m(E_k)} \frac{1}{m(B(x;r_k))} \int_{B(x;r_k)} |f - f(t)|_X\, dm,$$

and $\frac{m(B(x;r_k))}{m(E_k)} \leq \frac{1}{c}$. An obvious corollary is the convergence (in the norm of X)

$$\lim_{k\to\infty} \frac{1}{m(E_k)} \int_{E_k} f dm = f(t), \qquad t \in \mathfrak{L}(f), \qquad \{E_k\}_{k=1}^\infty \text{ shrink nicely to } t. \tag{A.7}$$

Now we turn to the proof.

Proof. Observe first that we loose no generality in assuming that $f \in L^1(\mathbb{R}^d;X)$: If f is locally integrable, then $f\mathbf{1}_{B(t;r)} \in L^1(\mathbb{R}^d;X)$, and the value of the left-hand side is the same for f and $f\mathbf{1}_{B(t;r)}$. Since countably many balls $B(t;r)$ cover $\mathbb{R}^d$, the asserted identity (A.6) is true almost every where on $\mathbb{R}^d$ if it is true almost everywhere on each of the countable number of balls. Thus assume $f \in L^1(\mathbb{R}^d;X)$.

We denote $Lf(t) := \limsup_{r\to 0} \frac{1}{m(B(t;r))} \int_{B(t;r)} |f - f(t)|_X\, dm$; we must show that $Lf = 0$ a.e. for $f \in L^1(\mathbb{R}^d;X)$. (Observe that the expression for $Lf(t)$ is the same as (A.6), with limit replaced by upper limit; for this latter one, the existence is guaranteed.)

L is not a linear operator, but it is readily seen that it is subadditive; $L(g+h) \leq Lg + Lh$ at each point. Since continuous compactly supported functions are dense in $L^1(\mathbb{R}^d;X)$ by Lemma A.19, we can decompose f as $g+h$, where $g \in C_c(\mathbb{R}^d;X)$ and $h = f - g$ can be made as small as one likes in the $L^1(\mathbb{R}^d;X)$ norm.

Then $Lf(t) = L(g+h)(t) \leq Lg(t) + Lh(t) = Lh(t) \leq Mh(t) + |h(t)|_X$, where the last inequality is immediate from the definitions of L and the maximal operator M. Therefore

$$m(Lf > 2\epsilon) = m(Lh > 2\epsilon) \leq m(Mh > \epsilon) + m(|h|_X > \epsilon) \leq \frac{c_d + 1}{\epsilon} |h|_{L^1(\mathbb{R}^d;X)}.$$

Since the L^1 norm of h is at our disposal and the left-hand side of the above estimate is independent of h, it follows that $m(Lf > 2\epsilon)$ must vanish. Since $\{Lf > 0\}$ is a countable union of such sets, it follows that Lf must vanish almost everywhere, as we asserted. ☐

The abundance of Lebesgue points also immediately yields positive information about differentiability properties of measurable functions. It is first in order to make the appropriate definitions, although these are essentially the same as in elementary calculus.

For any function $f : \mathbb{R}^d \to X$, we can consider the difference quotients $\frac{1}{h}(f(t+he_i) - f(t))$. If these converge (in norm) to a point of X as $h \to 0$, we call this point the **partial derivative** (with respect to the ith coordinate) of f at t, and denote it by $D_i f(t)$ (or $\frac{\partial f}{\partial t_i}(t)$ or other similar familiar symbols). If f has a certain partial derivative at all points of some set $E \subset \mathbb{R}^d$, these define a function, denoted by $D_i f$ on E. If this function is continuous on E, we say that f is continuously differentiable on E with respect to the ith coordinate. If (an X-valued) f is continuously differentiable on E with respect to all the coordinates, we denote $f \in C^1(E;X)$ and say that f is continuously differentiable. More generally, the class of those $f : E \to X$ for which $D_{i_1} \cdots D_{i_r} f$ is continuous for all $r \leq k$ and all choices of $i_1 \ldots, i_r \in \{1, \ldots, n\}$ is denoted by $C^k(E;X)$.

To facilitate notation involved in integration and partial differentiation with respect to some but not necessarily all the variables, we introduce the notation

$$f_i(\cdot; a) := f(a_1, \ldots, a_{i-1}, \cdot, a_{i+1}, \ldots, a_d)$$

for the restriction of f to one variable, the other variables held constant at the values given by the corresponding coordinates of a. More generally, we use $f_I(\cdot; a)$, $I \subset \{1, \ldots, n\}$ to denote the restriction of f to the coordinates $i \in I$; the other variables are held constant at a, which is interpreted as a parameter of f_I. Similarly, t_I denotes the vector $(t_i)_{i \in I} \subset \mathbb{R}^{\# I}$, for any $t \in \mathbb{R}^d$. Since in the definition of the partial derivative all variables except one are held constant, it is clear that $D_i f(a) = D_i f_I(a_I; a)$, for any $I \ni i$.

The fundamental theorem of calculus is naturally of interest. One half is easy:

Lemma A.21. *If $f \in L^1([a,b];X)$, then $D \int_a^x f dt = f$ on Lebesgue points of f on $[a,b] \subset \mathbb{R}$, thus a.e. on this interval, and everywhere, if $f \in C([a,b];X)$.*

Note that it follows from Fubini's theorem that $f_i(\cdot; a) \in L^1(\mathbb{R};X)$ for almost every a if $f \in L^1(\mathbb{R}^d;X)$.

Proof. Obviously the sets $E_k := (t, t+h_k)$ if $h_k > 0$, $E_k := (t+h_k, t)$ otherwise, shrink nicely to t for any $h_k \to 0$. Thus $\frac{1}{h_k}\left(\int_a^{t+h_k} f dm - \int_a^t f dm\right) = \frac{1}{m(E_k)} \int_{E_k} f dm \to f(t)$ for $t \in \mathfrak{L}(f)$ by (A.7) resulting from Lemma A.20. □

The version of the fundamental theorem with the order of differentiation and integration reversed is more involved (and needs additional assumptions) even in the real-valued case, and it is not surprising that the vectors do not simplify this. However, in the sequel we will concentrate on functions with more regularity properties, and in this case things can be settled quite readily.

Lemma A.22. *If, for $f : (a,b) \to X$, $Df = 0$ everywhere, then f is a constant.*

Recall that this result fails even for $X = \mathbb{R}$ if "everywhere" is replaced by "a.e.", unless more assumptions are imposed on f.

Proof. For each $t \in (a,b)$, each $x^* \in X^*$, we have

$$\left|\frac{1}{h}(\langle x^*, f(t+h)\rangle - \langle x^*, f(t)\rangle)\right| = \left|\left\langle x^*, \frac{1}{h}(f(t+h) - f(t))\right\rangle\right| \leq \left|\frac{1}{h}(f(t+h) - f(t))\right|_X \to 0$$

as $h \to 0$; thus the scalar-valued functions $\langle x^*, f(\cdot)\rangle$ have a vanishing derivative at each point of (a,b). Thus we know from elementary calculus that $\langle x^*, f(t_1)\rangle = \langle x^*, f(t_2)\rangle$ for any points $t_1, t_2 \in (a,b)$. Since X^* separates the points of X, we conclude that $f(t_1) = f(t_2)$, i.e. f is a constant. □

Lemma A.23. *If $f \in C^1([a,b];X)$, then $\int_a^t Df dm = f(t) - f(a)$ for all $t \in [a,b]$.*

Proof. Denote $g(t) := \int_a^t Df dm - f(t)$. Then $Dg = Df - Df = 0$ at every $t \in [a,b]$. Thus $g = g(a) = -f(a)$, i.e., $\int_a^t Df dm = f(t) - f(a)$ for all $t \in [a,b]$. □

Before we can properly use the usual notation related to partial derivatives of higher order, the commutativity of the partial derivative operators on a class of sufficiently smooth vector-valued functions must be established.

Lemma A.24. *If $D_i f$, $D_j f$, $D_i D_j f$ and $D_j D_i f$ are continuous in an open set G, then the two mixed derivatives coincide at each point of G.*

Iterative application of this result shows that operator products of at most k partial derivative operations on f can be performed in any order, given that $f \in C^k(\mathbb{R}^d; X)$; and for $f \in C^\infty$, the order of any finite number of partial derivatives is completely immaterial. When restricting the domain of the operators D_i to such functions, it follows that all operator products of partial derivative operations can be represented in the form $D_1^{\alpha_1} \dots D_d^{\alpha_d}$, $\alpha := (\alpha_i)_{i=1}^d \in \mathbb{N}^d$. This operator is given the usual short-hand notation D^α.

Proof. Let us fix a point $a \in G$; we must show that $D_i D_j f(a) = D_j D_i f(a)$. Since for any function g, $g_I(\cdot; a)$ is continuous whenever g is, and since $D_i g(a) = D_i g_I(a_I; a)$ for $I \ni i$, we can clearly consider $f_{\{i,j\}}(\cdot; a)$ instead of f. This is a function of two variables (which can obviously be named 1 and 2), and thus the claim reduces to showing that $D_1 D_2 g(a_1, a_2) = D_2 D_1 g(a_1, a_2)$ whenever the functions $D_1 g$, $D_2 g$, $D_1 D_2 g$ and $D_2 D_1 g$ are continuous in a neighbourhood of (a_1, a_2). So let the assumptions of this reduced assertion be satisfied and choose any two points (b_1, b_2), (c_1, c_2) such that $[b_1, c_1] \times [b_2, c_2]$ lies inside the above mentioned neighbourhood, say B, in which the four derivative functions are continuous.

Using Lemma A.23, we compute

$$\begin{aligned}\int_{b_1}^{c_1} dt_1 \int_{b_2}^{c_2} D_2 D_1 g(t_1, t_2) dt_2 &= \int_{b_1}^{c_1} \left(D_1 g(t_1, c_2) - D_1 g(t_1, b_2)\right) dt_1 \\ &= g(c_1, c_2) - g(b_1, c_2) - g(c_1, b_2) + g(b_1, b_2)\end{aligned}$$

and similarly $\int_{b_2}^{c_2} dt_2 \int_{b_1}^{c_1} D_1 D_2 g(t_1, t_2) dt_1$ yields the same result. But by Fubini's theorem, each of these integrals is equal to an integral over $[b_1, c_1] \times [b_2, c_2]$ with respect to the two dimensional Lebesgue measure $m_2 = m_1 \times m_1$. Thus we have deduced that $\int_Q D_1 D_2 g dm_2 = \int_Q D_2 D_1 g dm_2$ for any box $Q = [b_1, c_1] \times [b_2, c_2] \subset B$.

Let us now assume, contrary to the assertion, that $x := D_1 D_2 g(a) \neq D_2 D_1 g(a) =: y$ for some $a \in B$. By the assumed continuity, we can find a box Q around a such that $|D_1 D_2 g(t) - x|_X < \frac{1}{3}|x - y|_X$ and $|D_2 D_1 g(t) - y|_X < \frac{1}{3}|x - y|_X$ for $t \in Q$. But then

$$\begin{aligned}\left|\int_Q D_1 D_2 g dm - \int_Q D_2 D_1 g dm\right|_X &\geq \left|\int_Q (x - y) dm\right|_X - \int_Q \left(|D_1 D_2 g - x|_X + |D_2 D_1 - y|_X\right) dm \\ &\geq |x - y|_X\, m(Q) - 2 \cdot \frac{1}{3}|x - y|_X\, m(Q) > 0,\end{aligned}$$

and this contradicts the fact that the two integrals, the difference of which occurs on the left-hand side, are equal. ☐

It is obvious that the vector-valued partial derivatives satisfy the usual linearity property $D_i(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 D_i f_1 + \lambda_2 D_2 f_2$, whenever the right-hand side exists. The derivatives of a product of an operator-valued and a vector-valued function can be computed as in the case of two scalar functions:

$$\begin{aligned}&\frac{1}{h}\left(G(t + he_i) f(t + he_i) - G(t) f(t)\right) \\ &\quad = \frac{1}{h}\left(G(t + he_i) - G(t)\right) f(t) + G(t + he_i) \frac{1}{h}\left(f(t + he_i) - f(t)\right) \underset{h \to 0}{\longrightarrow} D_i G(t) f(t) + G(t) D_i f(t),\end{aligned}$$

given that G is strongly differentiable (i.e., the difference quotients converge strongly to the derivative) with respect to the ith coordinate and norm-continuous at t, and f is differentiable at the same point. Differentiability of both functions in norm is more than sufficient.

Iterative application of this result with the help of Lemma A.24 gives the usual Leibniz rule

$$D^\alpha(fg) = \sum_{\theta\le\alpha} \binom{\alpha}{\theta} D^{\alpha-\theta} f \cdot D^\theta g,$$

for $\alpha \in \mathbb{N}^d$, whenever $f \in C^{|\alpha|_1}(\mathbb{R}^d; X)$, $g \in C^{|\alpha|_1}(\mathbb{R}^d)$. (We use the multi-index notation $\binom{\alpha}{\theta} := \frac{\alpha!}{\theta!(\alpha-\theta)!}$, where $\alpha! := \alpha_1! \cdots \alpha_d!$.)

There is no difficulty in extending Taylor's theorem to the vector-valued setting. Observe, however, that the many proofs from elementary calculus which exploit the mean-value theorem are not applicable here. The mean-value theorem is not true in the vector-valued setting, even for $X = \mathbb{R}^2$, as the example $f(t) := (\sin t, \cos t - 1)$ shows: Clearly $f(0) = f(2\pi) = 0$, but $|f'(t)| = 1$ for all t.

Lemma A.25 (Taylor's theorem). *For $f \in C^{n+1}(\mathbb{R}^d; X)$,*

$$\left| f(t+h) - \sum_{k=0}^n \frac{1}{k!}(h\cdot D)^k f(t) \right|_X \le \frac{1}{(n+1)!} \max_{\lambda\in[0,1]} \left|(h\cdot D)^{n+1} f(x+\lambda h)\right|_X$$

Proof. Consider first the one-dimensional case, and denote

$$g(t,s) := \sum_{k=0}^n D^k f(s) \frac{(t-s)^k}{k!}.$$

Then $g(t,t) = f(t)$, and

$$D_2 g(t,s) = \sum_{k=0}^n D^{k+1} f(s) \frac{(t-s)^k}{k!} - \sum_{k=1}^n D^k f(s) \frac{(t-s)^{k-1}}{(k-1)!} = D^{n+1} f(s) \frac{(t-s)^n}{n!}.$$

Since $g(t,s) - g(t,t) = \int_t^s D_2 g(t,r) dr$, we have the estimate

$$\left| \sum_{k=0}^n D^k f(s) \frac{(t-s)^k}{k!} - f(t) \right|_X \le \int_t^s |D_2 g(t,r)|_X \, dr \le \max \left|D^{n+1} f(r)\right|_X \int_t^s \frac{|t-r|^n}{n!} dr$$

The d-dimensional case follows in the usual way by investigating the function $\mathbb{R} \ni h \mapsto f(t + he_i)$. $\square$

We conclude this section with a brief comment on the test function space $\mathcal{D}(\mathbb{R}^d; X) := C_c^\infty(\mathbb{R}^d; X)$. By the product rule of differentiation, functions of the form $\sum_{k=1}^d x_k \psi_k$, with $x_k \in X$, $\psi_k \in \mathcal{D}(\mathbb{R}^d)$, are in $\mathcal{D}(\mathbb{R}^d; X)$. Then Lemma A.18 provides us with many examples of $\mathcal{D}(\mathbb{R}^d; X)$ functions and Lemma A.19 shows that they are dense in each of the spaces $L^p(\mathbb{R}^d; X)$, $1 \le p < \infty$.

A.6 Fourier transform

The Fourier transform of $f \in L^1(\mathbb{R}^d; X)$ is defined as one could expect:

$$\mathcal{F}f(\xi) := \widehat{f}(\xi) := \int_{\mathbb{R}^d} f(t) e^{-\mathrm{i}2\pi t\cdot\xi} dt, \qquad \mathcal{F}^* f(\xi) := \check{f}(\xi) := \int_{\mathbb{R}^d} f(t) e^{\mathrm{i}2\pi t\cdot\xi} dt. \tag{A.8}$$

Observe that $\mathcal{F}^*(f) = \mathcal{F}(\widetilde{f}) = \widetilde{\mathcal{F}f}$ (by the change-of-variable formula; $\widetilde{f}(t) := \mathcal{R}f(t) := f(-t)$ is the reflection). This gives some taste of the inversion formula, which is the main reason for introducing the operator $\mathcal{F}^*$. The properties of the two transforms are essentially identical. In the sequel, we mostly consider $\mathcal{F}$ only.

It follows immediately from the definition that

$$\left|\widehat{f}\right|_{L^\infty(\mathbb{R}^d;X)} \leq \int_{\mathbb{R}^d} |f(t)|_X \, dt = |f|_{L^1(\mathbb{R}^d;X)} \, . \tag{A.9}$$

The convolution of two Borel functions f and g, one of which is operator-valued, is defined by

$$f * g(t) := \int_{\mathbb{R}^d} f(t-s)g(s)ds = \int_{\mathbb{R}^d} f(s)g(t-s)ds, \tag{A.10}$$

whenever the integrand is integrable. (The last equality follows from the change-of-variable formula and shows that $f*g$ is symmetric in f and g). Observe that scalar-functions can always be identified with operator-valued ones, with the identification of $\lambda \in \mathbb{C}$ and $\lambda \,\mathrm{id} \in \mathcal{B}(X)$.

Minkowski's integral inequality shows that $|f * g|_{L^p} \leq |f|_{L^1} |g|_{L^p}$ for $p \in [1, \infty)$:

$$\left(\int \left|\int g(t-s)f(s)ds\right|^p dt\right)^{\frac{1}{p}} \leq \int \left(\int |g(t-s)|^p \, dt\right)^{\frac{1}{p}} |f(s)| \, ds = |g|_{L^p} \, |f|_{L^1} \, ,$$

where the absolute values can be replaced by appropriate norms. (The corresponding inequality for $p = \infty$ is even easier.)

With the aid of the convolution, the Fourier transform can also be written $\widehat{f}(\xi) = f * e_\xi(0)$, where we denote by e_t, $t \in \mathbb{R}^d$, the function $e_t(\xi) := e^{\mathbf{i}2\pi t\cdot\xi}$. For the sequel, also recall that $\tau_h f := f(\cdot - h)$.

The following properties of the Fourier transform follow instantly for $f \in L^1(\mathbb{R}^d; X)$, $g \in L^1(\mathbb{R}^d)$, $\lambda > 0$, (using the vector-valued Fubini's theorem and the change-of-variable formula where appropriate):

$$\begin{aligned}
\widehat{\tau_h f}(\xi) &= \int f(t-h)e^{-\mathbf{i}2\pi\xi\cdot t}dt = \int f(t)e^{-\mathbf{i}2\pi\xi\cdot(t+h)}dt = e^{-\mathbf{i}2\pi\xi\cdot h}\widehat{f}(\xi) = (e_{-h}\widehat{f})(\xi) \\
\widehat{e_h f}(\xi) &= \int e_h f e_{-\xi} dt = \int f e_{-(\xi-h)} dt = \widehat{f}(\xi - h) = (\tau_h \widehat{f})(\xi) \\
\widehat{f * g}(\xi) &= \int dt \int ds f(t-s)g(s)e^{-\mathbf{i}2\pi\{(t-s)+s\}\cdot\xi} \\
&= \int ds \int du f(u)e^{-\mathbf{i}2\pi u\cdot\xi} g(s)e^{-\mathbf{i}2\pi s\cdot\xi} = (\widehat{f}\widehat{g})(\xi) \\
\widehat{f(\lambda\cdot)}(\xi) &= \int f(\lambda t)e^{-\mathbf{i}2\pi\xi\cdot t}dt = \int f(u)e^{-\mathbf{i}2\pi\xi\cdot\frac{1}{\lambda}u}\frac{du}{\lambda^d} = \lambda^{-n}\widehat{f}(\frac{1}{\lambda}\xi)
\end{aligned}$$

The following result is also immediate from the definition; it will play the central role in extending $\mathcal{F}$ beyond L^1.

Lemma A.26. *For $f \in L^1(\mathbb{R}^d; \mathcal{B}(X;Y))$, $g \in L^1(\mathbb{R}^d; X)$, we have*

$$\int_{\mathbb{R}^d} \widehat{f} g dm = \int_{\mathbb{R}^d} f \widehat{g} dm$$

Proof. Writing out the definition of the Fourier transform and using Fubini's theorem, the left-hand side is equal to

$$\int_{\mathbb{R}^d} d\xi \int_{\mathbb{R}^d} dt f(t) e^{-\mathbf{i}2\pi t\cdot\xi} g(\xi),$$

and this is clearly symmetric in f and g, thus also equal to the right-hand side. ☐

As in the Fourier analysis of real-valued functions, the **Schwartz space** $\mathcal{S}(\mathbb{R}^d; X)$ of **rapidly decreasing functions** turns out to be one of the spaces on which the Fourier transform works most "naturally". This space is defined, as in the case $X = \mathbb{C}$, as the class of those $\psi \in C^\infty(\mathbb{R}^d; X)$ for which each of the countable number of seminorms

$$|\psi|_{\alpha,\beta} := \sup_{t\in\mathbb{R}^d} \left|t^\beta D^\alpha \psi(t)\right|_X ,$$

$\alpha, \beta \in \mathbb{N}$ is finite.

(Recall the usual multi-index notation used here: $t^\beta := t_1^{\beta_1} \cdots t_d^{\beta_d}$, $D^\alpha := D_1^{\alpha_1} \cdot D_d^{\alpha_d}$.) We also refer to the elements of $\mathcal{S}$ as **test functions**. (Sometimes this name is reserved only for the smaller test function space $\mathcal{D}$ of all C^∞ functions with compact support, but we find it convenient to apply this name also to the functions of $\mathcal{S}$.)

$\mathcal{S}$ becomes a metric space when endowed with the (translation invariant) metric

$$\varrho(\psi, \phi) := \sum_{\alpha,\beta\in\mathbb{N}^d} 2^{-|\alpha|_1-|\beta|_1} \frac{|\psi-\phi|_{\alpha\beta}}{1+|\psi-\phi|_{\alpha\beta}}.$$

It is easy to see that $\varrho(\psi_k, \psi) \to 0$ if and only if $|\psi_k - \psi|_{\alpha,\beta} \to 0$ for all $\alpha, \beta \in \mathbb{N}^d$. This will be called convergence in $\mathcal{S}(\mathbb{R}^d; X)$; it is convergence in the topology induced by the metric ϱ. All topological notions related to $\mathcal{S}$ will always refer to this topology.

It is sometimes useful to observe that the same topology of $\mathcal{S}$ is generated by the metric ϱ' defined like ϱ but using the seminorms $\|\psi\|_{\alpha,r} := \sup_{t\in\mathbb{R}^d}(1 + |t|^2)^{\frac{r}{2}} |D^\alpha\psi(t)|_X$, where $\alpha \in \mathbb{N}^d$, $r \in \mathbb{N}$. Indeed, the following is true:

Lemma A.27. *For $\psi \in \mathcal{S}(\mathbb{R}^d; X)$, $\alpha, \beta \in \mathbb{N}^d$, $|\beta|_1 \le r$, we have*

$$|\psi|_{\alpha,\beta} \le \|\psi\|_{\alpha,r} \le A(n, r) \sum_{k=0}^{d} |\psi|_{\alpha,\theta_k(r)} .$$

The topological equivalence of ϱ and ϱ' follows from this: $\varrho(\phi_k, \phi) \to 0$ if and only if, for all $\alpha, \beta \in \mathbb{N}^d$, $|\phi_k - \phi|_{\alpha,\beta} \to 0$ if and only if, for all $\alpha \in \mathbb{N}^d$, $r \in \mathbb{N}$, $\|\phi_k - \phi\|_{\alpha,\beta} \to 0$ if and only if $\varrho'(\phi_k, \phi) \to 0$; the second "if and only if" follows from the assertion of the lemma.

Proof. We first observe that

$$(1 + |t|^2)^r = \left(1 + \sum_{j=1}^{d} t_j^2\right)^r = \sum_{i_0+\ldots+i_d=r} \binom{r}{i_0, \ldots, i_d} t_1^{2i_1} \cdots t_d^{i_d}$$

contains each $t^{2\beta}$, $\beta \le r$, with a coefficient not less than one; thus $(1 + |t|^2)^r \ge t^{2\beta}$, and after taking square roots it follows that $\|\psi\|_{\alpha,r} \ge |\psi|_{\alpha,\beta}$ for any $\beta \le r$. To derive an inequality of converse type, we recall first the following simple result:

For $0 < q < p < \infty$, $\xi \in \mathbb{R}^m$, we have

$$|\xi|_p := \left(\sum_{k=1}^{m} |\xi_k|^p\right)^{\frac{1}{p}} \le \left(\sum_{k=1}^{m} |\xi_k|^q\right)^{\frac{1}{q}} =: |\xi|_q . \tag{A.11}$$

This is easy to derive, since the homogeneity of the inequality allows us to assume that the right-hand side is equal to unity (after introducing $\eta := \xi/|\xi|_q$, if necessary). When this assumption holds, then $|\xi_k| \in [0, 1]$ for each k, and therefore $|\xi_k|^p \le |\xi_k|^q$ for p and q as in the assertion, whence $\sum_{k=1}^m |\xi_k|^p \le \sum_{k=1}^m |\xi_k|^q = 1$, and we can take the pth root to deduce the claim.

Applying (A.11) to $p = 2$, $q = 1$, we find that $\left(1 + \sum_{j=1}^d t_j^2\right)^{\frac{1}{2}} \le 1 + \sum_{j=1}^d |t_j|$, and combining this with Jensen's inequality $\left(\frac{1}{n+1} \sum_{j=0}^d a_j\right)^r \le \frac{1}{n+1} \sum_{j=0}^d a_k^r$, with $a_j := |t_j|$ for $j \ge 1$, $a_0 = 1$ and $r \ge 1$, we can estimate

$$\left(1 + \sum_{j=1}^{d} t_j^2\right)^{\frac{r}{2}} \le \left(1 + \sum_{j=1}^{d} |t_j|\right)^r \le (n + 1)^{r-1} \left(1 + \sum_{j=1}^{d} |t_j|^r\right) .$$

It then follows that

$$\left|(1+|t|^2)^{\frac{r}{2}} D^\alpha \psi\right|_{L^\infty} \leq (n+1)^{r-1} \left(|D^\alpha \psi|_{L^\infty} + \sum_{j=1}^{d} \left|t_j^r D^\alpha \psi\right|_{L^\infty} \right)$$

$$= (n+1)^{r-1} \left(|\psi|_{\alpha,0} + \sum_{j=1}^{d} |\psi|_{\alpha,re_j} \right).$$

The assertion is now proved. □

It is clear that $|\psi|_{0,0} = |\psi|_{L^\infty}$. From the previous lemma it also follows that

Lemma A.28. *For $\psi \in \mathcal{S}(\mathbb{R}^d; X)$, $|\psi|_{L^p(\mathbb{R}^d;X)} \leq A \sum_{k=1}^{d} |\psi|_{0,\beta_k}$, where $A = A(d,p)$ and $\beta_k = \beta_k(d)$.*

Proof. The pth power of the L^p norm can be written as

$$\int_{\mathbb{R}^d} |\psi(t)|_X^p \, dt \leq \left(\sup_{t \in \mathbb{R}^d} (1+|t|^2)^{\frac{1}{2}(n+1)} |\psi(t)|_X \right)^p \int_{\mathbb{R}^d} \frac{dt}{(1+|t|^2)^{\frac{1}{2}(n+1)p}} = c(n,p) \, \|\psi\|_{0,n+1}^p ,$$

where $c(n,p) := \int_{\mathbb{R}^d} (1+|t|^2)^{-\frac{1}{2}(n+1)p} dt < \infty$. The asserted norm estimate now follows from Lemma A.27. □

The convenience of the Schwartz space lies in two main attributes:

1. It behaves well with various common operations of analysis.
2. It is dense in many spaces of interest.

The density of $\mathcal{S}(\mathbb{R}^d; X)$ in $L^p(\mathbb{R}^d; X)$, $1 \leq p < \infty$, is immediate from the fact that $\mathcal{D}(\mathbb{R}^d; X) \subset \mathcal{S}(\mathbb{R}^d; X)$ is dense in these spaces. In the following, our purpose is to justify the former statement, once it is given more precise content, of course.

Several results are conveniently stated for **slowly increasing** functions, i.e., C^∞ mappings, all of whose derivatives are bounded by a polynomial. In particular, each test function and each polynomial is slowly increasing.

Lemma A.29. *Let Φ be slowly increasing and $\gamma \in \mathbb{N}^d$. Then the mappings $\psi \mapsto \Phi\psi$ and $\psi \mapsto D^\gamma \psi$ are continuous from $\mathcal{S}$ to $\mathcal{S}$.*

It is implicitly assumed that at least one of Φ and ψ is operator-valued in the first mapping. Although it is conventional to write operators to the left, we take the freedom not to always obey this rule in order to simplify the statement of some result. It is always understood that in any product, all functions, except possibly one, are operator-valued. Also recall the identification of scalar-functions with operator-valued ones.

Proof. First, obviously the product of two C^∞ functions is C^∞, as is any derivative of a C^∞ function. The computation

$$|\Phi\psi|_{\alpha,\beta} = \left|t^\beta D^\alpha (\Phi\psi)\right|_{L^\infty} \leq \sum_{\theta \leq \alpha} \left|D^{\alpha-\theta} \Phi\right|_{L^\infty} \left|t^\beta D^\theta \psi\right|_{L^\infty} = \sum_{\theta \leq \alpha} \left|D^{\alpha-\theta} \Phi\right|_{L^\infty} |\psi|_{\theta,\beta}$$

shows that $\Phi\psi \in \mathcal{S}$, and the same equation applied to $\psi_k - \psi$ in place of ψ shows that $\Phi\psi_k \to \Phi\psi$ in $\mathcal{S}$ whenever $\psi_k \to \psi$ in $\mathcal{S}$. The other assertion of the lemma follows similarly from the obvious identity $|D^\gamma \psi|_{\alpha,\beta} = |\psi|_{\alpha+\gamma,\beta}$. □

The above assertion could equally well have been established for the function class $\mathcal{D}$. The real benefit of $\mathcal{S}$ compared to $\mathcal{D}$ is its behaviour with the Fourier transform.

Lemma A.30. *The Fourier transform $\mathcal{F}$ is continuous from $\mathcal{S}$ to $\mathcal{S}$. Furthermore, if P is a polynomial, then*

$$\mathcal{F}(P(D)\psi)(\xi) = P(\mathbf{i}2\pi\xi)\widehat{\psi}(\xi) \qquad \text{and} \qquad P(D)\widehat{\psi}(\xi) = \mathcal{F}(P(-\mathbf{i}2\pi\cdot)f)(\xi).$$

Proof. Assume for the moment that the two formulae in the assertion have been verified. It then follows that

$$\left|\widehat{\psi}\right|_{\alpha,\beta} = \left|\xi^\beta D^\alpha \widehat{\psi}\right|_{L^\infty} = \left|\mathcal{F}((\mathbf{i}2\pi)^{-|\beta|_1} D^\beta(-\mathbf{i}2\pi)^{|\alpha|_1} t^\alpha \psi)\right|_{L^\infty}$$
$$\le (2\pi)^{|\alpha|_1 - |\beta|_1} \left|D^\beta t^\alpha \psi\right|_{L^1} \le C \sum_{k=1}^N |\psi|_{0,\beta_k},$$

where the first inequality is (A.9), and the second follows from Lemmas A.29 and A.28. Once we verify the two asserted formulae, the continuity of $\mathcal{F}$ from $\mathcal{S}$ to $\mathcal{S}$ follows.

For $\psi \in \mathcal{S}$ and $\gamma \in \mathbb{N}^d$, $\xi^\gamma \psi$ and $D^\gamma \psi$ are in $\mathcal{S}$ (by Lemma A.29) and thus integrable (by Lemma A.28). Since $D^\gamma e_\xi = (\mathbf{i}2\pi\xi)^{|\gamma|_1} e_\xi$, it follows easily from the dominated convergence theorem applied to appropriate difference quotients that the convolution $\psi * e_\xi$ can be differentiated by differentiating under the integral in A.10 either of the functions ψ or $g = e_\xi$. Thus (observe that ξ is regarded as a parameter and not a variable in the following computation)

$$(P(D)\psi) * e_\xi = \psi * P(D)e_\xi = \psi * P(\mathbf{i}2\pi\xi)e_\xi = P(\mathbf{i}2\pi\xi)(\psi * e_\xi);$$

in particular $\widehat{P(D)\psi}(\xi) = (P(D)\psi) * e_\xi(0) = P(\mathbf{i}2\pi\xi)(\psi * e_\xi)(0) = P(\mathbf{i}2\pi\xi)\widehat{\psi}(\xi)$.

The differentiation with respect to ξ in the definition (A.8) of the Fourier transform can be brought under the integral by exactly the same domination estimates as above in this proof, and this immediately yields the second formula in the assertion. □

In establishing the inversion formula, the following fixed point of $\mathcal{F}$ can be exploited:

Lemma A.31. $\mathcal{F}\left(e^{-\pi|t|^2}\right) = e^{-\pi|\xi|^2}$.

Proof. Since $e^{-\pi|t|^2}$ is a product of n functions of one real variable of the similar form, it suffices to consider the one-dimensional case. Then

$$\int_{-\infty}^{\infty} e^{-\pi t^2} e^{-\mathbf{i}2\pi t\xi} dt = \int_{-\infty}^{\infty} e^{-\pi(t+\mathbf{i}\xi)^2} e^{-\pi\xi^2} dt = e^{-\pi\xi^2} \int_{-\infty+\mathbf{i}\xi}^{\infty+\mathbf{i}\xi} e^{-\pi z^2} dz = e^{-\pi\xi^2} \int_{-\infty}^{\infty} e^{-\pi z^2} dz,$$

where Cauchy's theorem (e.g. Rudin [20], with some easy estimates related to the smallness of the integrand at infinity) was used in the last step; the familiar integral on the right has the value 1, and the proof is complete. □

Lemma A.32. *The Fourier transform is a bijection of $\mathcal{S}$ onto $\mathcal{S}$, with continuous inverse $\mathcal{F}^{-1} = \mathcal{F}^* = \mathcal{F}^3$. Furthermore, $\mathcal{F}^2 = \mathcal{R}$ (the reflection operator).*

Proof. Applying Lemma A.26 and the basic properties of the Fourier transform to the functions $\phi(\epsilon\cdot)$, $\phi \in \mathcal{S}(\mathbb{R}^d; X)$ and $\psi \in \mathcal{S}(\mathbb{R}^d)$, we deduce

$$\int \phi(\epsilon\cdot)\widehat{\psi} dm = \int \phi \epsilon^{-n} \widehat{\psi}(\epsilon^{-1}\cdot) dm = \int \phi \widehat{\psi(\epsilon\cdot)} dm = \int \widehat{\phi} \psi(\epsilon\cdot) dm.$$

As $\epsilon \to \infty$, $\phi(\epsilon\cdot) \to \phi(0)$ and $\psi(\epsilon\cdot) \to \psi(0)$ at each point. Since ϕ and ψ are bounded and $\widehat{\psi}$ and $\widehat{\phi}$ integrable, it follows from the dominated convergence theorem that $\phi(0)\int \widehat{\psi} dm = \psi(0) \int \widehat{\phi} dm$. Taking $\psi(t) := \widehat{\psi}(t) := e^{-\pi|t|^2}$ (Lemma A.31), it follows that $\phi(0) = \int \widehat{\phi} dm$. Then

$$\phi(t) = (\tau_{-t}\phi)(0) = \int \widehat{\tau_{-t}\phi} dm = \int e^{\mathbf{i}2\pi t\cdot\xi} \widehat{\phi}(\xi) d\xi = \mathcal{F}^*(\widehat{\phi})(t),$$

i.e., $\phi = \mathcal{F}^*\mathcal{F}\phi$.

The continuity of $\mathcal{F}$ was already proved, and the continuity of $\mathcal{F}^*$ follows from the fact that $\mathcal{F}^* = \mathcal{R}\mathcal{F}$ and the obvious continuity of $\mathcal{R}$.

The injectivity of $\mathcal{F}$ is also obvious now, for if $\mathcal{F}\phi = \mathcal{F}\psi$, then $\phi = \mathcal{F}^*\mathcal{F}\phi = \mathcal{F}^*\mathcal{F}\psi = \psi$. Using the formula $\mathcal{F}^* = \mathcal{R}\mathcal{F}$, we further have $\phi = \mathcal{R}\mathcal{F}^2\phi$, and applying once more $\mathcal{R}$ to both sides (recalling that $\mathcal{R}^2$ is the identity) we obtain $\mathcal{R}\phi = \mathcal{F}^2\phi$. Thus $\mathcal{F}^4 = \mathcal{R}^2 = \mathrm{id}$ and $\mathcal{F}^* = \mathcal{R}\mathcal{F} = \mathcal{F}^3 = \mathcal{F}^{-1}$.

The proof is complete. □

We obtain immediate corollaries involving convolutions.

Corollary A.33. *For $\psi \in \mathcal{S}$, $\phi \to \psi * \phi$ is continuous from $\mathcal{S}$ to $\mathcal{S}$. Furthermore, $\mathcal{F}(\psi\phi) = \widehat{\psi} * \widehat{\phi}$.*

Proof. From the elementary properties of the Fourier transform on L^1 we recall that $\mathcal{F}(\phi*\psi) = \widehat{\phi}\widehat{\psi}$. Applying $\mathcal{F}^*$ to both sides gives $\phi*\psi = \mathcal{F}^*(\widehat{\phi}\widehat{\psi})$, which shows that convolution is continuous, since $\mathcal{F}$, multiplication by an $\mathcal{S}$ function, and $\mathcal{F}^*$ are continuous.

The asserted formula is obtained by substituting $\widehat{\psi}$ and $\widehat{\phi}$ in place of ψ and ϕ in the L^1 convolution formula to yield:

$$\widehat{\psi} * \widehat{\phi} = \mathcal{F}^*(\mathcal{F}^2\phi \cdot \mathcal{F}^2\psi) = \mathcal{F}^*(\widetilde{\phi}\widetilde{\psi}) = \mathcal{F}^*\mathcal{R}(\phi\psi) = \mathcal{F}(\phi\psi).$$

□

Next we will establish the completeness of $\mathcal{S}$. A simple preliminary results dealing with the type of convergence involved is in order.

Lemma A.34. *Let $\{f_k\}_{k=1}^{\infty}$ be a sequence of continuous functions from $\mathbb{R}^d$ to X such that $D_i f_k$ is continuous for each k. Let $f_k \to f$ at each point of $\mathbb{R}^d$ and $D_i f_k \to g$ uniformly on $\mathbb{R}^d$. Then f is continuously differentiable with respect to the ith coordinate, and $D_i f = g$.*

Proof. The uniform limit of continuous functions is continuous; the usual argument applies as such to vector-valued mappings. Thus it suffices to show that $D_i f = g$ at each point.

By Lemma A.23, $f_k(t + he_i) - f_k(t) = \int_{t_i}^{t_i+h} D(f_k)_i(s;t)ds$ for each $k \in \mathbb{Z}_+$, $t \in \mathbb{R}^d$ and $h \in \mathbb{R}$. From the assumed convergence it follows that the left-hand side tends to $f(t+he_i) - f(t)$ as $k \to \infty$, and the uniform convergence of $D_i f_k$ to g ensures that the right-hand side tends to $\int_{t_i}^{t_i+h} g_i(s;t)ds$. But it now follows from the very definition of the partial derivative, together with Lemma A.21, that $D_i f(t) = g_i(t_i;t) = g(t)$. This proves the claim. □

Lemma A.35. *$\mathcal{S}(\mathbb{R}^d;X)$ is complete.*

Proof. Let $\{\phi_k\}_{k=1}^{\infty}$ be a Cauchy sequence on $\mathcal{S}(\mathbb{R}^d;X)$. In terms of the metric ϱ' of $\mathcal{S}$ (which is topologically equivalent to ϱ), this means that $\varrho'(\phi_k,\phi_j) \to 0$ as $k,j \to \infty$, i.e., $\|\phi_k - \phi_j\|_{\alpha,r} \to 0$ for all $\alpha \in \mathbb{N}^d$, $r \in \mathbb{N}$. Denoting $p_r(t) := (1+|t|^2)^{\frac{r}{2}}$, the assumed convergence means that $\{p_r D^\alpha \phi_k\}_{k=1}^{\infty}$ is Cauchy (in the supremum norm), and thus converges to a continuous function $g_{\alpha,r}$ uniformly on $\mathbb{R}^d$. In particular, $D^\alpha\phi_k \to g_{\alpha,0}$ uniformly, and iterative application of Lemma A.34 shows that $\phi := g_{0,0}$ is infinitely differentiable with $D^\alpha\phi = g_{\alpha,0}$. Thus we know that $p_r D^\alpha \phi_k \to p_r D^\alpha \phi$ uniformly for each α and r, and it follows that $\phi \in \mathcal{S}$ (since uniform convergence implies convergence of the supremum norms), and ϕ is the $\mathcal{S}$-limit of the sequence $\{\phi_k\}_{k=1}^{\infty}$. Thus an arbitrary Cauchy sequence converges in $\mathcal{S}$, and the proof is complete. □

We already know that $\mathcal{D} \subset \mathcal{S} \subset L^p$, $1 \le p < \infty$, and the two first spaces are dense in the third. It is also true that $\mathcal{D}$ is dense in $\mathcal{S}$; this is not immediate from the previous inclusions, since we require density in the sense of $\mathcal{S}$, where convergence is much stronger than L^p convergence. Nevertheless, this can be obtained quite readily with the aid of the following lemma.

Lemma A.36. *If $\psi \in \mathcal{S}(\mathbb{R}^d)$ and $\phi \in \mathcal{S}(\mathbb{R}^d;X)$, then $\psi(\epsilon\cdot)\phi \to \psi(0)\phi$ in $\mathcal{S}$ as $\epsilon \to 0$.*

Proof. Denote $a := \psi(0)$. The Leibniz rule yields the estimate:

$$\left|t^\beta D^\alpha(\psi(\epsilon t)\phi(t) - a\phi(t))\right|_X$$
$$\leq \sum_{0\neq\theta\leq\alpha} \binom{\alpha}{\theta} \left|t^\beta D^{\alpha-\theta}\phi(t)\right|_X \left|(D^\theta\psi)(\epsilon t)\right| \epsilon^{|\theta|_1} + \left|t^\beta D^\alpha \phi(t)\right|_X |\psi(\epsilon t) - a|\,.$$

In the summation, $\left|t^\beta D^{\alpha-\theta}\phi(t)\right|_X$ and $\left|(D^\theta\psi)(\epsilon t)\right|$ are bounded uniformly in t, since $\psi, \phi \in \mathcal{S}$; thus the supremum norm of the first term on the right vanishes as $\epsilon \to 0$.

To estimate the last term, we first note $|t|^2 t^\beta$ is slowly increasing, and thus $|t|^2 t^\beta D^\alpha\phi$ is in $\mathcal{S}$ by Lemma A.29. In particular, it is uniformly bounded, and thus $\left|t^\beta D^\alpha \phi(t)\right|_X \leq C|t|^{-2}$ for some C. It is then clear (since $|\psi(\epsilon t) - 1| \leq |\psi|_{L^\infty(\mathbb{R}^d)} + 1 < \infty$, that the last term above is less than any given $\eta > 0$ for all $|t|$ not less than some R. (This is true independently of the value of the parameter ϵ.) On the other hand, the continuity of ψ ensures that in a suitable r-neighbourhood of 0, $|\psi - a| \leq |\phi|_{\alpha,\beta}^{-1}\eta$. Therefore, the last term is also less than η in $t \leq R$, given that ϵ is sufficiently small so that $\epsilon R < r$. We have thus shown that even the supremum norm of the last term above can be made as small as desired by choosing sufficiently small ϵ.

The asserted convergence in each of the seminorms $|\cdot|_{\alpha,\beta}$ is now established. □

Corollary A.37. *$\mathcal{D}(\mathbb{R}^d; X)$ is dense in $\mathcal{S}(\mathbb{R}^d; X)$.*

Proof. For $\psi \in \mathcal{D}(\mathbb{R}^d)$ and $\phi \in \mathcal{S}(\mathbb{R}^d; X)$, clearly $\psi(\epsilon\cdot)\phi \in \mathcal{D}(\mathbb{R}^d; X)$ for $\epsilon > 0$. Choosing ψ so that $\psi(0) = 1$ (which is certainly possible by Lemma A.18), it follows from Lemma A.36 that the functions $\psi(\epsilon\cdot)\phi \in \mathcal{D}(\mathbb{R}^d; X)$ converge in $\mathcal{S}$ to ϕ. □

Lemma A.36, the essential ingredient of the previous proof showed that the functions $\psi(\epsilon\cdot)$, $\psi \in \mathcal{D}(\mathbb{R}^d)$ provide an approximation of the identity in the topology of $\mathcal{S}$ as $\epsilon \to 0$. Another result of similar kind will be of use, too:

Lemma A.38. *If $\psi \in C_c(\mathbb{R}^d)$, $\operatorname{ran}\psi \subset [0,\infty)$, $\int_{\mathbb{R}^d}\psi dm = 1$, and $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d; X)$, then for every $t \in \mathfrak{L}f$*

$$\lim_{r\to 0}\int_{\mathbb{R}^d} f(s)\epsilon^{-n}\psi(\epsilon^{-1}(s-t))ds = f(t).$$

In particular, this holds for $\psi \in \mathcal{D}(\mathbb{R}^d)$.

Proof. We compute

$$\int_{\mathbb{R}^d} |f(s) - f(t)|_X\, \psi(\epsilon^{-1}(s-t))\frac{ds}{\epsilon^d} = \frac{m(\operatorname{supp}\psi)}{\epsilon^d m(\operatorname{supp}\psi)}\int_{\epsilon\operatorname{supp}\psi + t} |f(s) - f(t)|_X\, \psi(\epsilon^{-1}(s-t))ds\cdot$$
$$\leq \frac{1}{m(\epsilon\operatorname{supp}\psi + t)}\int_{\epsilon\operatorname{supp}\psi+t} |f(s) - f(t)|_X\, ds\cdot m(\operatorname{supp}\psi)\sup_{s\in\mathbb{R}^d}\psi(s)$$

and the integral average here tends to zero as $\epsilon \to 0$: it is easy to see that the sets $\epsilon\operatorname{supp}\psi + t$ shrink nicely to t. (We used the properties of the Lebesgue measure via the identities $\epsilon^d m(\operatorname{supp}\psi) = m(\epsilon\operatorname{supp}\psi) = m(\epsilon\operatorname{supp}\psi + t)$.) □

Yet another density result, useful in Fourier analysis, is the following:

Lemma A.39. *The functions $\varphi \in \mathcal{S}(\mathbb{R}^d; X)$ satisfying $\mathcal{F}\varphi \in \mathcal{D}(\mathbb{R}^d; X)$ and $0 \notin \operatorname{supp}\mathcal{F}\varphi$, are dense in $L^p(\mathbb{R}^d; X)$ for $p \in (1,\infty)$. The functions in $\mathcal{F}^{-1}\mathcal{D}(\mathbb{R}^d; X)$ without the support condition are dense even in $\mathcal{S}(\mathbb{R}^d; X)$, thus in $L^p(\mathbb{R}^d; X)$, $p \in [1,\infty)$.*

Proof. Since $\mathcal{D}$ is dense in $\mathcal{S}$ and $\mathcal{F}^{-1}$ is continuous from $\mathcal{S}$ to $\mathcal{S}$, it is clear that $\mathcal{F}^{-1}\mathcal{D} \subset \mathcal{S}$ is dense in $\mathcal{S}$, thus in L^p, $p < \infty$. Thus it suffices to show that, given $\varphi \in \mathcal{F}^{-1}\mathcal{D}$, there exist $\varphi_\epsilon \in \mathcal{F}^{-1}\mathcal{D}$ such that $\widehat{\varphi}_\epsilon(0) = \widehat{\varphi}(0)$, and $|\varphi_\epsilon|_{L^p(\mathbb{R}^d;X)} \to 0$ as $\epsilon \to 0$.

This is readily achieved, since taking $\widehat{\phi} \in \mathcal{D}$ such that $\widehat{\phi} = 1$ in a neighbourhood of zero, it is clear that $\widehat{\phi}(\epsilon^{-1}\cdot)\widehat{\varphi} = \mathcal{F}(\epsilon^d \phi(\epsilon\cdot) * \varphi)$ is in $\mathcal{D}$ and coincides with $\widehat{\varphi}$ in a neighbourhood of zero; thus the difference of these two is a function with required properties. Furthermore,

$$\left|\epsilon^d \phi(\epsilon\cdot) * \varphi\right|_{L^p(\mathbb{R}^d;X)} \leq \left|\epsilon^d \phi(\epsilon\cdot)\right|_{L^p(\mathbb{R}^d)} |\varphi|_{L^1(\mathbb{R}^d;X)}$$

and

$$\left|\epsilon^d \phi(\epsilon\cdot)\right|^p_{L^p(\mathbb{R}^d)} = \int_{\mathbb{R}^d} \left|\epsilon^d \phi(\epsilon t)\right|^p dt = \int_{\mathbb{R}^d} \epsilon^{np} \left|\varphi(t)\right| X^p \frac{dt}{\epsilon^d} = \epsilon^{d(p-1)} |\varphi|_{L^p(\mathbb{R}^d)},$$

and this clearly tends to zero as $\epsilon \to 0$. □

A.7 Distributions

So far we have analyzed the Fourier transform on a rather restricted set of functions. The extension of this concept by duality to a significantly larger domain is our next goal. Other applications of the related notion also appear.

In the theory of scalar-valued distributions, the duality extension is somewhat more obvious, since one simply considers the usual dual space $\mathcal{S}^*(\mathbb{R}^d)$ of $\mathcal{S}(\mathbb{R}^d)$, i.e., the continuous linear transformations from $\mathcal{S}(\mathbb{R}^d)$ to $\mathbb{C}$. As is well known, this class includes all L^p functions, $1 \leq p \leq \infty$, (with the suitable identification $\langle f, \phi\rangle = \int_{\mathbb{R}^d} f\phi dm$, $f \in L^p(\mathbb{R}^d)$, $\phi \in \mathcal{S}(\mathbb{R}^d)$). Two modifications are needed to build a class of vector-valued distributions for which similar identifications can be made:

1. The space of test functions on which we make pairings $\langle u, \phi\rangle$, u a distribution, ϕ a test function, must consist of scalar-valued functions in order to be able to identify vector-valued functions with distributions in the usual way.

2. We cannot stick to a dual space in the usual sense, i.e., mappings to the complex plane $\mathbb{C}$, since vector-valued functions will yield vector outcomes when identified with distributions.

The following definition satisfying the requirements above has proved to be useful:

Definition A.40. *For a Banach space X, the space of vector-valued* ***tempered distributions*** *is defined by $\mathcal{S}^*(\mathbb{R}^d; X) := \mathcal{B}(\mathcal{S}(\mathbb{R}^d); X)$.*

Observe that this does coincide with the usual dual space definition when $X = \mathbb{C}$.

Continuity of $u \in \mathcal{S}^*(\mathbb{R}^d; X)$ on $\mathcal{S}(\mathbb{R}^d)$ naturally means that $\langle u, \psi_n\rangle \to \langle u, \psi\rangle$ (in the norm of X) whenever $|\psi_n - \psi|_{\alpha\beta} \to 0$ for all $\alpha, \beta \in \mathbb{N}^d$, i.e., whenever $\varrho(\psi_n, \psi) \to 0$. This continuity can also be characterized as follows:

Lemma A.41. *A linear transformation $u : \mathcal{S}(\mathbb{R}^d) \to X$ is a tempered distribution if and only if for some finite C and N,*

$$|\langle u, \psi\rangle|_X \leq C \sum_{|\alpha|_1, |\beta|_1 \leq N} |\psi|_{\alpha,\beta}, \tag{A.12}$$

for all $\psi \in \mathcal{S}(\mathbb{R}^d)$.

Proof. If u satisfies (A.12), it is obvious that the convergence $\phi_n - \phi \to 0$ in $\mathcal{S}(\mathbb{R}^d)$ implies the convergence $|\langle u, \phi_n\rangle - \langle u, \phi\rangle|_X \to 0$.

To show the converse, suppose $u \in \mathcal{S}^*(\mathbb{R}^d; X)$. Since u is continuous, there is a $\delta > 0$ such that $|\langle u, \phi\rangle|_X < 1$ whenever $\varrho(\phi, 0) < \delta$. Having chosen such a δ, pick an N large enough so that $\sum_{|\alpha|_1, |\beta|_1 \geq N} 2^{-|\alpha|_1 - |\beta|_1} < \frac{1}{2}\delta$. (This is possible, since the series is absolutely convergent.) But then $\varrho_N(\phi, 0) := \sum_{|\alpha|_1, |\beta|_1 \leq N} 2^{-|\alpha|_1 - |\beta|_1} \frac{|\phi|_{\alpha,\beta}}{1 + |\phi|_{\alpha,\beta}} \leq \frac{1}{2}\delta$ implies $|\langle u, \phi\rangle|_X \leq 1$. Since $\varrho_N(\phi) \leq \sum_{|\alpha|_1, |\beta|_1 \leq N} |\phi|_{\alpha,\beta}$, it follows that $|\langle u, \phi\rangle|_X \leq 1$ whenever $\sum_{|\alpha|_1, |\beta|_1 \leq N} |\phi|_{\alpha,\beta} \leq \frac{1}{2}\delta$.

But for any (non-zero) $\psi \in \mathcal{S}(\mathbb{R}^d)$, the previous inequality is certainly satisfied by $\phi := \frac{1}{2}\delta\psi\left(\sum_{|\alpha|_1,|\beta|_1\le N}|\psi|_{\alpha,\beta}\right)^{-1}$; therefore

$$1 \ge |\langle u,\phi\rangle|_X = |\langle u,\psi\rangle|_X \frac{\delta}{2}\left(\sum_{|\alpha|_1,|\beta|_1\le N}|\psi|_{\alpha,\beta}\right)^{-1}$$

for all $\psi \in \mathcal{S}(\mathbb{R}^d)$, and the claim follows with $C = \frac{2}{\delta}$. $\square$

Example A.42. *Every measurable $g : \mathbb{R}^d \to X$, for which $(1+|t|^2)^{-N}g \in L^p$ for some N, is in $\mathcal{S}^*(\mathbb{R}^d;X)$ with the identification $\langle g,\phi\rangle := \int_{\mathbb{R}^d} g\phi dm$.*

In particular, every slowly increasing function is a tempered distribution.

Proof. We have $\langle g,\phi\rangle = \left\langle (1+|t|^2)^{-N}g, (1+|t|^2)^N\phi\right\rangle$. Since $\phi \mapsto (1+|t|^2)^N\phi$ is continuous from $\mathcal{S}(\mathbb{R}^d)$ to $\mathcal{S}(\mathbb{R}^d)$, it suffices to prove that $L^p(\mathbb{R}^d;X) \subset \mathcal{S}^*(\mathbb{R}^d;X)$. By Hölder's inequality, $|\langle g,\phi\rangle|_X \le |g|_{L^p(\mathbb{R}^d;X)}|\phi|_{L^{\overline{p}}(\mathbb{R}^d;X)}$, and since $|\phi|_{L^{\overline{p}}(\mathbb{R}^d;X)}$ is bounded by a finite number of the seminorms $|\cdot|_{\alpha,\beta}$ of $\mathcal{S}$ by Lemma A.28, the assertion follows. $\square$

Whenever there is, for a certain $u \in \mathcal{S}^*(\mathbb{R}^d;X)$, a locally integrable function $f : \mathbb{R}^d \to X$ such that $\langle u,\phi\rangle = \langle f,\phi\rangle$, we identify u with the function f. We sometimes emphasize this by saying that u is a proper function (as opposed to a generalized function, another name occasionally used for distributions).

The standard operations of analysis are defined for $u \in \mathcal{S}^*$ "by duality". Thus we have, for each $u \in \mathcal{S}^*$, $\phi,\psi \in \mathcal{S}$

1. the Fourier transform $\langle \widehat{u},\psi\rangle := \left\langle u,\widehat{\psi}\right\rangle$,
2. multiplication by a slowly increasing function $\langle u\phi,\psi\rangle := \langle u,\phi\psi\rangle$,
3. convolution by a test function $\langle u*\phi,\psi\rangle := \left\langle u,\widetilde{\phi}*\psi\right\rangle$,
4. differentiation $\langle D^\alpha u,\psi\rangle := (-1)^{|\alpha|_1}\langle u,D^\alpha\psi\rangle$.
5. reflection $\langle \widetilde{u},\psi\rangle := \left\langle u,\widetilde{\psi}\right\rangle$.

Since the right-hand sides of each of the above equalities are well-defined for and continuous in $\psi \in \mathcal{S}$, we see that $\widehat{u}$, $u\phi$ etc., as defined, are proper tempered distributions (in particular, continuous on $\mathcal{S}$, so S^*, too, is closed under these operations. Recall that an element of the dual $\mathcal{S}^*$ is determined by its action on each element of the Schwartz space $\mathcal{S}$, justifying the way of definition used here.

The definitions above agree with the ones for functions, when $u \in \mathcal{S}^*$ is identifiable with a proper function f, i.e., $\langle u,\psi\rangle = \langle f,\psi\rangle = \int_{\mathbb{R}^d} f\psi dx$. Another important class of tempered distributions consists of those corresponding to finite Borel measures by $\langle u,\psi\rangle = \int_{\mathbb{R}^d}\psi d\mu$.

The operations on $\mathcal{S}^*$ defined above inherit many of the familiar properties of these operations on test functions, as illustrated by the following simple lemma.

Lemma A.43. *For $u \in \mathcal{S}^*$, $\phi \in \mathcal{S}$ we have $\widehat{u*\phi} = \widehat{u}\widehat{\phi}$.*

Proof. For $\psi \in \mathcal{S}$ we compute $\left\langle \widehat{u}\widehat{\phi},\psi\right\rangle = \left\langle \widehat{u},\widehat{\phi}\psi\right\rangle = \left\langle u,\mathcal{F}(\widehat{\phi}\psi)\right\rangle = \left\langle u,\mathcal{F}^2\phi*\widehat{\psi}\right\rangle = \left\langle u,\widetilde{\phi}*\widehat{\psi}\right\rangle = \left\langle u*\phi,\widehat{\psi}\right\rangle = \left\langle \widehat{u*\phi},\psi\right\rangle$. $\square$

In fact, we can say a lot more about the convolution of a distribution with a test function: the result turns out to be a rather well-behaving proper function, as demonstrated next.

Lemma A.44. *If $u \in \mathcal{S}^*$, $\phi \in \mathcal{S}$, then $u*\phi$ can be identified with a slowly increasing (in particular, C^∞ smooth) function f. The value of f at each x is given by $f(x) = \left\langle u, \tau_x \widetilde{\phi} \right\rangle$.*

Proof. Since, for $h = h_j e_j$, $\frac{1}{h_j}(\tau_{x+h}\widetilde{\phi} - \tau_x \widetilde{\phi}) \to -\tau_x D_j \widetilde{\phi}$ in $\mathcal{S}$ and u is a continuous linear functional, we see that f defined as in the assertion satisfies $\frac{1}{h_j}(f(x+h) - f(x)) = \left\langle u, \frac{1}{h_j}(\tau_{x+h}\widetilde{\phi} - \tau_x\widetilde{\phi}) \right\rangle \to - \left\langle u, \tau_x D_j \widetilde{\phi} \right\rangle$. By iteration, f has derivatives of all orders and $D^\alpha f(x) = (-1)^{|\alpha|_1} \left\langle u, \tau_x D^\alpha \widetilde{\phi} \right\rangle$. Since $D^\alpha \widetilde{\phi} \in \mathcal{S}$, we also see, from the expression for $D^\alpha f$, that it is only required to prove that $f(x) = \left\langle u, \tau_x \widetilde{\phi} \right\rangle$ is bounded by a polynomial for any $\phi \in \mathcal{S}$, and the same will follow for the derivatives.

To show that f is bounded by a polynomial, use (A.12) to get $C, N > 0$ such that $|f(x)| = \left| \left\langle u, \tau_x \widetilde{\phi} \right\rangle \right| \le C \sum_{|\alpha|_1, |\beta|_1 \le N} |\tau_x \phi|_{\alpha\beta}$. We then observe that each of the finite number of terms in the right-hand side is bounded by

$$|\tau_x \phi|_{\alpha\beta} = \sup_y \left| y^\beta D^\alpha \phi(y - x) \right| = \sup_t \left| (x+t)^\beta D^\alpha \phi(t) \right| \le \sum_{\gamma \le \beta} \binom{\beta}{\gamma} |x^\gamma| \, |\phi|_{\alpha, \beta - \gamma} ,$$

a polynomial in (the absolute values of the components of) x.

We now know that f has the asserted properties. In particular, since f is slowly increasing, we know that it can be identified with a tempered distribution. It remains to show that this distribution coincides with the convolution $u*\phi$, i.e., that $\langle u * \phi, \psi \rangle = \int_{\mathbb{R}^d} f\psi dx$ for each $\psi \in \mathcal{S}$. To this end we compute $\langle u * \phi, \psi \rangle = \left\langle u, \widetilde{\phi} * \psi \right\rangle = \left\langle u, \int_{\mathbb{R}^d} \widetilde{\phi}(\cdot - x)\psi(x)dx \right\rangle = \left\langle u, \int_{\mathbb{R}^d} (\tau_x \widetilde{\phi})\psi(x)dx \right\rangle = \int_{\mathbb{R}^d} \left\langle u, \tau_x \widetilde{\phi} \right\rangle \psi(x)dx = \int_{\mathbb{R}^d} f(x)\psi(x)dx$. Passing u inside the integral is legitimate, if we can verify that the Riemann sums of the integral converge in $\mathcal{S}$. Indeed, the linearity of u guarantees that we can bring it inside a finite sum, and the rest follows from the continuity of u on $\mathcal{S}$.

For this last task, we first note that $D^\alpha \int_{\mathbb{R}^d} \widetilde{\phi}(\cdot - x)\psi(x)dx = \int_{\mathbb{R}^d} (D^\alpha \widetilde{\phi})(\cdot - x)\psi(x)dx$; indeed, $\int_{\mathbb{R}^d} \left| \frac{1}{h_i}(\widetilde{\phi}(y + h - x) - \widetilde{\phi}(y - x)) \right| |\psi(x)| \, dx \le \int_{\mathbb{R}^d} \max \left| \frac{\partial^2 \widetilde{\phi}}{\partial y_i^2} \right| \frac{|h|}{2} |\psi(x)| \, dx \to 0$ as $|h| \to 0$, and the claim for general D^α follows by iteration. Our intention is to show that

$$\sum_i y^\beta D^\alpha \widetilde{\phi}(y - x_i)\psi(x_i)m(\Delta_i) \to \int_{\mathbb{R}^d} y^\beta D^\alpha \widetilde{\phi}(y - x)\psi(x)dx$$

as $\max m(\Delta_i) \downarrow 0$ and $\cup_i \Delta_i \uparrow \mathbb{R}^d$, where Δ_i are disjoint volumes such that $x_i \in \Delta_i$. Furthermore, this convergence should take place uniformly in $y \in \mathbb{R}^d$. This is exactly what is meant by the convergence of Riemann sums in the topology of $\mathcal{S}$.

For convenience, we denote $t := y - x$, $\varphi := D^\alpha \widetilde{\phi} \in \mathcal{S}$. Then our task is to approximate $\int_{\mathbb{R}^d} \sum_{\gamma \le \beta} \binom{\beta}{\gamma} t^{\beta-\gamma}\varphi(t)x^\gamma \psi(x)dx$ by $\sum_i \sum_{\gamma \le \beta} \binom{\beta}{\gamma} t^{\beta - \gamma}\varphi(t)x_i^\gamma \psi(x_i)m(\Delta_i)$ uniformly in $t \in \mathbb{R}^d$.

Since $\varphi \in \mathcal{S}$, $\left| t^{\beta-\gamma}\varphi(t) \right| \le C \left| t^\eta \right|$ for any η, for some C, and it follows from the fact that $x \mapsto x^\gamma \psi(x)$ is in $\mathcal{S} \subset L^1$ that both the integral and the Riemann sums above become as small as one likes for large enough $|t|$. Thus it is sufficient to establish the approximation property in an arbitrary compact set, and due to the continuity in t of both the integral and the Riemann sums, the task further reduces to approximation on the points of a finite ϵ-net.

For a fixed t, the integral is continuous and vanishes at infinity faster than the inverse of any polynomial; thus the desired convergence follows from the elementary properties of the Riemann integral. For the finite number of t in our ϵ-net, the convergence is certainly uniform.

The proof is now complete. □

A.8 Fourier series

A useful fact in the study of Fourier series is the density in $L^p(\mathbb{T}^d; X)$, $p \in [1, \infty)$, and in $C(\mathbb{T}; X)$, of the trigonometric polynomials. For $C(\mathbb{T}; X)$, this can be established as in the case $X = \mathbb{C}$

by considering convolutions with an appropriate sequence of trigonometric functions, such as the Fejér kernel
$$\frac{1}{n+1}\frac{\sin^2((n+1)\pi t)}{\sin^2(\pi t)},$$
or some other sequence providing an approximation of the identity in the limit (see e.g. [20], pages 89–91).

The case $C(\mathbb{T}^d;X)$ can be reached by induction: Assume the density of the trigonometric polynomials in $C(\mathbb{T}^{d-1};X)$. Since we already showed that they are dense in $C(\mathbb{T};X)$, for any Banach space X, we can take $C(\mathbb{T}^{d-1};X)$ in place of X to deduce the density of functions of the form
$$\sum_{k=-n}^{n} e^{\mathbf{i}2\pi k\cdot} f_k, \qquad f_k \in C(\mathbb{T}^{d-1};X),$$
on $C(\mathbb{T};C(\mathbb{T}^{d-1};X))$. Since the f_k can be approximated by trigonometric polynomials, we deduce, after identifying $C(\mathbb{T};C(\mathbb{T}^{d-1};X))$ with $C(\mathbb{T}^d;X)$, that trigonometric polynomials are also dense in $C(\mathbb{T}^d;X)$. The conclusion on $L^p(\mathbb{T}^d;X)$ follows from the density of continuous functions (Lemma A.19; this is easily modified to work on $\mathbb{T}^d$ instead of $\mathbb{R}^d$).

Before stating the first basic results concerning vector-valued Fourier series, we observe the following null-criterion similar to Corollary A.5:

Lemma A.45. *If, for $f \in L^1(\mathbb{T}^d;X)$, the integral $\int_\Omega fg\,dm$ vanishes for every $g \in C(\mathbb{T}^d)$, then $f = 0$ (a.e.).*

Proof. If the assumption is true, then $\int_\Omega \langle x^*, f\rangle\, g\,dm = 0$ for every $x^* \in X^*$, thus $\langle x^*, f(t)\rangle = 0$ for all t, except possibly those on a null-set Z_{x^*}. Taking $\{\xi_k^*\}_{k=1}^\infty$ as in Lemma A.4, we find that $Z := \cup_{k=1}^\infty Z_{\xi_k^*}$ is a null-set, and $\langle \xi_k^*, f(t)\rangle = 0$ for all ξ_k^*, $k \in \mathbb{Z}_+$, for $t \in Z^c$. By the properties of the sequence $\{\xi_k^*\}_{k=1}^\infty$, $f = 0$ on Z^c, thus almost everywhere. ☐

The following basic results then follow:

Lemma A.46. *Let $f \in L^1(\mathbb{T}^d;X)$.*

1. *If $\int_{\mathbb{T}^d} e^{\mathbf{i}2\pi\kappa\cdot t} f(t)dt = 0$ for all $\kappa \in \mathbb{Z}^d$, then $f = 0$ a.e. Consequently, no two functions have the same Fourier series.*
2. *If the Fourier series $\sum_{\kappa\in\mathbb{Z}^d} x_\kappa e^{\mathbf{i}2\pi\kappa\cdot(\cdot)}$ of f converges absolutely, then it defines a continuous functions which is equal to f (a.e.).*

Proof. 1. By linearity we have $\int_{\mathbb{T}^d} gf\,dm = 0$ for all trigonometric polynomials g on $\mathbb{T}^d$, and finally, by density, for all $g \in C(\mathbb{T}^d;X)$. Thus $f = 0$ (a.e.) by Lemma A.45.

2. By assumption, $\sum_{\kappa\in\mathbb{Z}^d} x_\kappa e^{\mathbf{i}2\pi\kappa\cdot(\cdot)}$ is the uniform limit of continuous functions (trigonometric polynomials), and thus continuous. The Fourier coefficients of $\sum_{\kappa\in\mathbb{Z}^d} e^{\mathbf{i}2\pi\kappa\cdot(\cdot)} x_\kappa$ are the same as those of f, so the functions coincide (a.e.) by part 1. ☐

Some observations regarding the connection between Fourier series and the Fourier transform are in order. If $f \in L^1(\mathbb{R}^d;X)$, then
$$\sum_{\kappa\in\mathbb{Z}^d}\int_{[0,1)^d} |f(t+\kappa)|_X\,dt = \sum_{\kappa\in\mathbb{Z}^d}\int_{[0,1)^d+\kappa} |f(t)|_X\,dt = \int_{\mathbb{R}^d} |f(t)|_X\,dt = |f|_{L^1(\mathbb{R}^d;X)} < \infty,$$
so $\sum_{\kappa\in\mathbb{Z}^d} f(\cdot+\kappa)$ converges absolutely on $L^1(\mathbb{T}^d;X)$; thus it defines a function on the torus. The Fourier coefficients of this function are given by
$$\begin{aligned}\int_{[0,1)^d}\left(\sum_{\kappa\in\mathbb{Z}^d} f(t+\kappa)\right)e^{-\mathbf{i}2\pi\nu\cdot t}dt &= \sum_{\kappa\in\mathbb{Z}^d}\int_{[0,1)^d} f(t+\kappa)e^{-\mathbf{i}2\pi\nu\cdot t}dt \\ &= \sum_{\kappa\in\mathbb{Z}^d}\int_{[0,1)^d+\kappa} f(t)e^{\mathbf{i}2\pi\nu\cdot(t-\kappa)}dt = \int_{\mathbb{R}^d} f(t)e^{-\mathbf{i}2\pi\nu\cdot t}dt = \widehat{f}(\nu).\end{aligned}$$

In the second to last step we used the fact that $e^{\mathbf{i}2\pi\nu\cdot\kappa} = 1$, since $\nu \cdot \kappa \in \mathbb{Z}$. The change of the order of summation and integration is valid due to the absolute convergence shown above.

These ideas lead to the important Poisson summation formula:

Lemma A.47 (Poisson summation formula). *For $\phi \in \mathcal{S}(\mathbb{R}^d; X)$, we have*

$$\sum_{\kappa \in \mathbb{Z}^d} \phi(t+\kappa) = \sum_{\kappa \in \mathbb{Z}^d} \widehat{\phi}(\kappa) e^{\mathbf{i}2\pi\kappa\cdot t}$$

The assumption that ϕ be rapidly decreasing is quite a bit stronger than what is required to derive the conclusion, but we only need the result in the present form. Once the lemma is proved, a slight generalization is obtained by substituting $\phi(\lambda\cdot)$ in place of ϕ (and recalling that $\mathcal{F}(\phi(\lambda\cdot)) = \lambda^{-d}\widehat{\phi}(\lambda^{-1}\cdot))$, and then $\lambda^{-1}t$ in place of t. This yields

$$\sum_{\kappa \in \mathbb{Z}^d} \phi(t+\lambda\kappa) = \lambda^{-d} \sum_{\kappa \in \mathbb{Z}^d} \widehat{\phi}(\lambda^{-1}\kappa) e^{\mathbf{i}2\pi\kappa\cdot\lambda^{-1}t}.$$

Proof. Since $\widehat{\phi}$ is rapidly decreasing, $\sum_{\kappa\in\mathbb{Z}^d} \widehat{\phi}(\kappa) e^{\mathbf{i}2\pi\kappa\cdot(\cdot)}$ converges absolutely, and so defines a continuous function, which is equal to $\sum_{\kappa\in\mathbb{Z}^d} f(\cdot+\kappa)$ (a.e.). Since this series is also absolutely convergent for $f \in \mathcal{S}(\mathbb{R}^d; X)$, the latter function is also continuous, and we actually have the equality everywhere. □

A.9 Notes and comments

The integration theory in Banach spaces (Section A.2) is developed following Neveu [16], with minor influence from de Pagter [5]. This is a little different from the standard references such as Hille and Phillips [9] and Diestel and Uhl [6]; more details on this are given below.

Section A.3 follows de Pagter [5] and Diestel and Uhl [6]. The reason for omitting the proof of Lemma A.11 is not that it would be beyond the level of our treatment; however, the proof requires a series of results from the theory of compact operators, which are certainly interesting on their own, but which would be of no other use in the present context. The material required for the proof is mostly concentrated on pages 59–76 of Diestel and Uhl [6]; some additional results concerning the Bôchner integral are also needed.

Section A.4 again follows de Pagter [5]. An important extension result omitted is the theorem of Marcinkiewicz and Zygmund (1939), which states that every $T \in \mathcal{B}(L^p(\Omega); L^q(\Omega))$, $p, q \in [1, \infty)$, can be extended to any Hilbert space $\mathcal{H}$ to yield an operator of norm at most $C_{p,q} |T|_{\mathcal{B}(L^p(\Omega);L^q(\Omega))}$. Here $C_{p,q}$ are universal constants depending only on p and q, and $C_{p,q} = 1$ if $1 \leq p \leq q < \infty$. This could be used to prove that every Hilbert space is UMD from the knowledge that $\mathbb{C}$ is UMD; it also follows that $M_p(\mathcal{H}) = M_p(\mathbb{C})$ for any Hilbert space $\mathcal{H}$.

Section A.5 is based on Rudin [20] and Spivak [22], Section A.6 on Rauch [17] and Rudin [19], and Section A.8 on Stein and Weiss [24]. These books deal with the scalar-valued theory, but the vector-valued generalizations above are rather straightforward.

The vector-valued distributions (Section A.7) are defined as they appear in Hieber and Prüss [8]. Our results are generalized to the vector-valued situation from those appearing in Rudin [19] and Stein and Weiss [24]. The classical treatment of the theory of distributions is Schwartz [21].

Notes on vector-valued integration. We shortly quote Hille and Phillips [9], pages 71–85, for alternative definitions of measurability and integration in the vector-valued setting; we also refer to this work for the proofs of the results cited in the following.

The multiplicity of notions of integration results from the various topologies available in a Banach space. One often introduces at least two distinct notions of measurability: $f : \Omega \to X$ is called *weakly measurable* if the scalar-valued functions $\langle x^*, f(\cdot)\rangle$, $x^* \in X^*$ are measurable. A *strongly measurable* function is one which is the limit a.e. of a sequence of countably-valued functions, each value of which is attained on a measurable set. Since Lemma A.1 shows that a

Borel measurable function is a limit of simple, i.e., finitely-valued, functions taking each distinct value on a measurable set, it is clear that Borel measurability implies strong measurability in separable spaces. To see the converse, observe that each countably-valued function $\sum_{k=1}^{\infty} x_k \mathbf{1}_{E_k}$, with each E_k disjoint and measurable, is Borel measurable. Indeed, the preimage of any Borel set (in fact, of any set) is the (countable) union of the preimages E_k of the x_k it contains; since each E_k is measurable, so is their countable union. Thus each strongly measurable function is the limit a.e. of Borel measurable functions, thus it is Borel measurable itself by Lemma A.2. Hence Borel measurability coincides with strong measurability at least in separable Banach spaces.

Furthermore, one can show that weak and strong measurability are also equivalent in the separable setting. In general, a function is strongly measurable if and only if it is weakly measurable and has an essentially separable range. Thus our definition of measurability (Remark A.7 and the discussion following it) also agrees with strong measurability in the non-separable case.

Several definitions of the integral can also be made. Ours is equivalent to the *Bôchner integral*: integrable functions f are precisely those that are strongly measurable and satisfy $\int_\Omega |f|_X \, d\mu < \infty$ ([9], Theorem 3.7.4, or [6], Theorem II.2.2).

Another way of defining the Bôchner integral is the following: Integrable functions are those having an expansion $f = \sum_{k=1}^{\infty} x_k \mathbf{1}_{E_k}$, with E_k measurable and $\sum_{k=1}^{\infty} x_k \mu(E_k)$ absolutely convergent. Then $\int_E f d\mu := \sum_{k=1}^{\infty} x_k \mu(E \cap E_k)$ ([6], page 55). This notion is exploited systematically in Mikusiński [15]. On $\mathbb{R}^d$, the regularity properties of the Lebesgue measure allow one to restrict to expansions where each E_k is a finite box $[a_1, b_1] \times \cdots \times [a_d, b_d]$, and we still get the same class of integrable functions.

Bibliography

[1] W. Arendt and S. Bu. The operator-valued Marcinkiewicz multiplier theorem and maximal regularity. In *Ulmer Seminare über Funktionalanalysis und Differentialgleichungen, Heft 5*, pages 36–58. 2000.

[2] D. L. Burkholder. Martingales and Fourier analysis in Banach spaces. In A. Dold and B. Eckmann, editors, *Probability and Analysis*, volume 1206 of *Lecture Notes in Mathematics*, pages 61–108. Springer-Verlag, 1985.

[3] P. Clément, B. de Pagter, F. A. Sukochev, and H. Witvliet. Shauder decompositions and multiplier theorems. *Studia Mathematica*, 138(2):135–163, 2000.

[4] V. H. de la Peña and E. Giné. *Decoupling, From Dependence to Independence.* Probability and its Applications. Springer-Verlag, New York, 1999.

[5] B. de Pagter. Martingales, UMD-spaces and the vector-valued Hilbert transform. Seminar notes, Technische Universiteit Delft, 1993–1994.

[6] J. Diestel and J. J. Uhl, Jr. *Vector Measures.* Number 15 in Mathematical Surveys. American Mathematical Society, Providence, Rhode Island, 1977.

[7] J. Duoandikoetxea. *Fourier Analysis*, volume 29 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, Rhode Island, 2001.

[8] M. Hieber and J. Prüss. Maximal regularity of parabolic problems. Monograph in preparation, 2000.

[9] E. Hille and R. S. Phillips. *Functional Analysis and Semi-Groups*, volume 31 of *Colloquium publications.* American Mathematical Society, Providence, Rhode Island, 1957.

[10] R. C. James. Weak compactness and reflexivity. *Israel Journal of Mathematics*, 2:101–119, June 1964.

[11] R. C. James. Bases in Banach spaces. *American Mathematical Monthly*, 89(9):625–640, November 1982.

[12] J.-P. Kahane. Sur les sommes vectorielles $\sum \pm u_n$. *Comptes Rendus de l'Academie des Sciences*, 259:2577–2580, octobre 1964.

[13] R. Latała and K. Oleszkiewicz. On the best possible constant in the Khintchine–Kahane inequality. *Studia Mathematica*, 109:101–104, 1994.

[14] T. R. McConnell. On Fourier multiplier transformations of Banach-valued functions. *Transactions of the American Mathematical Society*, 285(2):739–757, October 1984.

[15] J. Mikusiński. *The Bochner Integral*, volume 55 of *Matematische Reihe.* Birkhäuser, Basel, Stuttgart, 1978.

[16] J. Neveu. *Discrete Parameter Martingales*, volume 10 of *North-Holland Mathematical Library*. North-Holland Publishing Company, 1975.

[17] J. Rauch. *Partial Differential Equations*, volume 128 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991.

[18] J. L. Rubio de Francia. Martingale and integral transforms of Banach space valued functions. In *Probability in Banach spaces*, volume 1221 of *Lecture Notes in Mathematics*, pages 195–222. Springer-Verlag, 1986.

[19] W. Rudin. *Functional Analysis*. McGraw–Hill series in higher mathematics. McGraw–Hill, 1973.

[20] W. Rudin. *Real and Complex Analysis*. McGraw–Hill, Singapore, 3rd edition, 1987.

[21] L. Schwartz. *Théorie des distributions*. Publications de l'Institute de Mathématique de l'Université de Strasbourg. Hermann, Paris, 1966.

[22] M. Spivak. *Calculus on Manifolds*. Mathematics monograph series. W. A. Benjamin, Inc., New York, 1965.

[23] E. M. Stein. *Singular Integrals and Differentiability Properties of Functions*. Number 30 in Princeton Mathematical Series. Princeton University Press, Princeton, New Jersey, 1970.

[24] E. M. Stein and G. Weiss. *Introduction to Fourier Analysis on Euclidean Spaces*. Number 32 in Princeton Mathematical Series. Princeton University Press, Princeton, New Jersey, 2nd edition, 1975.

[25] L. Weis. A new approach to maximal L_p-regularity. In G. Lumer and L. Weis, editors, *Proceedings of the "6th International Conference on Evolution Equations and their Applications in Physical and Life Sciences" in Bad Herrenalb 1998*, volume 215 of *Lecture Notes in Pure and Applied Mathematics*. Marcel Dekker, November 2000.

[26] L. Weis. Operator-valued Fourier multiplier theorems and maximal L_p-regularity. Preprintreihe der Fakultät für Mathematik, Universität Karlsruhe, 2000.

[27] D. Williams. *Probability with Martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, 1991.

[28] H. Witvliet. *Unconditional Schauder decompositions and multiplier theorems*. PhD thesis, Technische Universiteit Delft, November 2000.

[29] F. Zimmermann. On vector-valued Fourier multiplier theorems. *Studia Mathematica*, 93(3):201–222, 1989.

ISBN 951-22-5477-8
ISSN 0784-6460
Espoo, 2001